Collaborations in polar research: a bibliometric study

Bev Ager, bevger@bas.ac.uk
Andrew Gray, andrew.gray@ucl.ac.uk
Polar Libraries Colloquy, June 2022

Abstract

This paper reviews recent developments in bibliometric studies of polar research, building on papers presented to the 2016 and 2018 Colloquies. It extends existing work to consider the patterns of collaboration between countries, institutes, and individuals, with an emphasis on Antarctic research. It also considers the way in which material is reused and synthesised for major international collaborations, in particular the way polar literature on climate change is brought into the United Nations Intergovernmental Panel on Climate Change (IPCC) reports that are used to inform the current debate on climate change and government policy.

Part 1 - polar bibliometrics, 2012-21

Defining polar bibliometrics

A key requirement for any bibliometric analysis is to identify the set of papers to be examined. This is particularly challenging for polar science, which is a broad, fuzzily-defined, and heterogenous field cutting across many formal disciplines, and only really defined by the geographic scope of its subject. Gray (2019), to which this paper is essentially an update, summarised the pros and cons of the different approaches available here as of 2018 - manual selection via topical databases; analysis of "polar journals"; analysis by affiliation to a "polar institution"; topic or subject indexing within a broader database; and finally conventional keyword analysis.

In the four years since that paper, the fourth option (topic indexing) has had some minor developments, but is still sadly not practically useful for identifying polar research. The "topics of prominence" used in Scopus - algorithmically clustered groups of papers - continue to be available within the SciVal analysis tool and are loosely referred to within Scopus, but are not searchable there. An initial analysis suggests that some of these can potentially be mapped to polar research, although not exhaustively. Web of Science does not have a direct equivalent, but has introduced "citation topics" within the InCites analysis tool, again calculated by algorithmic citation clustering; however, these are grouped at a higher level and as such no polar science topics can be distinguished.

One other development to note is that in 2019, the High North Research Documents project was relaunched as Open Polar (openpolar.no), a thematic database for all open access polar research. The restriction to just open access material means that it was not a suitable database to draw from in this analysis, but it will be looked at in further research.

As with the earlier analysis, we focused on keyword searching. This is the most commonly used method for bibliometric analysis of polar research, and probably the most effective and reliable. The search term used was the same as in Gray (2019), building on keywords presented in Gray & Hughes

(2016) for Antarctic papers and in Aksnes & Hessen (2009) for Arctic papers. In Web of Science syntax, the two searches were:

```
TS=((antarc* NOT (candida OR "except antarctica" OR "not antarctica"
OR "other than Antarctica")) OR "transantarctic" OR "ross sea" OR
"amundsen sea" OR "weddell sea" OR "southern ocean")
```

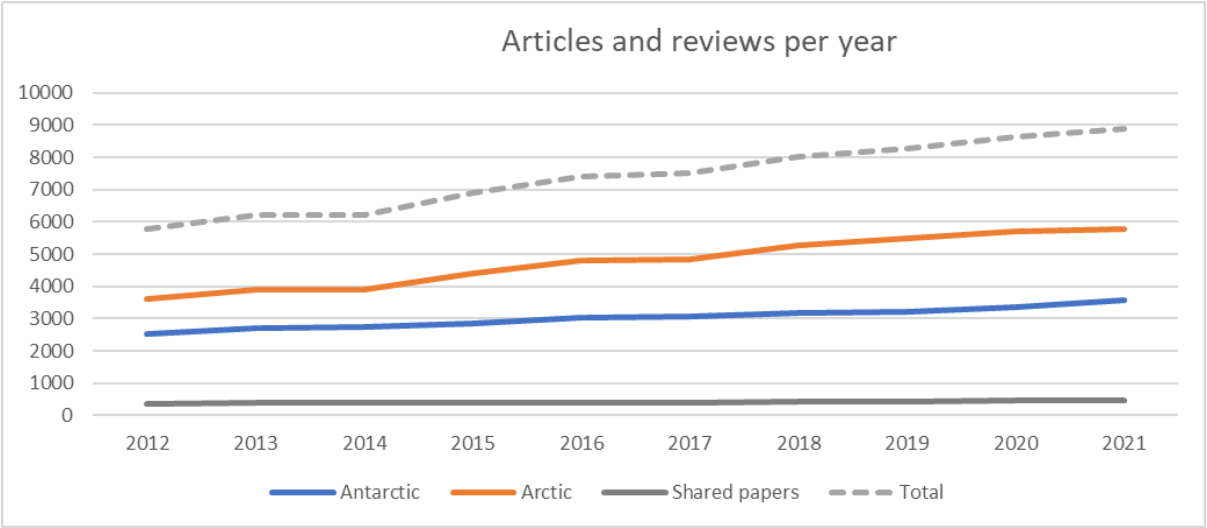
```
TS=(("Arctic" NOT "arctic bramble" NOT "sub-Arctic") OR "Svalbard" OR
"Spitsbergen" OR "Longyearbyen" OR "Ny-Alesund" OR "Hornsund" OR
"Barentsburg" OR "Kongsfjord" OR "Hopen" OR "Bjornoya" OR "Bear
Island" OR "Greenland" OR "Baffin Island" OR "Queen Elizabeth Islands"
OR "Ellesmere Island" OR "Devon Island" OR "Somerset Island" OR
"Prince of Wales Island" OR "Banks Island" OR "Ellef Ringnes Island"
OR "Amund Ringnes Island" OR "Bathurst Island" OR "Axel Heiberg
Island" OR "Prince Patrick Island" OR "King William Island" OR "Prince
Charles Island" OR "Bylot Island" OR "Bathurst Island" OR "Southampton
Island" OR "Brooks Range" OR "St Lawrence Island" OR "St Matthew
Island" OR "Seward Peninsula" OR "Nunivak Island" OR "Novaya Zemlja"
OR "Severnaja Zemlja" OR "Novosibirskije Ostrova" OR "Jan Mayen" OR
"Victoria Islands" OR "Nunavut" OR "Fram Strait" OR "Beaufort Sea" OR
"Davis Strait" OR "Barents Sea" OR "Kara Sea" OR "Storfjorden" OR
"Baffin" OR "Hudson Bay" OR "Siberian Sea" OR "Laptev Sea" OR
"ChukchiSea" OR "Bering Strait" OR "Bering Sea" OR "Karskoje Sea")
```

As can be seen, the larger Arctic search reflects the much more fragmented nature of Arctic geography. We did not do a distinct keyword search for "polar" as it seemed unlikely that many papers could be identified which discussed the polar regions and did not also use one of the above keywords.

All data was gathered from Web of Science with one of the two search terms above, filtered to only articles or reviews, published between 2012 and 2021 inclusive. Initial assessment and data analysis was carried out through InCites; this caused a small number of papers to be omitted as the databases used for Web of Science and InCites are not completely identical. This gave us a total set of 47,910 Arctic papers and 30,347 Antarctic. 4,162 papers were identified as belonging to both sets, giving an overall deduplicated total of 74,095 papers across ten years.

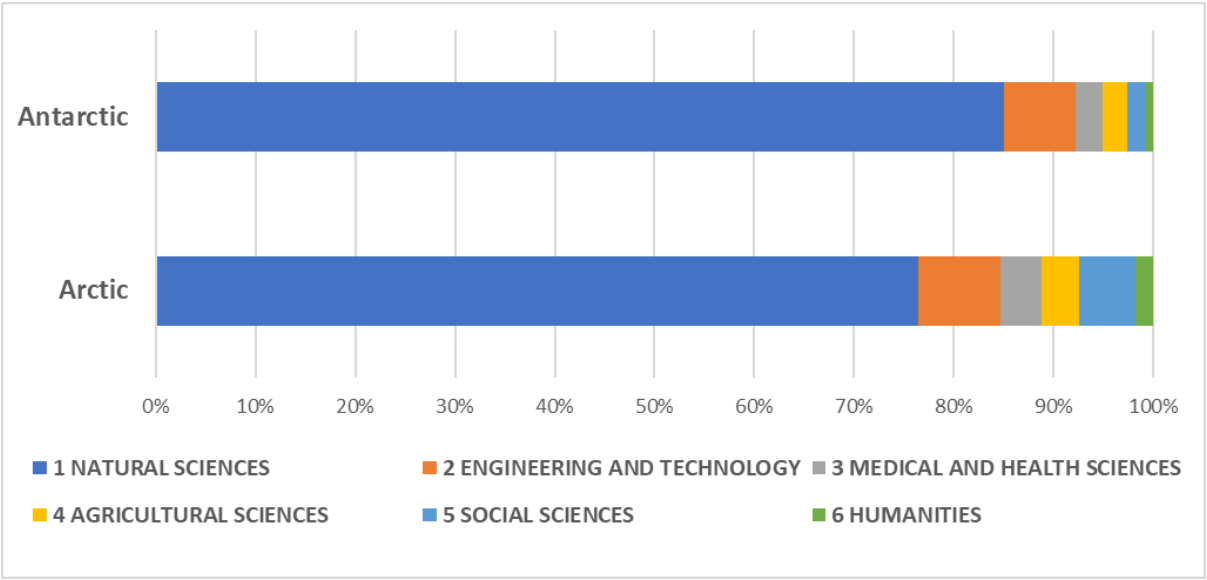
Yearly changes

The total output per year continues to rise, at a rate of around 4% per year for Antarctic research and 5.5% per year for Arctic research. However, both can be sporadic, with periods of faster growth and then retrenchment. This can be compared to the overall increase in scientific activity (estimated at 8-9% per year; Bornmann & Mutz 2015). The volume of research which is identifiably "bipolar" is growing at a slightly slower rate, around 2.5% per year, and so is making up a progressively smaller proportion of the total over time.



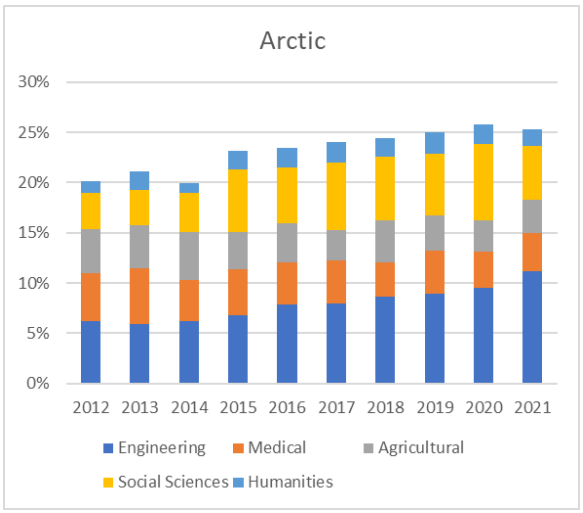
Distribution by field

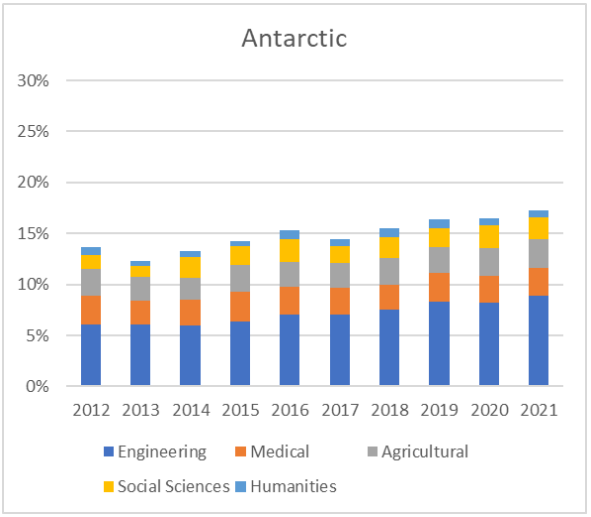
Here, we have used the OECD classification schema, breaking research down into six broad categories and then subdividing those. Papers are assigned by inference from the journal in which they are published, which while potentially confusing in some cases, is reasonably reliable at this broad a scale.



Both areas are dominated by the natural sciences (biology and the physical sciences, with engineering and technology in a distant second place. The residue (around 8% for the Antarctic, 15% for the Arctic) varies - most notably, in the Arctic, a sizable proportion is given over to the social sciences, while this is less prominent in the Antarctic. The surprising volume for "agricultural sciences" is likely due to this classification including fisheries.

A detailed breakdown of the topics making up the six individual fields is shown in the appendix.

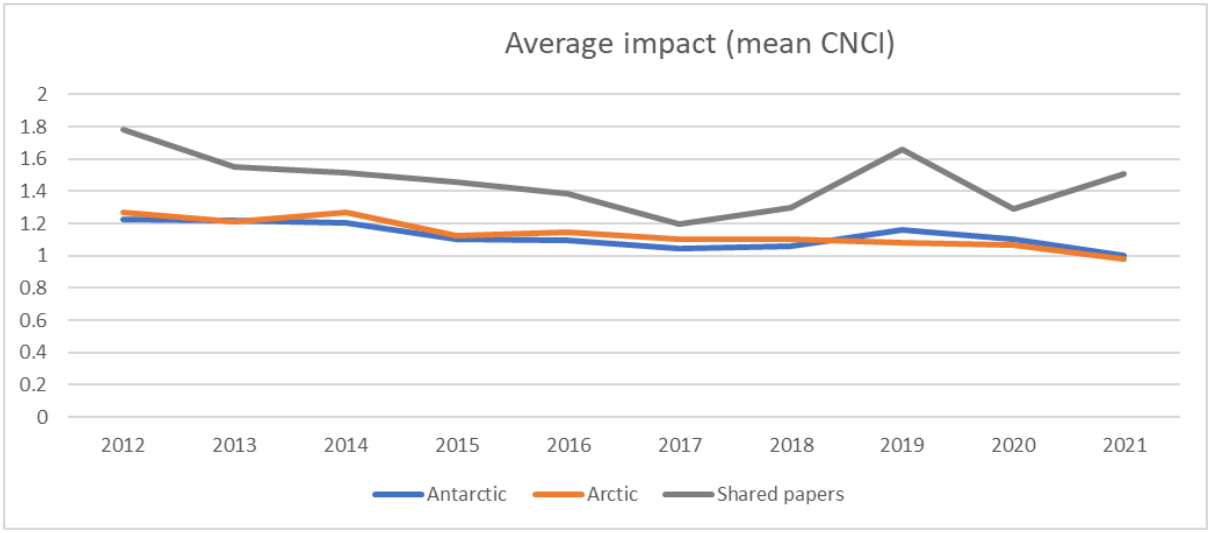




Looking at the development over time, we can see some differences between fields. In both regions, the share of fields other than natural sciences has steadily increased, with this change being driven primarily by engineering research in both cases. In the Arctic, the social sciences are tending to increase and agricultural research tending to decrease.

Relative impact

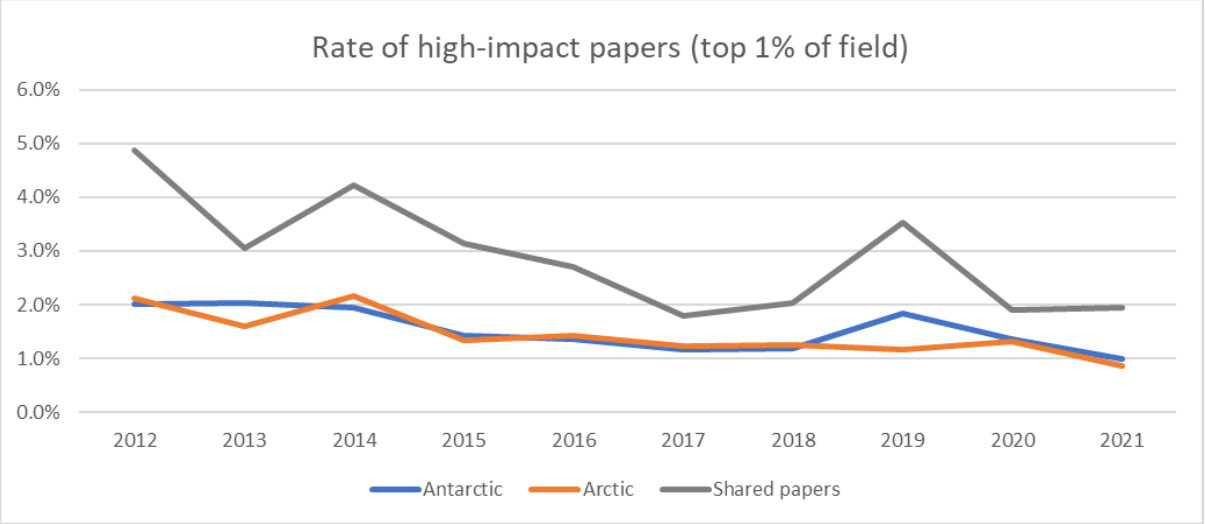
We can assess the relative impact of papers by looking at their normalised citation rate. This compares the average number of citations a paper has received to those of a similar age, document type (article, review, etc), and subject area. Here, we have used the Web of Science subject areas rather than the OECD groupings; these are somewhat more granular, with papers divided into approximately 250 topics such as “oceanography”, “medicinal chemistry”, or “language and linguistics”.



It is immediately evident that the shared papers are substantially higher cited than those which address just one of the regions, while the relative citation rates for Arctic and Antarctic papers are generally consistent. However, there is also a decline - Arctic and Antarctic research has decreased from around 1.2x the world average in 2012-14 down to approximately equal to the world average. Shared papers are somewhat more volatile, but there is an overall downwards trend at a similar rate.

Another way of assessing impact, rather than focusing on the average - which can be distorted by outliers - is to look at the rate of papers which are very highly cited. Here, we consider the papers

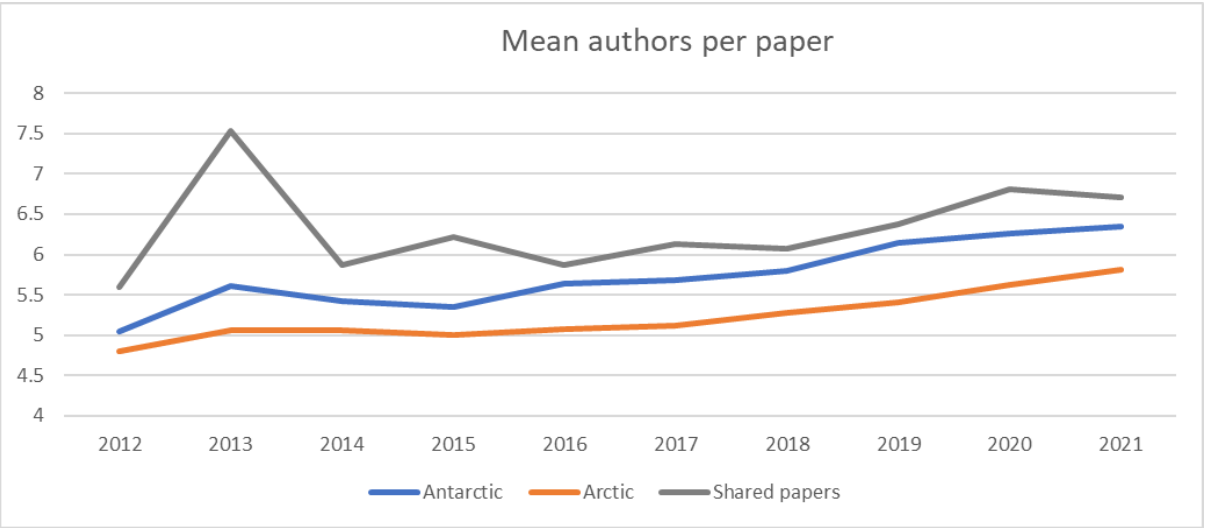
which have a citation rate that would place them in the top 1% of comparable papers, normalised as above.



The general patterns are quite similar. We see that Arctic and Antarctic papers start out above the world average (2% of their papers are in the global top 1%) but gently decline to around average by 2021. The shared papers have a higher citation rate than the regional ones, but again that rate drops consistently over time.

Collaborative papers

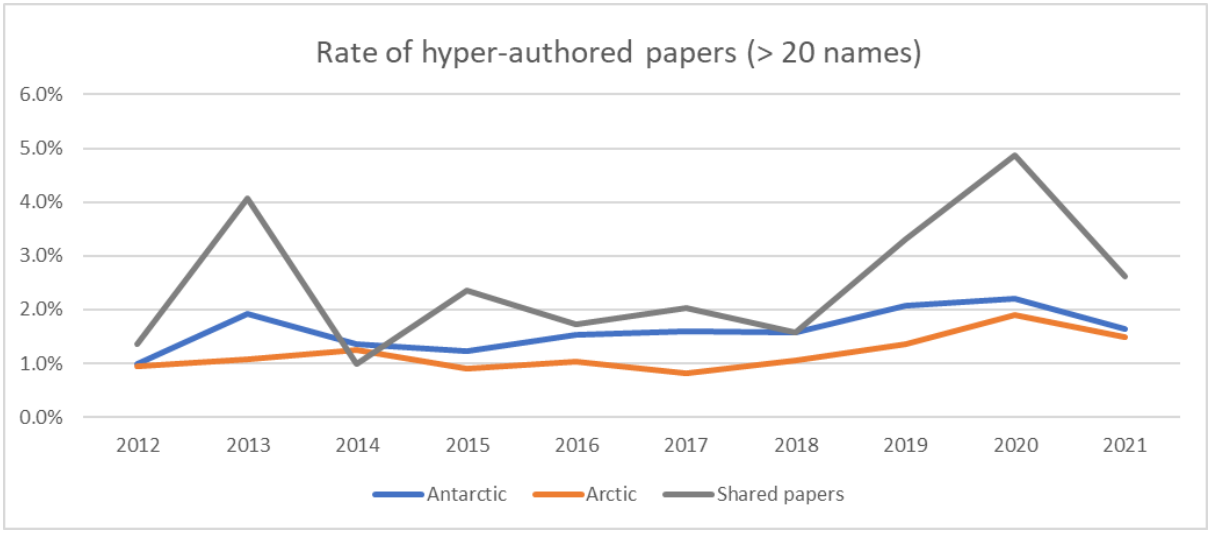
Another potentially interesting angle to consider is that of collaboration. There is a general trend to increase the level of collaboration in many fields of science, and polar research is no exception.



All three groups of papers showed a marked increase in collaboration rates over the ten years, with Arctic and Antarctic papers going from around five authors on average to around six - a 20% increase. Joint papers had a slightly higher level of authorship but the same trend was visible. There was also a consistent difference between Arctic and Antarctic papers, with the latter having slightly longer author lists by around 0.5 authors.

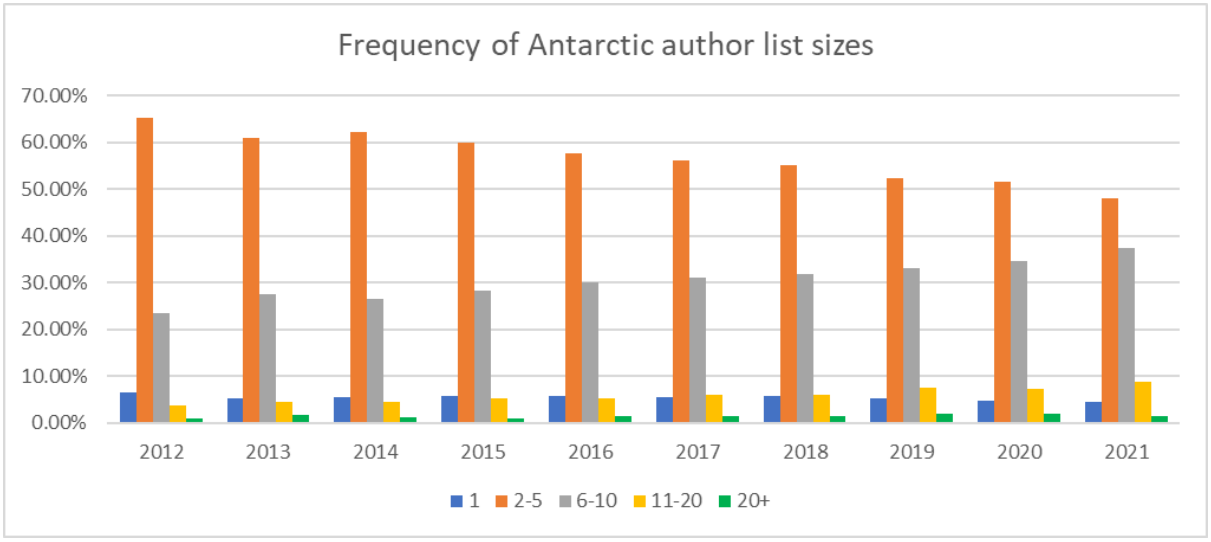
A similarly growing phenomenon is that of massively authored papers, with very long author lists. This is most common in subjects such as high-energy physics, with massive research teams of

experimenters and theoreticians leading to hundreds of names on a paper, but is growing in many other fields. In some areas of medicine, it is now common to see fifty or more researchers named.

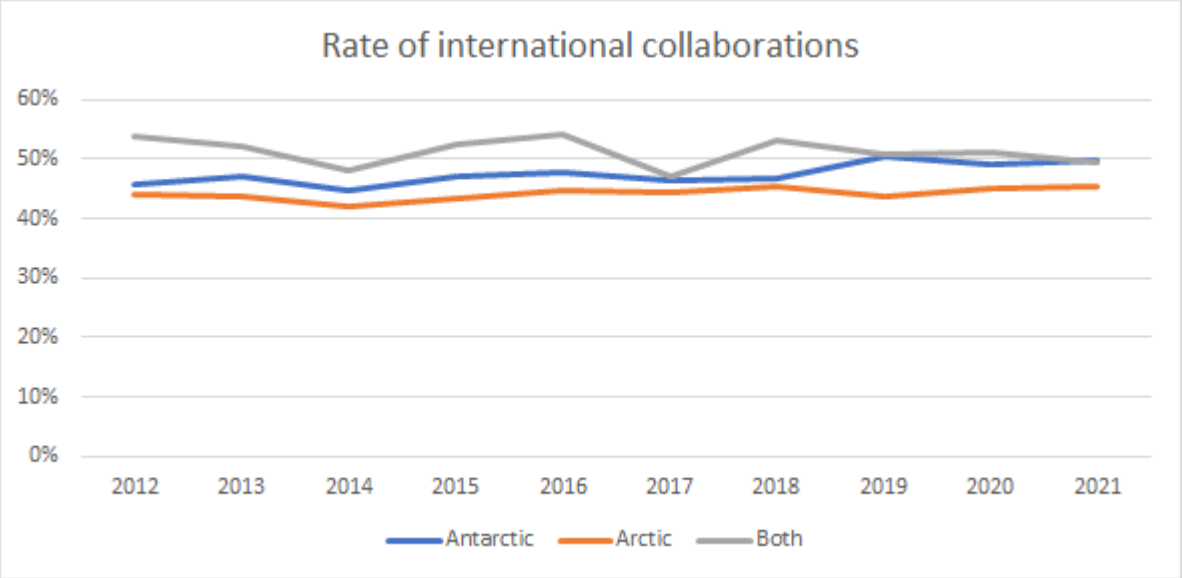


What constitutes "hyper-authorship" is very variable, but setting a lower threshold at an arbitrary but reasonable twenty papers, we find that there is indeed a slow growth over time. Both Arctic and Antarctic papers grow from around 1% to 1.5-2% having > 20 named authors. Again, the Antarctic papers show a tendency to longer author lists. Shared papers are substantially more volatile, but tend to have a higher rate of hyper-authorship than the others.

We can also look at the relative frequency of author lists in various ranges. The data suggests a slow shift to slightly larger research groups, rather than a significant change in very large groups or solo authorships.



We can see that the change is heavily driven by a rise in papers with 6-10 authors at the expense of those with 2-5 authors. Papers with 11+ authors are growing but at a relatively slower rate, and those with 20+ remain very rare. Single-authored papers are decreasing slowly, but were always relatively uncommon.



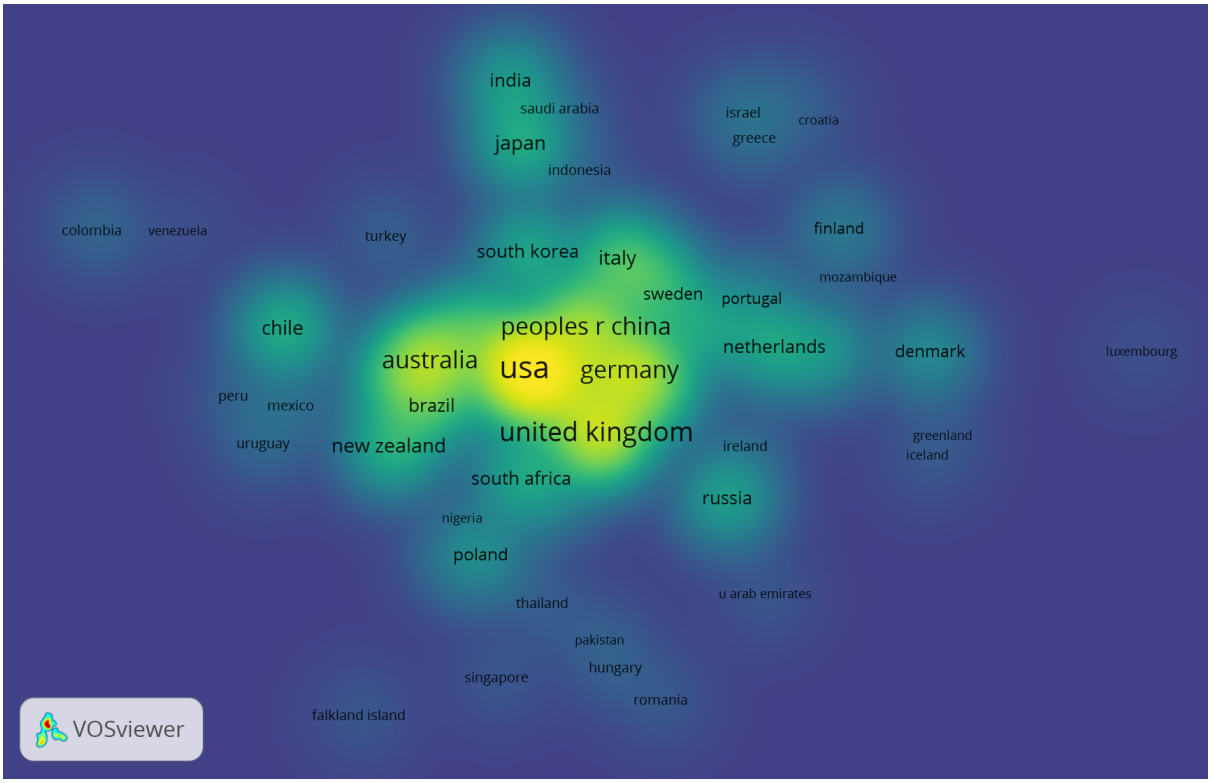
The frequency of international collaborations - papers with authors from at least two countries - is high in all three groups, around 50%, though slightly higher for Antarctic papers than Arctic ones, and as expected slightly higher still for joint papers.

Part 2 - collaborations

International collaborations & special relationships

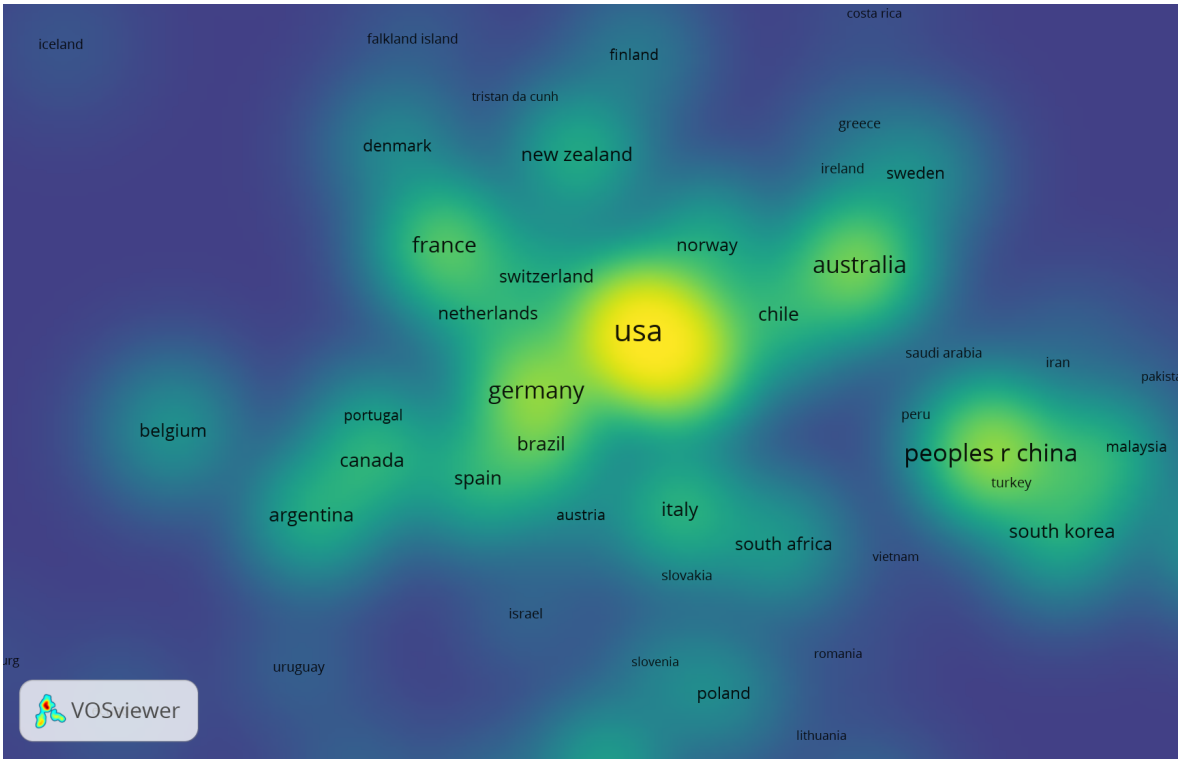
International collaborations, i.e. papers with at least one other country affiliation, are highly important in polar science. In order to analyse how these collaborations may develop or change over this same period, annual datasets of global Antarctic and Arctic research were generated by exporting full Web of Science records using the search strings and filters described above.

As full Web of Science records contain country data within the author affiliation field, it was relatively straightforward to then analyse international collaborations based on co-occurrences of countries on a paper. The example heat-density map for Antarctica 2018-2019 below was created using VOSViewer – a free software tool used to create visualisations of networks of bibliometric data.

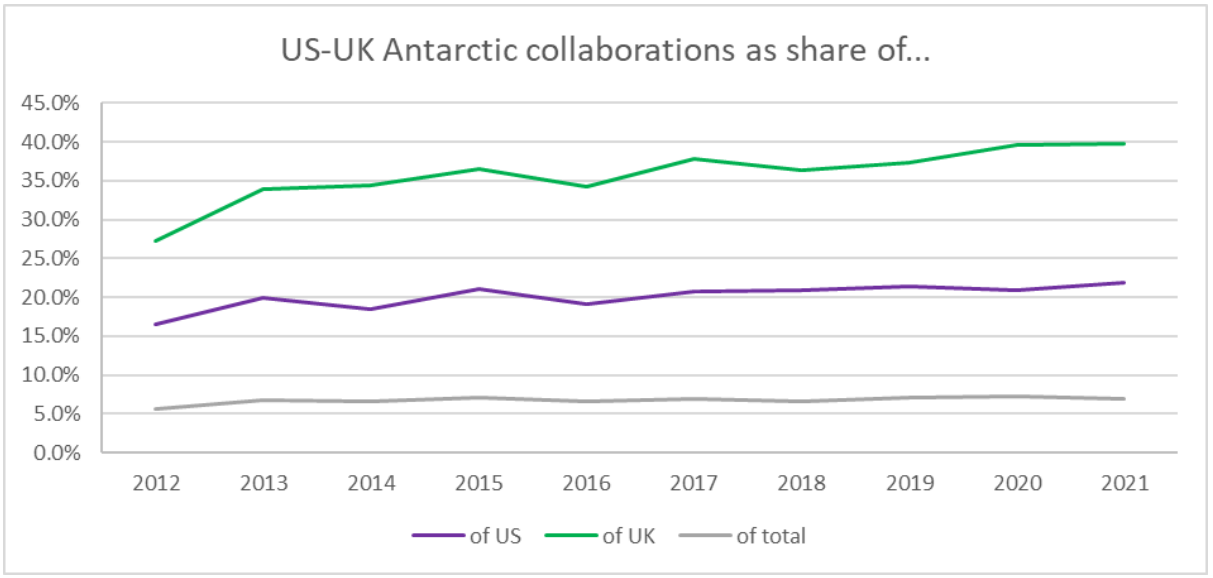


Broadly speaking, the bigger and brighter the heat-spot, the higher the volume of papers, and the more closely connected the cluster, the stronger the collaborative relationship. At the centre is a tightly knit grouping of collaborations, primarily dominated by long-standing Antarctic Treaty members - USA, Australia, UK, and Germany. Additionally, non-Treaty countries, such as China, which continues to maintain a high-level of Antarctic output (Gray 2019). Situated around the periphery are some of the Arctic countries - Canada, Norway, and Sweden. This is likely indicative of research that is bi-polar than specifically Antarctic-focussed. There are also clear clusters based on language, such as the grouping of Spanish-speaking Central and South American countries.

Consistently, across the years studied, the US-UK constitutes a key Antarctic collaborative partnership, to the extent that in the most recent 2020-21 heat-map, shown below, the UK is no longer visible, subsumed by its proximity to the larger partner, the US.

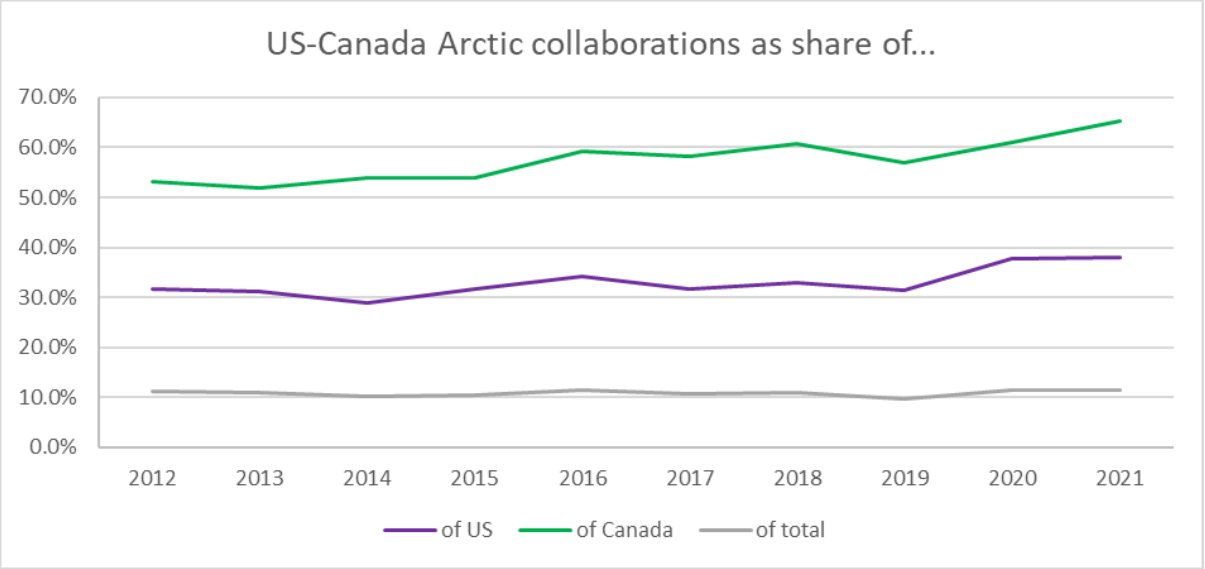


Clearly, not all partnerships are created equal and its relative importance to the respective party will vary. One way of viewing this would be by seeing what proportion of overall output a particular collaboration represents for each country. Again, focussing on the US-UK relationship, the graph below shows that US collaborations made up between 28 and 41% of total UK Antarctic output between 2012 and 2021. The trend shows a steady incline over years, with US collaborations making up a growing proportion of total UK Antarctic output. For the US, UK collaborations constitute a more modest 16-21% of their total Antarctic output. Overall, for this period, US-UK collaborations represent on average 6.9% of total global Antarctic output.



The same method was followed for Arctic papers and a key collaboration that stood out - unsurprisingly, given their geographical proximity and shared language - was the US and Canada. As shown in the graph below, papers that include a US collaboration represent between 50 and 65% of

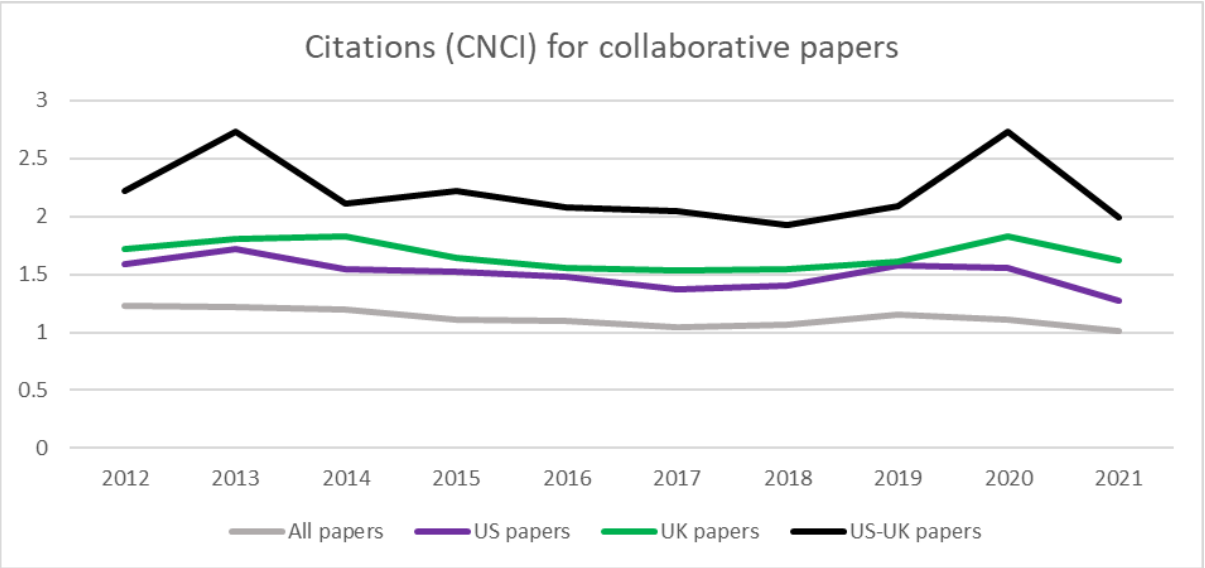
Canadian Arctic output. For the US, this was a more conservative 30-38% of total US Arctic output. Across the whole body of Arctic research, US-Canada papers represent around 10%.



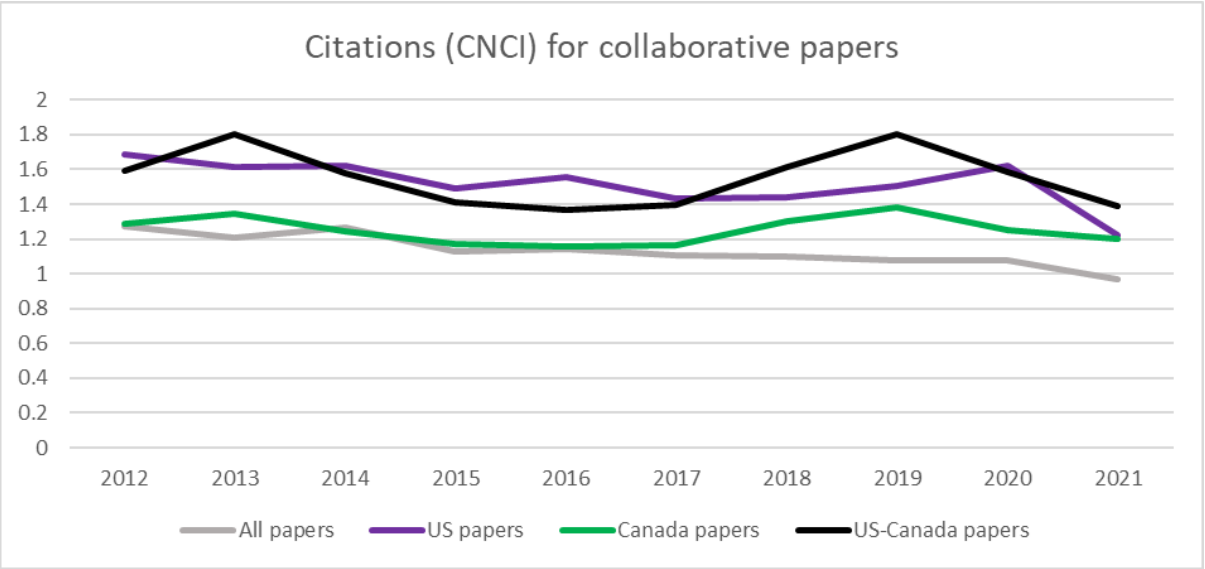
Relative impact of collaborative papers

In general, papers with increased collaboration - increased numbers of authors, or increased numbers of countries represented - tend to have higher overall rates of citation. The mechanisms for this are not always clear and may vary by field. (Thelwall & Maflahi, 2020)

Here, we assess the citation impact of polar papers, using the InCites category-normalised citation impact (CNCI) measure. Antarctic papers across all Web of Science research topics have an averaged CNCI of 1.13 for the period. As shown in the line-graph below, both the US and UK papers perform similarly. However, the combined impact of US-UK papers is significantly higher with clear peaks in 2013 and 2020. Overall US papers have an average CNCI of 1.63, 1.77 for the UK. However, US-UK collaborations have an average of 2.21 - i.e. papers carrying US-UK affiliations receive more than twice as many citations as the baseline, which demonstrates the value of these collaborations.



The pattern is markedly different for Arctic US-Canada papers. As shown in the CNCI line graph below, there is a slight uplift for the US in joint collaborations with Canada - US-Canadian papers have an averaged CNCI of 1.89, as opposed to US Arctic papers, which have an averaged CNCI of 1.71. For Canada up until 2017, the CNCI roughly follows the baseline for all Arctic papers, with an averaged CNCI across the period of 1.55, so there is more of a marked increase for combined US-Canada papers.



Part 3 - polar research in the IPCC

About the IPCC

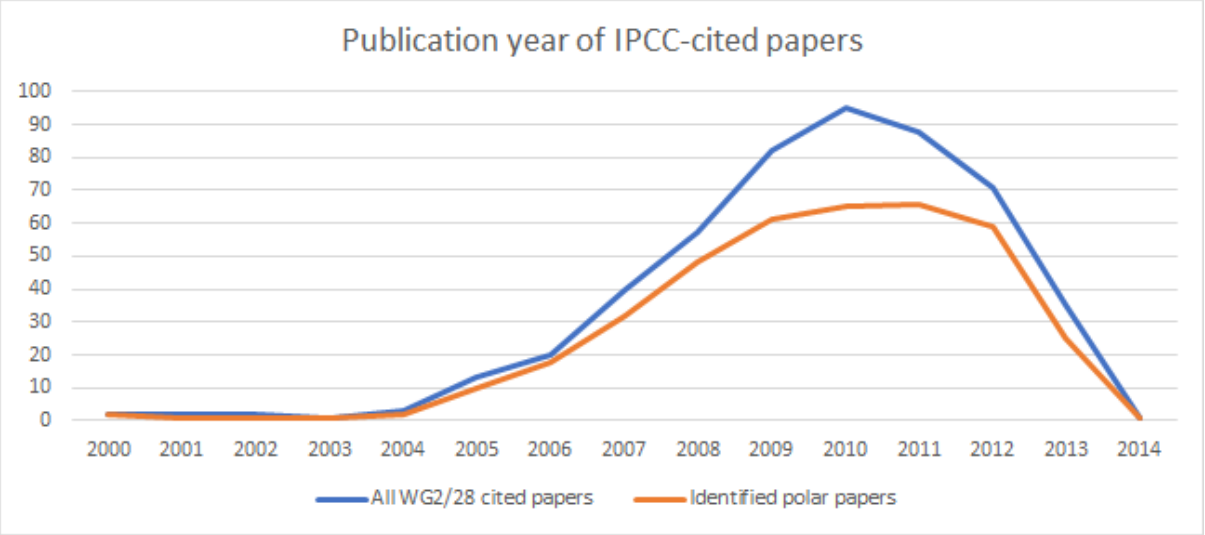
The Intergovernmental Panel on Climate Change (IPCC) is the United Nations body for assessing scientific research related to climate change. The Working Group reports produced by an international panel of researchers provide Member Governments and observer organisations with the consensus on climate change and are used to underpin international and national climate policies. Inclusion within these reports could therefore be used as a measure of research quality, i.e. the international standing or perceived authority by peers.

As part of the Fifth Report (2014), Working Group 2 looked at the effects of climate change, with Chapter 28 focussing specifically on the polar regions (Larsen et al, 2014). The references from this chapter were extracted for analysis, and out of 620 citations 512 (83%) were picked up by the Antarctic and Arctic search strings for analysis. The remaining references could generally be categorised as either non-polar specific - literature more broadly related to climate change - or not in publications indexed by Web of Science.

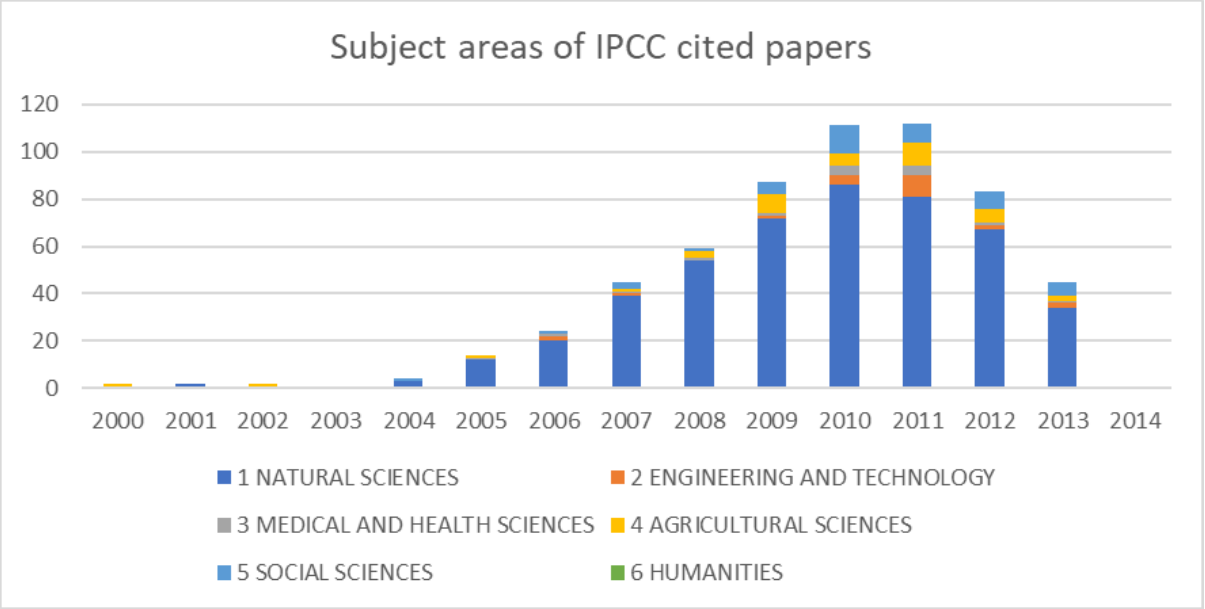
Profile of papers

A preliminary trawl of the citations by publication year also showed that those included in the Polar Regions chapter did not fit the 2012-21 period analysed. As indicated by the graph below, the majority of papers cited were from 2010, four years prior to the report's publication date. In all, roughly two-thirds of chapter citations were published before 2010. This would suggest that although climate change is a rapidly evolving topic, the IPCC reports give more weight to well-established research over the latest findings, which would be less highly cited. The majority of

polar-specific citations were published prior to 2012, leaving a considerable body of polar research that could have been cited that was not.



The OECD topic areas of the Polar Regions chapter again highlights that most papers fit broadly within Natural Sciences, with the breakdown mirroring the profile of Antarctic and Arctic papers, as shown in the Appendix: Earth and related environmental sciences forming the largest sub-topic, followed by Biological sciences. Agricultural sciences forms the next largest topic overall - as mentioned previously, this includes fisheries as a sub-topic. Social Sciences is primarily made up of the sub-topic Social and economic geography, which explores the impact of climate change on human populations.

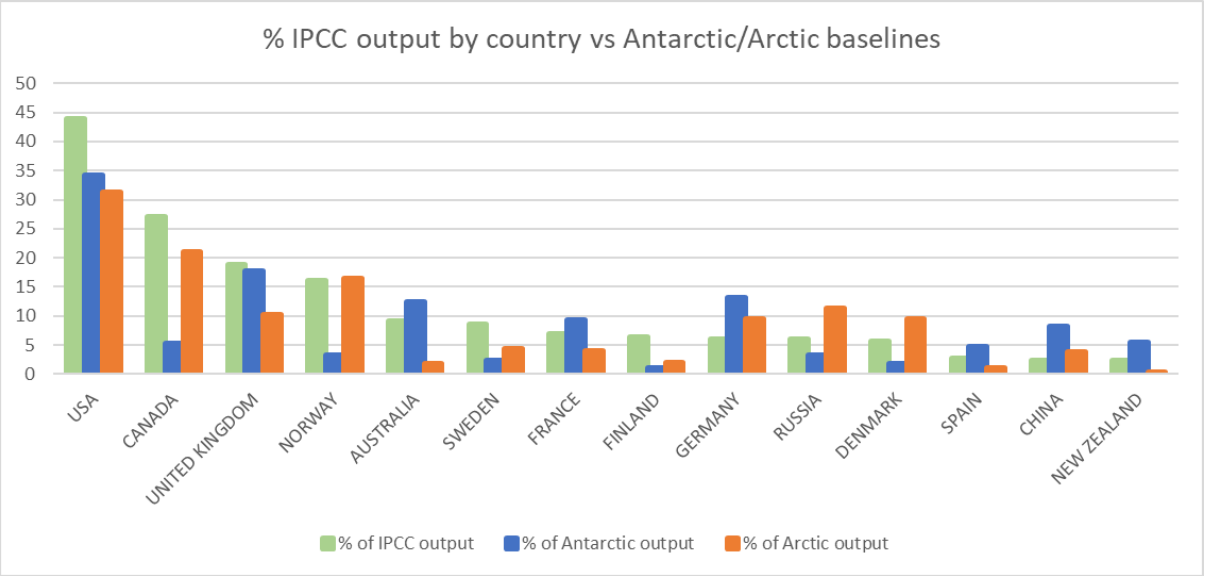


In terms of article types, the Polar Regions chapter appears to favour review papers or research syntheses over original research articles. Reviews made up a significant proportion of the polar citations - ~12% - as opposed to ~4% of all polar research produced over the 2000-2014 period. Papers averaged 5.5 authors, compared to an average of 5.3 for Arctic and 5.8 for Antarctic papers

2012-21, but it should be noted that there is a long-term trend for the number of authors to increase.

It might be expected that given the international standing of the IPCC and the extra visibility of international papers, that the papers cited here would be more international than average. However, international collaboration rates were lower than might have been anticipated: 42% of papers were international collaborations, slightly below the overall average of 45-48% for Arctic & Antarctic papers 2012-21.

The papers were, however, very highly cited; a CNCI of 3.5 and 12% in the top 1% of their field, compared to the 1-1.2 and 1-2% for Arctic and Antarctic papers in the later sample. It strongly suggests that the material cited by the IPCC is among the most visible research.



Further research

This paper is primarily an update to earlier research, but the focus on collaborations offers some interesting potential for further analysis. One approach would be to compare bilateral partnerships to multilateral collaborations (cf Adam & Gurney 2018). This could potentially show whether, for example, Antarctic papers that carry only UK and US affiliations perform better than those with multiple country affiliations. Another area of interest, but perhaps more difficult to achieve, would be to explore whether there is an increase over time in the number of individual countries per paper - i.e. is polar research becoming more geographically diverse?

The analysis of the IPCC Polar Regions chapter literature was limited as they largely fell outside of the 2012-2021 parameters of the earlier part of this study. As the 2022 IPCC Sixth Assessment Report is now available, a follow-up would establish whether the profile of literature selected for report inclusion continues to follow the same pattern, i.e. favouring older, highly cited research etc.

References

Adams & Gurney (2018) *Bilateral and Multilateral Coauthorship and Citation Impact: Patterns in UK and US International Collaboration.* Frontiers in Research Metrics and Analysis. 3:12. https://doi.org/10.3389/frma.2018.00012

Aksnes & Hessen (2009). *The Structure and Development of Polar Research (1981–2007): a Publication-Based Approach*. Arctic, Antarctic, and Alpine Research, 41:2, 155-163, https://doi.org/10.1657/1938-4246-41.2.155

Bornmann & Mutz (2015) *Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references*. Journal of the Association for Information Science & Technology, 66(11), 2215-2222 https://onlinelibrary.wiley.com/doi/abs/10.1002/asi.23329

Gray (2019) *Bibliometrics for studying polar research*. In: Hallikainen, Liisa; Parikka, Susanna, (eds.) Polar Libraries Colloquy 2018, Proceedings. Developing polar networks: Ideas and possibilities for the future. Rovaniemi, University of Lapland. https://lauda.ulapland.fi/handle/10024/63992

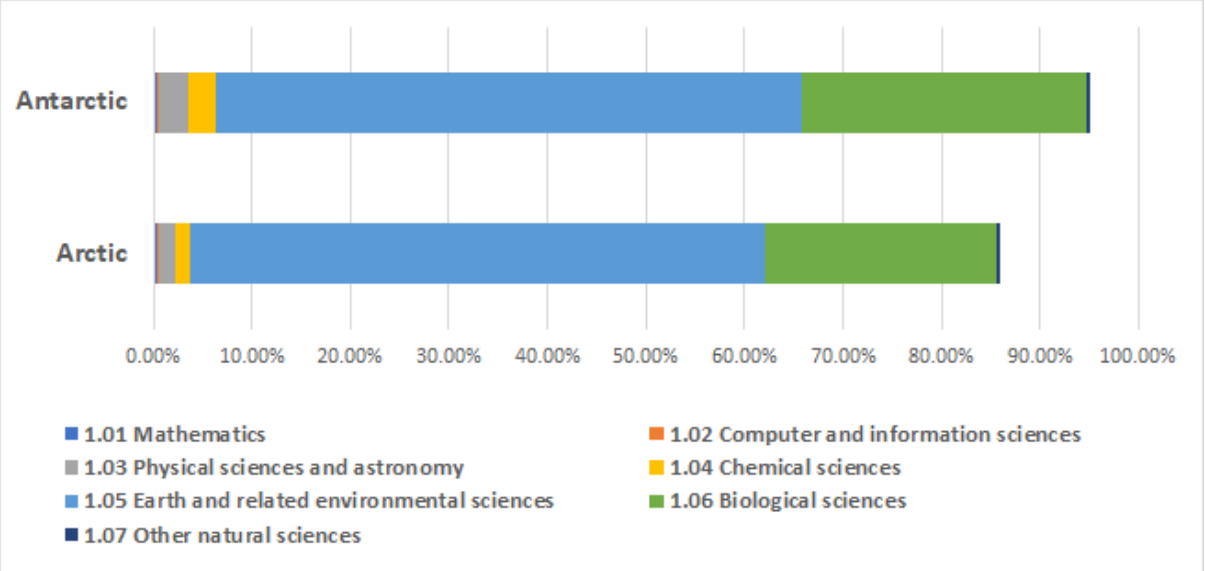
Gray & Hughes (2016). *Demonstration of "substantial research activity" to acquire consultative status under the Antarctic Treaty*. Polar Research, 35, 34061. https://doi.org/10.3402/polar.v35.34061

Larsen et al (2014). *Polar regions*. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Barros et al (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1567-1612

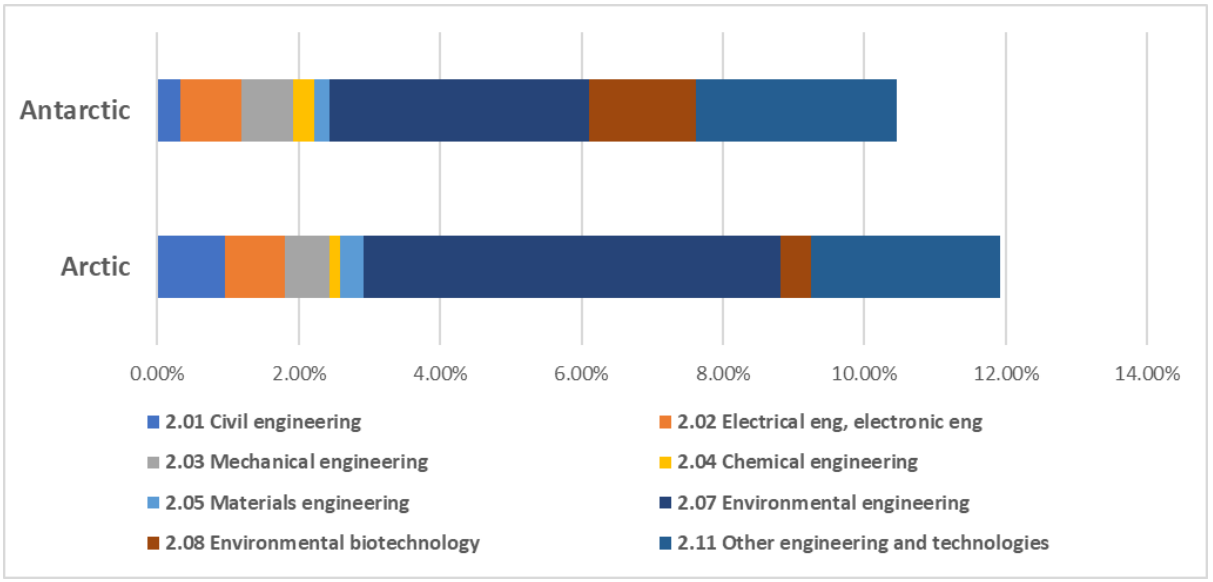
Thelwall & Maflahi (2020). *Academic collaboration rates and citation associations vary substantially between countries and fields.* Journal of the Association for Information Science & Technology, 71(8), 968-978. https://doi.org/10.1002/asi.24315

Appendix

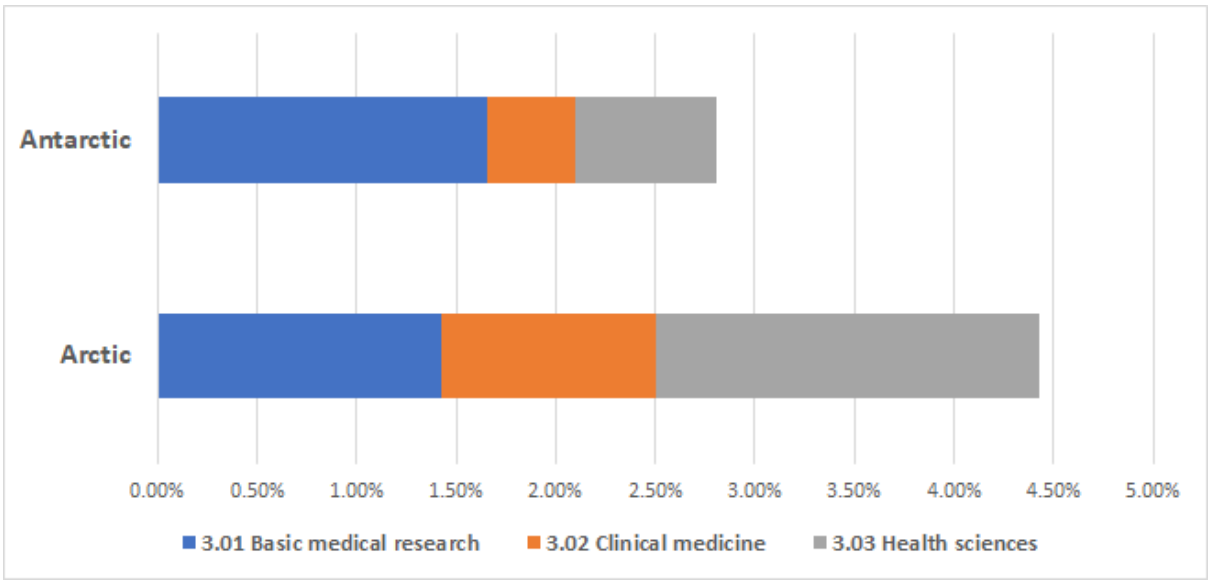
The relative breakdowns of research by field in the major subject areas are shown below. The percentages listed are the share of total research across 2012-2021.



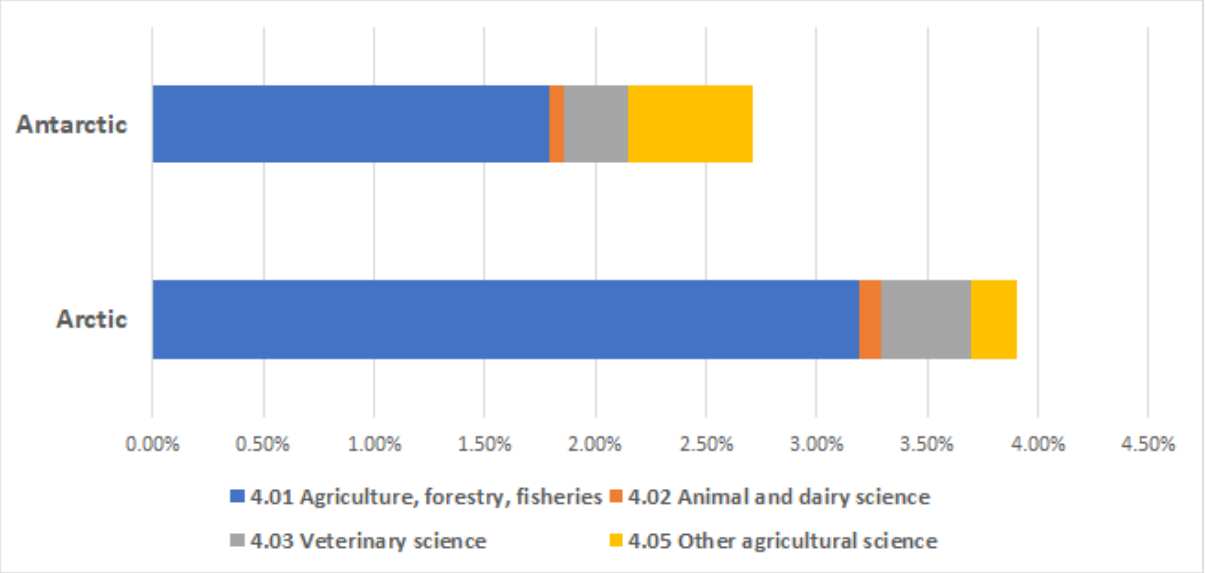
Within the natural sciences, which dominate both fields, the earth sciences make up a similarly sized slice of both, just under 60%. However, the biological sciences are more significant in the Antarctic, with 29% vs 23.5%. The Antarctic also has around twice the levels of the physical sciences and chemical sciences.



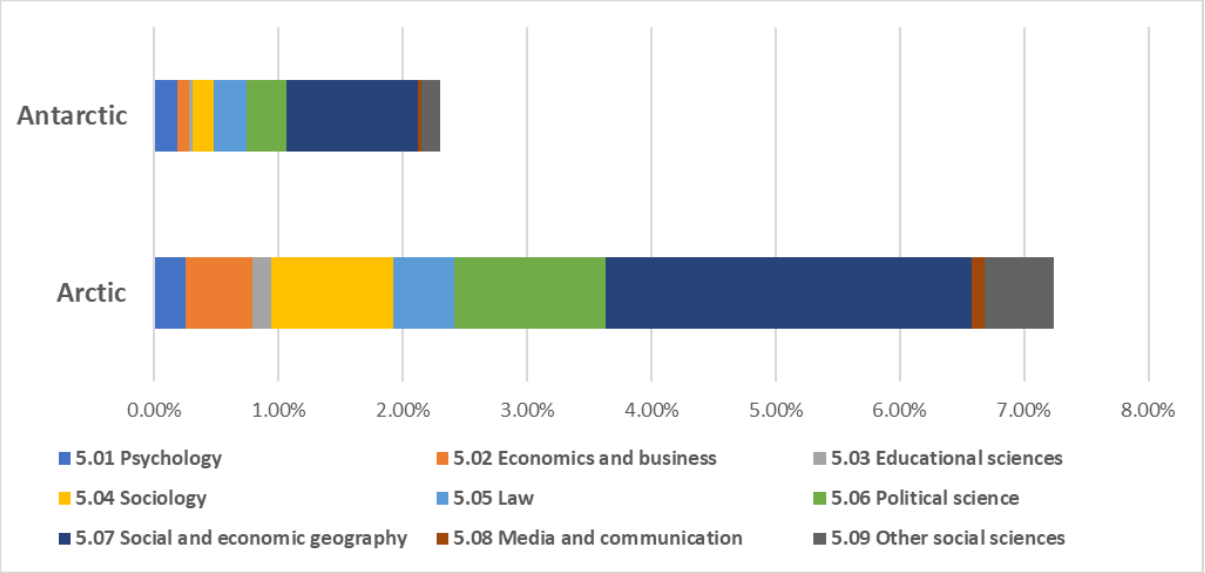
Within engineering, there are again some strong discrepancies - while the overall totals are similar, the Arctic has noticeably higher levels for civil engineering and environmental engineering, as befits an area where more emphasis is focused on human activities living in that environment, and much less for environmental biotech.



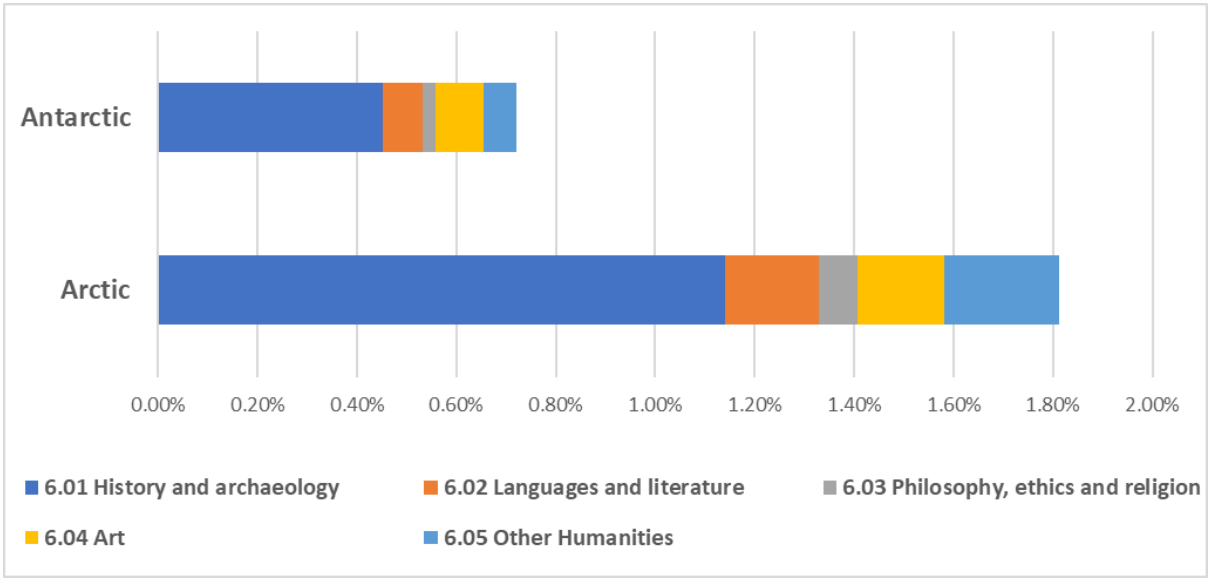
The level of basic medical research is similar in both, but the applied fields of clinical and health sciences are much more pronounced in the Arctic. Again, this is very consistent with assuming research driven by the needs of the populations living in that environment.



Within the agricultural sciences, the Arctic is more prominent in all fields except the catch-all "other". It is perhaps reasonable to assume that the share given to 'agriculture, forestry and fisheries' is more concentrated on fisheries for the Antarctic, but the data is not available at that level of granularity.



The broad social sciences category shows a very expected skew, with the Arctic heavily represented in all fields, but it is interesting to note some of the relative differences. The Antarctic has almost as much psychology research, perhaps a heritage from the long tradition of using Antarctic stations as isolated locations for human factors research, but much lower rates of economics, education, and sociology, all of which would focus on the existing populations. Meanwhile, both have strong showings in social/economic geography, reflecting research on these areas as geographic concepts.



The humanities are the smallest subset overall, less than 1% of Antarctic research and 2% of Arctic, so it is difficult to draw too many conclusions. In both, however, it is clear that history is the dominant segment.