

We would like to thank the four Reviewers and the Editor for their constructive suggestions, which helped us to improve the manuscript. Specific answers are given in blue text and manuscript modifications related to the Reviewer's and Editor comments are given in green text. Line numbers correspond to the revised manuscript without tracked changes.

Reviewer #1

This manuscript presents an important and timely assessment of diversity, equity, and inclusion (DEI) in the cloud physics community, combining a metadata analysis of 6,987 peer-reviewed papers (1970–2020) with a community survey of 198 respondents. The topic is highly relevant, and the effort to integrate quantitative and qualitative perspectives is commendable. However, several conceptual and structural issues limit the manuscript's clarity and impact in its current form. In particular, the terminology used to describe researchers from the Global South, the framing of gender analysis, and the interpretation of small-sample survey results require substantial revision. Addressing these concerns would considerably strengthen the analytical rigor, inclusivity, and global sensitivity of the paper.

Major Comments

1. **Terminology – “Tropical Scientists”:** In my opinion, the use of the term “tropical scientists” is highly problematic. I find it both unnecessary and ill-considered that the authors have chosen to introduce a new label that lacks conceptual grounding and adds no clarity to the discussion. As someone working within this geographic zone, I find this terminology technically inaccurate and, frankly, counterproductive—it conflates latitude with social and institutional realities, and risks exoticizing researchers from the Global South. In my view, this framing reflects a misunderstanding of the issue it seeks to address. Established terms such as Global North/South, high-, middle-, or low-income countries (as defined by the UN or World Bank), or developed/developing countries—though imperfect—are far more appropriate, as they rest on recognized metrics and are widely understood in academic and policy contexts. I strongly recommend removing the term entirely and clarifying the distinction between tropical study sites and researchers from the Global South, which are not interchangeable categories.

A/ We agree with the reviewer that the term “tropical scientist” is not completely appropriate and we thank you for the suggestion to replace it with “Global south” or “low-income countries”. However, we decided to replace it with “scientist working in a tropical country”. Although “Global south” and “low-income countries” could work, neither of these terms align with the purpose of why we were interested in scientists working in tropical countries in our study. We added the following paragraph to walk the audience

through our thinking process of why we study “*scientists working in the Tropics*” and how it relates to studying historically marginalized populations.

Lines 102-114: “In this study, we explore the regional representation bias of collaborations in the cloud physics community for scientist working in the tropics (anyone from whatever country who lives and works in a tropical country), i.e. within the latitudes of 23.5° N and 23.5° S. Multiple studies show that this region is one of the least studied in the field (Yakobi-Hancock et al., 2013, Petter and Wright, 2015; Kanji et al., 2017; Irish et al., 2017; Welti et al., 2020; Metcalfe et al., 2025). For example, input observations are often calculated rather than sampled for the Tropics in studies that evaluate the role of aerosols and cloud formation in climate models (e.g., Cesana et al., 2023; De Vera et al., 2024). The overlooked observations from tropical countries is a perpetuation of the historical marginalization of the region and its populations, based on historical legacies of colonialism, natural resources extractivism, and slavery (Rini, 2020; Pereiria and Tsikata, 2021; Atilas and Rojas-Paez, 2022; Feijo and Orre, 2024; Meyfroidt et al., 2024), and that today the effects of this history are still present in scientific practices (Godrie, 2025; Arosoaie et al., 2026). Therefore, this study not only focuses on the biases based on the location of scientists, but expands the discussion to evaluate the overall state of DEI in scientific spaces in the cloud physics community”.

2. **Use of “Minorities” vs “Marginalized”:** The manuscript frequently uses the term “minorities.” Given the global context of the study, where numerical minorities may differ by region, the term “marginalized” would be more appropriate and inclusive. It emphasizes systemic disadvantage rather than numerical representation. This change would also align better with the ethical framing of equity in the paper’s conclusion.

A/ We thank the reviewer for pointing this out and agree. Considering this comment and the comments from Reviewer #2 and Reviewer #4, we have modified the text replacing “minority” and “marginalized” with “historically excluded”.

3. **Gender Analysis and Non-Binary Representation:** The gender analysis appears to be strictly binary (men/women). While this may have been constrained by data availability (e.g., use of genderize.io), the discussion section should at least acknowledge and briefly discuss the exclusion of trans and non-binary individuals. For instance, the authors could note the limitations of algorithmic gender inference, and mention issues such as misgendering, the use of dead names in publication records, and the broader underrepresentation of trans scientists in the geosciences. This would strengthen the inclusivity and awareness reflected in the DEI framing.

A/ Section 2.2.2 was modified as follows:

Lines 210-215: “It is important to note that genderize.io assigns a binary (female or male) gender based on the person’s first name. This can only be a good approximation, as names cannot always be used to infer the gender (e.g., Sebo, 2021; Lockhart et al., 2023;

Marty et al., 2023). Critically, the binary gender assumption inherent in this methodology reinforces cisgender normativity and erases non-binary and possibly also trans identities. While being aware of these important limitations, the superficial approach necessitated by the quantitative analysis of the large amount of data cannot avoid them.”

We also added the following text to Section 3.1:

Lines 261-268: “The methodology used in the metadata analysis to assign gender to authors (genderize.io) assumes a gender binary, risking the misclassification of non-binary and gender-diverse individuals. We acknowledge this as a limitation of the present study that could have introduced a gender exclusion bias. As shown by the results in the survey described in section 3.2, at least three respondents self-identified outside of the gender-binary system, which shows the active participation of non-binary individuals in the cloud physics community. In section 3.3, we expanded the discussion on the experiences related to gender in the cloud physics community. Future metadata analysis should incorporate self-identification methods to include non-binary genders. For example, Son and Bell (2022) analyzed publications in which individuals self-identify their gender during manuscript submission.”

4. **Survey Interpretation and Sample Size Limitations:** The metadata analysis is the stronger and more robust component of the paper. However, the survey-based results (N=198) present significant limitations when further divided by gender, country, and ethnicity. When broken down into subcategories, the sample size per group, in my opinion, becomes too small for reliable interpretation or for drawing broad conclusions about regional or demographic patterns. I suggest restructuring the Results section so that the metadata analysis comes first—forming the backbone of the findings—followed by the survey results as complementary qualitative insight. The discussion could then focus on how these two strands inform each other, rather than treating both as equally representative.

A/ Thank you for this suggestion. We have reorganized the paper accordingly.

Reviewer #2

In this study, the authors present the current state and over time development of Diversity, Equity, and Inclusion (DEI) in the cloud physics community by considering two different data sets: metadata of 6,987 published papers between 1970 and 2020, and the survey responses from 198 participants. This research is significant for the advancement of the geosciences as it helps identify areas that need improvement. However, there are some issues with the framing of certain aspects, including the terminology used. Therefore, my general comments align with the Rev1.

General comments:

- **Underrepresented instead of the word minority or marginalised:** Minority means that the smaller number or part. As Rev1 mentioned, minorities differ from region to region. It is important to note that labelling the global south as a minority is a Western-centric viewpoint. The groups considered in this context are not smaller in number; they are underrepresented, overlooked, or lacking visibility. Additionally, "marginalised" implies insignificance or a peripheral status, which also doesn't adequately convey the issue at hand.

A/ Thank you for pointing this out. Considering this comment and the comments from Reviewer #1 and Reviewer #4, we have modified the text replacing "minority" and "marginalized" with "historically excluded".

- **Terminology tropical scientists/countries:** The term "tropical" carries connotations that could present an "exotic" portrayal, which is not appropriate here. I completely agree with Rev1. Additionally, framing a comparison of "tropical countries" with so-called global north is a false dichotomy. One is defined by economic and historical factors under neo-colonial structures, while the other is geographical.

A/ We agree with the reviewer that the terminology implemented may not be appropriate for the discussion. We have replaced the term "tropical scientist" by "scientist working in a tropical country". Please see our reply to question 1 for Reviewer #1 for modifications made to the manuscript.

- The gender analysis presented is binary and fails to include non-binary and transgender individuals. This omission is not addressed in the paper. Although the potential for misgendering due to the algorithm is acknowledged, it still contains uncategorized (gender-blind) names. Furthermore, such gender-inferencing algorithms often struggle with the accuracy of non-Western names.

I will try to elaborate on what I mentioned in the general comments by giving specific examples from the first couple of pages, which are valid throughout the entirety of the paper.

A/ We appreciate the discussion on this limitation of the study. Please see our reply to question 3 for Reviewer #1 for further reference and modifications made to the manuscript.

line 46: "... minority representation..." This gap exists for any underrepresented group, not just minorities. You might not be a minority, but underrepresented, you might be a minority if you are looking from a Western lens. We are great in numbers, but our work or existence is not represented or acknowledged.

A/ The sentence was changed as follows. Lines 47-51: "Although various programs, policies, and inclusive environments are being developed to address issues such as gender inequality in academia (Ranganathan et al., 2021; Bernstein, 2024), many gaps in other historically excluded groups such as Hispanic, black, indigenous communities, scientists with disabilities, and LGBTQIA+ scientists still remain (Bernard and Cooperdock, 2018; NCSES, 2019; Gewin, 2020; Odekunle, 2020; Berhe et al., 2022; Marin-Spiotta et al., 2023)."

line 65: same issue here with the use of minorities

A/ The term "minorities" was changed to "historically excluded groups"

line 70: Women are underrepresented but larger in number, exactly like the scientists from the so-called global south. They are underrepresented, and there is a lack of acknowledgement.

A/ We agree. The sentence was changed as follows. Lines 71-72: "Geosciences are among the least diverse fields, in which historically excluded groups are exposed to systemic biases that prevents them from equal progress and recognition compared to their more privileged counterparts"

line 81: In the cited paper, this is mentioned as "USA, UK, ... However, these are not necessarily the most collaborative countries, if measured by their proportion of collaborative output... The analysis provided evidence that countries rated high in terms of scientific development were more likely to collaborate." This does not necessarily mean low income with low scientific development.

A/ The sentence was changed as follows. Lines 88-89: "are less likely to cooperate with countries with low scientific development."

line 86&96: Egypt and South Africa are not situated in the tropics but the subtropics. When comparing these regions with the global north, framing it as a dichotomy could be

misleading, aka a false dichotomy. Additionally, the term "tropical" carries connotations that could present an "exotic" portrayal, which is not appropriate here.

A/ Line 96 was deleted and the sentence from line 86 was modified for clarity. Lines 95-96: "The existing collaboration networks mentioned by Wagner (2005) and Gazni et al. (2012) highlight the limited connections between Global North and Global South research groups"

line 189: Latitude range includes Australia, which is considered to be a "global north" country. If the intention is to analyse underrepresentation, this could lead to misrepresentation in the analysis. Although the countries are highlighted later in the table, it would be beneficial to clarify this point earlier on.

A/ We agree with the Reviewer. The sentence was modified for clarity. Lines 218-220: "For a country to be considered tropical in this study, at least 50% of its inhabitants need to live in the tropics, i.e., between 23.5° N and 23.5° S and be considered a low-income country, with the exception of Taiwan, Singapore, and Hong Kong (See Table 1)."

Reviewer #3

This paper presents an analysis into the current climate and feelings of belonging within the specific geoscience community of cloud physics. It utilizes two different approaches to address this question: a meta-analysis of peer-reviewed publications and an online survey. The question addressed in the paper is a unique contribution to the field, however there are several significant issues that need to be revised and addressed before publication.

In addition to the previous two referee comments above, I will add the following:

1. There has been a surplus of more recent work that highlights the their point in lines 45-47 (Although various programs, policies, and inclusive environments are being developed to address issues such as gender inequality in academia, many gaps in other minority representation still remain (e.g., Hispanic, black, and indigenous communities; Bernard and Cooperdock, 2018; NCSES, 2019; Gewin, 48 2020; Odekunle, 2020). Please include an updated literature review from the last 5 years in this.

A/ We agree and have added to the revised manuscript the following recent studies: Ranganathan et al. (2021), Berhe et al. (2022), Hall et al. (2022), Marin-Spiotta et al. (2023), Bernstein (2024), and Rubbia (2025).

2. I agree with the previous two referees that “marginalized” is a more appropriate term than minority. Based on the descriptors provided, another alternative would be to say “historically excluded” groups. This is a better descriptor that is commonly used and recognizes the fact that these groups have been intentionally excluded by the system, as opposed to just individual actions.

A/ Thank you for pointing this out and we agree. Considering this comment and the comments from Reviewer #1 and Reviewer #2, we have modified the text replacing “minority” and “marginalized” with “historically excluded”.

3. In line 55 – able-bodied individuals are also significantly more represented in all fields than in the general population. There is literature on this as well and should be included in the literature review.

A/ We agree and have modified the following sentences to clarify these points:

Lines 54-57: “Many fields lack diversity in gender, ethnicity, disability, and language including health (Wenneras and Wold, 1997; Landman and Dandolu, 2009; McDermott et al., 2018), science, technology, engineering and mathematics (STEM)(Nadis, 1999; Watt, 2005; Haak, 2002; Barber et al., 2020; Ceci and Williams, 2011; Leslie et al., 2015), as well as humanities, in particular philosophy (Nielsen, 2016; Botts et al., 2014; Herfeld et al., 2022).”

Lines 77-81: “Ford et al. (2019) reported that historically excluded groups including women are less often invited to give talks, which are essential to establish a network,

funding, and academic collaborations. Similarly, individuals with disabilities report not only feeling less safe and included in Geoscience space like fieldwork (Lawrence and Dowey, 2021; Castro and Atchison, 2024), but also receiving less research assistantships, internships, fellowships, scholarships and grants than their able-bodied counterparts (NCSES, 2021).”

4. Lines 58-59 (The lack of diversity and inclusion across academic communities manifests itself in systematic bias of under-represented minorities). I disagree with how this comment is written – the system was intentionally designed to prevent lack of diversity and inclusion from the very beginning. The system is the problem, which further exacerbates the lack of diversity and inclusion we see today.

A/ We agree with the Reviewer and modified the corresponding paragraph as follows:

Lines 57-62: “The lack of diversity, equity, and inclusion in academia is one of the manifestations of a system originally structured to privilege dominant groups and exclude marginalized populations, a pattern that has historically characterized most fields in science and that continues to influence their institutional practices. Studies on practices in academia have found bias against under-represented groups (Gates, et al., 2019) that affect hiring and promotion processes (McDermott et al., 2018; Nielsen, 2016) and other career outcomes”

5. I think it would also be beneficial to expand on the background of cloud physics as a research field. It seems that your assumption is that everyone in this field trained in geosciences, or is there a combination of physics, chemistry, geography, etc? Each of these fields has presented its own variations in why they have issues related to DEI and should be included as a background variable when considering the climate of cloud physics.

A/ The following text was added to the revised manuscript to acknowledge this.

Lines 114-117: “It is important to note that the cloud physics community is a multidisciplinary field where meteorologists, physicists, biologists, oceanographers, chemists, environmental engineers, and chemical engineers, among others, work together in cloud and climate modelling, field, theoretical, and laboratory studies which implies variabilities in issues related to DEI.”

6. Your response rates are not clearly defined and need to be included. How many individuals were included in the ICCP listserv and the other announcement sent? How many participants actually registered for the 2021 meeting? Comparing this number to other membership numbers is not sufficient – especially when the data is anonymous and you cannot guarantee that individuals did not take the survey more than one time. I don’t agree with the statement that it represents 30-47% of the community as a whole – you should only report on response rates from who the survey was disseminated to.

A/ The survey was sent to a maximum of 938 email addresses (as of 2023); however, there is no way of verifying that all these email addresses were active because of many retired past members, and many people who have passed away. In addition, we do not have the exact numbers for 2021, since the listserv was migrated to a different server so the earliest numbers we have are for 2023. Thus, we report that the upper limit of the listserv numbers is 938. We have now added this information to Lines 153-154 of the revised manuscript.

For the 2021 meeting in Pune there were 570 attendees which we have already indicated in the original manuscript (previously Line 137, now Line 159). We believe comparing the response rates to the past three meeting attendee numbers is also useful because it captures the active participation of respondents in the survey. However, it does bias against those who could not attend due to schedule conflicts, distance, pricing. This is why we do not use *only* the 2021 numbers to calculate the response rates, but also use numbers from the 2012 and 2016 meetings which were not in Asia. This brings us to the % coverage of 30-47%. This information is now included in addition to the lower limit of the response rate calculated as a percentage of the listserv which is $198/938 = 21.1\%$ (Lines 160-161 of the revised manuscript).

7. Lines 137-139: you say the number of respondents who were Indian was higher – but how do you know that if the data is anonymous? Please clarify.

A/ We were able to reach this conclusion because we have the data on ethnicities in the survey, so we can calculate the percentage of ethnically Indian respondents which was 14% (n=27). Furthermore, when we compare the 2021 meeting figures (28% n=160) to the 2012 (3% n=20) and 2016 (1%, n=4) meetings it becomes clear that the Indian respondents had a higher percentage representation than their percentage presence at previous conferences. This is explained by the fact that the conference was indeed held in India in 2021 making it more accessible to the local participants.

8. Lines 206-208: What are the actual stats presented by Dutt, 2020? The ratio of white/non-white seems higher than what the ratio is in the US.

A/ Dutt (2020) claimed that “almost 90% of doctoral degrees are awarded to White people”. The sentence was rephrased as follows. Lines 361-362: “This trend agrees with the overall geoscientist PhD awardees in the US, where White people represent >90% (Dutt, 2020)”

9. Figure 1: The color purple is used to signify two different things – please use different colors for the “total” versus “other” gender category.

A/ The figure was modified accordingly.

10. The paper published by Marin-Spiotta et al (2023) in Earth’s Future provides survey results on the workplace climate in the geosciences. You should include this in your discussion/analysis as a comparison since their survey results include over 2000 survey responses.

A/ We thank the Reviewer for the suggestion. This interesting study was added along several sections of the revised manuscript and the following text was added to the discussion.

Lines 457-459: “The present results are in somewhat in agreement with those reported by Marin-Spiotta et al. (2023), who found in a much larger survey applied to geoscientists, that harassment (specifically sexual harassment) is experienced by historically excluded groups at higher rates.”

11. Did you ask if survey respondents had experienced bullying or harassment within a specified time frame or ever? This should be further clarified and justified on either approach.

A/ We thank the reviewer for this very important comment, unfortunately at the time of the survey, we did not ask the participants about a specific time frame where they experienced bullying or harassment. So, the question implied “ever”. Which is why we did the age analysis in Figure 6 to show that for the younger age group the % of respondents experiencing harassment is lower, which could imply two things, more recent awareness, or less time spent in the field. Both scenarios have already been acknowledged in the manuscript.

12. Line 324 – the leaky pipeline analogy is problematic. Berhe et al (2024) describes this in their Nature Geoscience paper and recommends a better analogy.

A/ Thank you for pointing us to the flaws of this analogy. We have reformulated our discussion as follows:

Lines 495-500: “This aspect may result in a negative gender feedback, discouraging women from remaining in academia and research in the cloud physics community (see Section 3.5). The diminishing number of women scientists as they move up the career ladder has been described as a leaky pipeline (Wolfinger et al. 2008). However, as Berhe et al. (2022) argue, a hostile obstacle course more aptly reflects the structural and systemic problems, the fact that barriers have at least unconsciously been put in place and sustained, and that the obstacles “selectively slow down scholars from historically excluded groups”.

We have also removed the reference to the leaky pipeline in Line 454 in the original manuscript since this analysis did not take academic positions into account, we cannot make a definitive statement on the retention and success of women in cloud physics.

13. Line 335-336. I disagree that you have the numbers to state that you are representing the cloud physicists globally, based on my previous comment about response rates. This should be removed.

A/ We agree with the reviewer, and have removed the respective sentences (Lines 333-336 in the original manuscript).

14. Given the broad demographics of who this survey was disseminated to, I think you should expand on potential limitations of your findings on feeling included.

Your question was simply “Do you feel included in the cloud physics community?” This is really vague and could be interpreted in so many different ways – I am less inclined to say that this is a valuable finding with this approach. What do you mean when you say included? Just in the workplace, in professional societies, when in school? This is not well defined.

A/ We agree that this is an open-ended question and not well defined. However, what it feels to be included means different things to different people, in different situations and we wanted to encompass the individual feeling of inclusion within the context being the cloud physics community, regardless of the locations being the workplace, education or social situations. We wanted respondents to answer based on their perceptions of feeling included and not define different spheres of inclusion a priori. We believe it is a valuable finding as it speaks to the sense of belonging (or lack thereof) from the respondents.

We acknowledge this limitation in the following statement. Lines 392-393: “It is important to note that the feeling of inclusion means different things to different people. Therefore, these results must be interpreted with caution”.

15. Why did you only ask about working with women in your survey? This is not inclusive to the entire gender spectrum.

A/ We agree with the reviewer that this leaves out the nonbinary conforming genders. However, we wanted to focus on women because that has been a huge focus in the geosciences in the last decade (Stokes et al., 2015; Popp et al., 2019; Pico et al., 2020; Ranganathan et al., 2021; Koffman et al., 2023; Bernstein, 2024), to bring in gender equality and the gender parity has largely been assessed on men and women using binary genders. We have now clarified this point in the revised manuscript:

Lines: 414-417: “Note that we have targeted women because there has been a huge focus in the geosciences in the last decade (Stokes et al., 2015; Popp et al., 2019; Pico et al., 2020; Ranganathan et al., 2021; Koffman et al., 2023; Bernstein, 2024), to bring in gender equality and the gender parity has largely been assessed on men and women using binary genders.”

16. What does “sensitive to gender issues” mean? Again as above – this is so vague and could be taken to mean so many things. I don’t agree that this should be presented in this dataset, it’s not a well designed survey question.

A/ Similar to comment 14 above, we did not want to specify specific gender issues because everyone’s perception of sensitivity and gender related issues are subjective. In general, it is meant to give an overview if work places of the respondents are supportive with regards to different genders working in the same space. So, we believe that this question provides a good overview of gender related issues in the workplace without delving deeper into specific details.

17. Question 40 in your survey – were you only talking about race/ethnicity here? Or was this applicable to any and all identities? Not clearly defined, which is problematic.

A/ We did not define the term minority for the respondents in the survey, because we wanted again to capture the perception of the respondents as minority. Specifically, we did not want to focus only on race and ethnic minorities but on minorities in any category such as disability, gender, sexual orientation, religion, and any other perceived minority status. The global nature of this survey made it challenging to define race/ethnicity as a historically excluded group, because for example an ethnically Indian person in Switzerland would identify as a minority, however, in India this person would not identify as a minority. As such we believed the phrasing of minority should include any perceived minority by the respondent.

18. You inherently introduced a lot of bias in your survey by using the term “tropical” to describe an entire subset of people. This is problematic.

A/ We agree with the reviewer. This was a bias motivated by the fact that this study is part of a larger project aimed at quantifying the lack of aerosol and cloud sampling in geographically defined tropical regions and its impacts on climate change modeling. Research regarding aerosol-cloud interactions and cloud physics in the tropical regions is paramount for improving climate forecasts, but yet it is lacking. We agree that, without this context, the term may sound problematic; however, we decided to keep the term and add a qualifier to explain the use of the phrase tropical scientists and tropical regions (Lines 102-104 in the revised manuscript).

19. You say that the majority of people said no to question 67. Why did you not include an additional clarifier on that to explain their perspective? There could be many reasons why participants selected no, I don't think it's appropriate to speculate on that in your paper.

A/ We did not speculate as to why people answered “no” in question 67. We just state that 60% are not planning to have collaborations with researchers from developing countries. Precisely because we did not ask for reasons, we just stated the results of that question and put it in context of the other questions (see Lines 325 - 332 in original manuscript, now Lines 500-509 in revised manuscript).

Reviewer #4

The following is a revision to my first review that was flagged as "AI-generated." My process for that version was to use the internal commenting tool while reading the manuscript, and then use AI to consolidate my comments into a single, readable file. When I was alerted that my review had been flagged, I ran that text through an AI detector, and it flagged quotes from the manuscript, equations, and associated comments on statistical assumptions and limitations, as well as basic textbook knowledge.

The following text is based on the original comments that I left in the downloaded manuscript, which I modified somewhat after re-reading for this review. My biggest concerns are the lack of acknowledgment of the limitations of statistical methods and the language around representation.

PART 1: This review is both significant and timely. However, it demonstrates the tendency within academia to use archaic language and framing when analyzing geographic, racial, gender, and sexuality DEI issues globally. To address these, significant revisions are needed before publication.

Include a Limitations section that explicitly addresses how categorizing gender as binary and grouping racial/ethnic reflects, and potentially perpetuates, systemic bias (this is why intersectionality is a critical parallel thread to your analysis).

A/ We thank the reviewer for this suggestion and we believe that it is better to mention the limitations of our study in each subsection so that they are immediately clear to the reader rather than having a single section for this. We highlight the limitations of our study in the following paragraphs. We also added Section 3.6 "Limitations" in Lines 578-590.

Lines 210-215: "It is important to note that genderize.io assigns a binary (female or male) gender based on the person's first name. This can only be a good approximation, as names cannot always be used to infer the gender (e.g., Sebo, 2021; Lockhart et al., 2023; Marty et al., 2023). Critically, the binary gender assumption inherent in this methodology reinforces cisgender normativity and erases non-binary and possibly also trans identities. While being aware of these important limitations, the superficial approach necessitated by the quantitative analysis of the large amount of data cannot avoid them."

Lines 261-268: "The methodology used in the metadata analysis to assign gender to authors (genderize.io) assumes a gender binary, risking the misclassification of non-binary and gender-diverse individuals. We acknowledge this as a limitation of the present study that could have introduced a gender exclusion bias. As shown by the results in the survey described in section 3.2, at least three respondents self-identified outside of the gender binary system, which shows the active participation of non-binary individuals in the cloud physics community. In section 3.3, we expanded the discussion on the experiences related to gender in the cloud physics community. Future metadata analysis should incorporate

self-identification methods to include non-binary genders. For example, Son and Bell (2022) analyzed publications in which individuals self-identify their gender during manuscript submission.”

Lines 319-323: “It is important to note that our conclusion on parachute science needs to be caveated as the non-English papers were excluded from the present analysis. Therefore, it is possible that a different scenario would emerge if the excluded papers were considered.”

Lines 392-393: “It is important to note that the feeling of inclusion means different things to different people. Therefore, these results must be interpreted with caution”.

Lines 436-437: “Although the terms bullying and harassment are not strictly synonyms (Einarsen, 2000), the way the question was formulated in the survey could have led participants to understand that these concepts were the same.”

Lines 455-457: “The reference to intersectionality is interpretive, it offers a possible lens for understanding, not a statistically proven explanation. Therefore, the reported percentages are to be taken with a caveat because of the small numbers.”

Lines 558-560: “Since this analysis did not take academic positions into account, we cannot make a definitive statement on the retention and success of women in cloud physics”

Include Position Statements from the authors. This clarifies how author backgrounds inform survey design and interpretation. Position statements are standard operating procedures in education and sociological research.

A/ A collective position statement was added to the methods (Section 2.3) of the revised manuscript.

Lines 224-241: “The author team for this article comprises scientists with a range of disciplinary backgrounds, upbringing, and experiences related to DEI. We were born, and have lived or worked in Canada, Colombia, Ecuador, Finland, France, India, Israel, Kenya, Mexico, the United States of America, and Switzerland. Our disciplines span from chemistry, over atmospheric science, to sociology and not all of us are currently working in the cloud physics field. We are able-bodied and cis-gender. Our academic positions range from postdocs to professor.

In one regard or another all of us are part of a historically excluded group, be it as born in a tropical South American country, first generation who attended university, person of color, woman, indigenous person, or in relation to sexual orientation, religion or nationality. This background has undoubtedly sensitized us for DEI, albeit from different angles.

We are committed to building bridges that help his community to be more diverse, equitable, and inclusive. For many of us, being part of the global majority and excluded

nationality, our lived experiences from intersectionality are different from the Global North. Also, many of us do hold privileged positions and acknowledge that the access to western education systems and structures have afforded us privileges. Our experiences thereby may both sensitize and limit our engagement with the full range of intersectional and geographic dimensions of marginalization explored in this study. The focus of the survey and study design on tropical countries and gender effects may reflect our experiences, but our differing backgrounds also allowed us to add many and question each other's perspectives."

Lines 50-54: "The word 'minority' denotes negatively stigmatized, ostracized, oppressed, and outcast individuals or groups (Blanz et al., 1995). In the context of DEI, it refers to a numerically smaller population of a group within a larger population based on characteristics such as gender, race, religion, ethnicity, disability, sexual orientation, or other attributes. In the present study, 'minority' is used to describe an underrepresented or disadvantaged group, unless stated otherwise." Issue: I appreciate that you have acknowledged the negative connotation of the term "minority". However, the term is appropriately used when discussing percent representation. When a group of people is not in the majority, they are often marginalized/oppressed due to systemic power structures.

A/ Thank you for pointing this out. Considering your comment and the comments from Reviewers #1, #2 and #3, we have modified the text replacing "minority" and "marginalized" with "historically excluded".

Lines 66-69: "Systemic bias, under-representation of minorities, and a lack of recognition of researchers in and from developing countries (the so-called 'Global South') are reflected in the limited scientific collaborations between research groups from developed (the so-called 'Global North') and developing countries (Gewin, 2023)." Issue: Drop the use of "so-called".

A/ We agree, we have deleted "so-called" in lines 66 and 67 of the original manuscript.

Lines 150-153: "Therefore, in further analysis, we classify the Caucasian, White or European categories as 'White', while the other ethnicities are classified as 'non-white' with the exception of participants who preferred not to identify their ethnicity." Issue: This binary centers whiteness as the default and groups all other identities into a single "other" category, which obscures important differences in experiences among diverse racial/ethnic groups. This approach reinforces white supremacy frameworks.

A/ We agree with the reviewer. However, we assume the reviewer means White and non-White, as we did not categorize all the historically excluded respondents into the "other" category, but rather into the non-White category, which still implies White as a reference. We thank the reviewer for pointing this out and instead we use the term people of global majority (POGM) or simply global majority (GM) coined by Campbell-Stevens (Campbell-Stevens 2009; 2021a), which acts to de-center White as the reference and encompasses Black, Asian, Brown, Indigenous, and mixed-heritage people, comprising approximately 80–85% of the world's population (Campbell-Stevens 2021b) highlighting their status as

the global majority. We have now changed all figures and text to change non-White to (PO)GM.

Campbell-Stephens, R.: Investing in Diversity: changing the face (and the heart) of educational leadership. *School Leadership and Management*, 29, (3), July 2009, 321-33, 2009.

Campbell-Stephens, R.M.: Introduction: global majority decolonising narratives. In: *Educational leadership and the global majority: decolonising narratives*. Switzerland: Palgrave Macmillan, pp. 1–21, 2021a.

Campbell-Stephens, Rosemary M.: *Educational Leadership and the Global Majority: Decolonising Narratives*. Springer Nature. p. 8–9, 57–58. ISBN 978-3-030-88282-2, 2021b.

Lines 182-186: "It is important to note that genderize.io assigns the gender based on the person's first name. Although this could be a good approximation, names cannot always be used to infer the gender (e.g., Sebo, 2021; Lockhart et al., 2023; Marty et al., 2023). Therefore, the results presented below must be taken with caution and considering the aforementioned bias." Issue: This perpetuates binary gender assumptions and doesn't acknowledge how this methodology erases non-binary identities and reinforces cisgender normativity. The authors should explicitly state this limitation.

A/ Thank you for pointing us to the need to explicitly name context and limitations. The first name to gender methodology certainly is superficial, but we take it from this comment and the other Reviewer comments that need for the superficiality is in this regard is clear when attempting the large data gender analysis. We have added the considerations suggested as follows:

Lines 210-215: "It is important to note that genderize.io assigns a binary (female or male) gender based on the person's first name. This can only be a good approximation, as names cannot always be used to infer the gender (e.g., Sebo, 2021; Lockhart et al., 2023; Marty et al., 2023). Critically, the binary gender assumption inherent in this methodology reinforces cisgender normativity and erases non-binary and possibly also trans identities. While being aware of these important limitations, the superficial approach necessitated by the quantitative analysis of the large amount of data cannot avoid them."

Lines 276-278: "A higher fraction of women, 44% compared to 27% of men, reported having experienced harassment, a difference that was statistically significant ($\chi^2(1) = 5.70$, $p = 0.017$, Cramér's $V = 0.18$), although the size of the gender difference is modest." Issue: In this reviewer's opinion, a 17% difference is not "modest". Sexism and misogyny, which is well-understood to be underreported in STEM fields, create hostile environments and power differentials and leads to work-place environments prone to harassment.

A/ We agree with the reviewer and have removed the sentences as the term "modest" as a descriptor as it trivializes the concept of sexism, misogyny and harassment and rather focuses on the significant difference between men and women.

Lines 284-286: "A higher fraction of non-White women reported experiencing bullying and harassment, which aligns with the concept of intersectionality, where overlapping

social identities and positions can increase the likelihood of discrimination (Crenshaw, 1991). However, the sample size prevents statistical verification of this pattern." Issue: Intersectionality is mentioned but not systematically applied throughout the analysis. The entire study design should have incorporated intersectional analysis from the beginning, not as an afterthought.

A/ The reference to intersectionality was purely descriptive and interpretive, no formal intersectional analysis or statistical testing was performed. As such, no pre-registration or power calculations were applicable. The subgroup percentages are not statistically robust and are presented only to highlight potential patterns warranting future investigation with larger samples. We believe it is worth reporting because this has not been done for the cloud physics community before. The following text was added to acknowledge this:

Lines 455-457: "The reference to intersectionality is interpretive, it offers a possible lens for understanding, not a statistically proven explanation. Therefore, the reported percentages are to be taken with a caveat because of the small numbers."

Lines 393-408: As a practice, parachute science is a continuation of colonialism. Be sure to include how biases (implicit and explicit) around the "ownership of knowledge and expertise" contribute to parachute science. Consider FAIR data practices and cite Ramírez-Castañeda, et al. (2022). PNAS 119 (34). At the very least, identify colonialism as the founding principle of parachute science.

A/ Thank you for pointing us to the lack of context in discussing parachute science practice. We have expanded as follows:

Lines 314-320: "Research practices in which scientists from wealthier countries conduct studies in lower-income countries, often without meaningful collaboration or involvement with local researchers or communities and then publish the results without authorship acknowledgement or benefit to the locals, have been termed helicopter or parachute science. These practices have a colonial and extractivist legacy and may be perpetuated by the belief that Global North institutions have knowledge to be transferred to partners in the Global South (Gewin, 2023; Stefanoudis et al., 2021; Odeny and Bosurgi, 2022; de Vos and Schwartz, 2022, and references therein). Our results suggest that parachute science is unfortunately common in the cloud physics community".

Lines 331-333: "Although there is not a clear way to eliminate these practices, as highlighted by Ramírez-Castañeda et al. (2022) the following four principles should stand out when performing field work: be collaborative, be respectful, be legal, and be safe."

PART 2. The dual-method approach (metadata analysis and survey) is valuable. However, in my review of procedures and statistical analyses, I've identified several significant methodological and statistical problems.

Throughout the document, chi-square tests are used (e.g., lines 244-245, 277-278, 291-293), but there is a lack of verification regarding required assumptions (cell counts

are greater than 5). Consider using Fisher's exact test (Kim, H. Y. (2017). Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test. *Restorative dentistry & endodontics*, 42(2), 152.) for small subgroups where cell counts are less than 5. For example, the subgroup for LGBTQIA+ = 8%, (line 200).

Be sure to apply statistical tests consistently. Report all chi-square tests with effect sizes, percentages with (n=xx) values and describe differences as significant or not significant (support with appropriate test). Be clear about when you are using descriptive vs. applied statistics.

A/ We started with broad descriptives to set the scene, show overall trends, and highlight groups that might be worth deeper look. For example, finding that 30% of respondents don't feel included prompted us to break the data down further by age, gender, and identity. We then tested meaningful comparisons, but only when group differences were large enough to warrant inference (for instance, running a χ^2 test on gender vs. career impact where the result was significant. Finally, we qualified our limitations by explicitly stating when a small sample size prevented formal testing. For example, with only 15 LGBTQIA+ respondents, no statistical test was run, though the finding was still noted descriptively and considered alongside external context.

To elaborate a bit more:

- Descriptive reporting (e.g., Lines 230–237 in the original manuscript): When presenting subgroup proportions without formal hypothesis testing (e.g., “35% of those <40 did not feel included”), we were highlighting observed patterns for qualitative or exploratory context, not making inferential claims. These were not intended as statistical tests and were explicitly framed as descriptive.
- Statistical testing (e.g., Lines 244–245, 277–278, 291–293 in the original manuscript): We applied chi-square tests with effect sizes (Cramér's V) only when comparing two or more groups on a categorical outcome where a meaningful association was hypothesized (e.g., gender vs. perceived career impact). This was done to assess whether observed differences were likely due to chance.
- Explicit “not statistically significant” statements (e.g., Line 257–258 in the original manuscript): These were included only when a formal test was conducted and the result did not reach conventional thresholds: to avoid overinterpreting non-significant trends.
- No testing for small subgroups (e.g., LGBTQIA+ N=15): We intentionally avoided statistical testing for very small groups (e.g., n<20) because chi-square tests are unreliable with low cell counts. Instead, we presented these as descriptive observations, often with reference to external literature (e.g., Le Bras, 2021) to contextualize them.

Therefore, we consider that there was no arbitrary “decision rule” rather, a principled distinction between descriptive illustration (for exploratory or small-sample observations) and inferential testing (for comparisons with adequate power).

Lines 75-78: "the inclusion of researchers from new geographical locations was not significant from 1990 to 2000 (Wagner, 2005) but it was significant between 1990 and 2013 (Wagner et al., 2017)." Issues/Suggestions: This combines 50 years of data (1970-2020) without addressing dramatic changes in academic publishing, databases, and geopolitics over this period. Consider time-stratified analyses, the effects of geopolitics on collaboration, and the rise of the Internet and digitization effects.

A/ We thank the reviewer for this suggestion; however, we consider that this analysis is beyond the scope of the present study.

Lines 136-143: Claims representativeness based on conference attendance but conference attendance is not a good measure for STEM practitioners, as it excludes industry, researchers who cannot attend due to cost, distance, timing, etc. Additionally, the 3 conferences included were from different years (2012, 2016, 2021) and then compared to 2021 survey and any demographic shifts (especially the aging out of white males). Shishkova, E., Kwiecien, N. W., Hebert, A. S., Westphall, M. S., Prenni, J. E., & Coon, J. J. (2017). Gender diversity in a STEM subfield—analysis of a large scientific society and its annual conferences. *Journal of the American Society for Mass Spectrometry*, 28(12), 2523-2531.

A/ We have now included additionally the percentage response rate based on the entire listserv (n=938) to whom the survey was sent out, which is $198/938 = 21.1\%$ in Lines 160-161 of the revised manuscript as suggested by previous reviewers with a similar comment.

Lines 142-144: "Although the ICCP membership is representative of the international cloud physics community, there could be some unintended biases in the survey data coming from the overrepresentation of respondents born in India, Germany, and US (above 15% each) as well the overrepresentation of respondents younger than 40 years old (64%)." Issues/Suggestions: Consider applying some weighting or stratified sampling corrections, conducting a sensitivity analysis to test how this potential overrepresentation affects conclusions. The geographic bias is particularly problematic for a study examining Global North/South dynamics

A/ The survey was sent by email, and not distributed at the conferences. As such it was emailed to the entire listserv of the ICCP community mailing list regardless of location. The responses reflect where most of the cloud physics research is taking place and thus where most of the scientific community comes from. This in itself is a result, and we do not consider such observation should be weight corrected as is important to highlight the uneven distribution. We are not examining dynamics of Global North/South. We are rather examining the influence of cloud physics research in the tropics.

Lines 157-171: "Although 8199 papers were pre-selected (from the documents available in Scopus on December 2021), those papers published in Astronomical and Astrophysical journals and other papers not related to cloud physics or written in a language different than English were removed leaving us with a total of 6987 papers." Issue/Suggestion: (1) Why didn't the authors utilize translation software or library services to translate papers? They state, "We note that the number of papers (i.e., 97)

removed because of language could slightly impact the outcome of the present analysis; however, this was necessary because the authors of the present study were unable to translate them.” This systematically excludes non-English publications, likely disproportionately affecting Global South researchers. I suggest that the authors include some text about how they would recommend addressing this challenge in the future. (2) In (lines 168-171): “the conclusions raised from the metadata analysis could be biased as Scopus, and most databases, does not cover all areas of interest leaving out important information from e.g., non-English-speaking countries or Global South countries.” My concern here is that this is exactly the population the study aims to understand. This is a methodological flaw for a DEI study.

A/ We agree with the reviewer that including those 97 papers would have been ideal; however, this would require hiring several people to either translate the documents or to corroborate that the online translation software did a good job, which was unfeasible for our study that had no dedicated funding. Since these papers present about 1% of the total number of papers, their analysis would not have a major effect on our conclusions.

Lines 180-186: “Each co-author's gender was evaluated employing the genderize.io web server (Demografix ApS, 2022) and included in the analysis if the certainty estimate provided by the tool was larger than 75%.” Issue/Suggestion: This reviewer suggests that you provide some justification for the 75% threshold and discuss the potential error rates for (1) using AI to determine gender from names (given that you are using the male/female binary system to sort names); (2) cultural variations in naming; and (3) explain how limitations may affect the validity of conclusions.

A/ As mentioned above, we acknowledge this limitation in Lines 210-215: “It is important to note that genderize.io assigns a binary (female or male) gender based on the person’s first name. This can only be a good approximation, as names cannot always be used to infer the gender (e.g., Sebo, 2021; Lockhart et al., 2023; Marty et al., 2023). Critically, the binary gender assumption inherent in this methodology reinforces cisgender normativity and erases non-binary and possibly also trans identities. While being aware of these important limitations, the superficial approach necessitated by the quantitative analysis of the large amount of data cannot avoid them.”

We also modified Lines 208-209: “included in the analysis if the certainty estimate provided by the tool was larger than 75% (a somewhat arbitrary but conservative choice)”

Lines 196-198: “Sixty seven percent of the participants were < 40 years old and 57.1% of the participants identified themselves as men, 41.4% as women, with 1.5% preferring not to respond.” Issues/Suggestions: Provide some text on what this non-response could mean (see Porter, J. R., & Ecklund, E. H. (2012). Missing data in sociological research: An overview of recent trends and an illustration for controversial questions, active nonrespondents and targeted samples. *The American Sociologist*, 43(4), 448-468.). Is there any pattern of non-response to other variables?

A/ To provide an explanation of what the non-response could mean, we summarize below (see table S2) the number of respondents who chose “prefer not to respond” as the

option for all questions where this was possible, we then identify if the % of no responses to specific questions is higher than the prevailing averages of no response.

We have now added this information to the revised manuscript in Lines 351-358 and table below (Table S2 now added to the supplementary material). “The highest missing data rates from “prefer not to respond” options came from questions about perception and experience (8-19%, n=15-35 from Qs 27, 31 and 47). These are rather qualitative questions in the survey and therefore could be a reason for the high missing data, if respondents were not confident that their perceptions of experience were in fact too qualitative for a response. In other words, the highest missing data were from questions that sought a yes/no qualitative opinion. The missing data for questions that may be considered private or intrusive is quite low (Table S2) with the highest being that about ethnicity (n=4, 2.2%), whereas the highest in the survey came from a question on perception of the increase in women in research past the PhD level (n=35, 19%).”

Table S2. Summary of missing data from questions in survey.

Questions	Number (n) of “Prefer not to respond” and % of 198 respondents
Q1. Please indicate your age	1, 0.5%
Q6.How do you identify your race/ethnicity?	4, 2.2%
Q7. Please select one of the following options related to how you identify yourself (Agender, Man, Nonbinary, Two-spirited, Woman, prefer not to answer, other)	1, 0.5%
Q8. Which of the following do you identify as? (Cisgender,Intersex, Transgender, Prefer not to respond)	6, 3.3%
Q9.What is your sexual orientation	10, 5.5%
Q27. In your opinion/perception, has the number of women past the PhD level (i.e. post-doc and above) increased in cloud physics research over the course of your career?	35, 19.1%
Q31. In your experience, do women get disproportionately assigned to academic/university/industry service work in order to represent their gender?	15, 8.2%
Q32.Do you consider yourself as part of a minority group?	7, 3.8%
Q35. Do you have any disability or accessibility needs?	3,1.6%
Q47. In general, do you consider students or colleagues from a minority/disability group are usually equitably treated in your working environment?	15, 8.2%
Q55.Have you ever experienced a hostile environment because of your gender/race/ethnicity/sexual orientation/disability?	3,1.6%
Q56.Have you ever encountered harassment or bullying by anyone from the cloud physics community?	6, 3.3%

Lines 284-287: "A higher fraction of non-White women reported experiencing bullying and harassment, which aligns with the concept of intersectionality, where overlapping social identities and positions can increase the likelihood of discrimination (Crenshaw, 1991). However, the sample size prevents statistical verification of this pattern."
Issues/Suggestions: I agree that the “non-White women” phrase indicates an

intersectionality issue. To support the statement about the sample size, include a parenthetical phrase (n=xx).

A/ We thank the reviewer for the suggestion and have now added to Line 284 original manuscript (Lines 449-450 of the revised manuscript) the sample size of non-White (now global majority) women (30.5%, n=36).

I am concerned that no power calculations are presented, nor are corrections for multiple rounds of testing. Also, in line 269, where the authors refer to respondents who reported disabilities, there should be some statement about how disabilities are underreported (Schneiderwind, J., & Johnson, J. M., 2021). Broadening the equity lens for STEM teacher education: The invisibility of disability. *AAAS bulletin*.)

A/ The reference to intersectionality was purely descriptive and interpretive, no formal intersectional analysis or statistical testing was performed. As such, no pre-registration or power calculations were applicable. The subgroup percentages are not statistically robust and are presented only to highlight potential patterns warranting future investigation with larger samples. We believe it is worth reporting because this has not been done for the cloud physics community before. The following text was added to acknowledge this.

Lines 455-457: "The reference to intersectionality is interpretive, it offers a possible lens for understanding, not a statistically proven explanation. Therefore, the reported percentages are to be taken with a caveat because of the small numbers."

The following text about how disabilities are underreported was also added.

Lines 434-435: "However, is it important to note that this number can be higher as adults with disabilities are typically underreported (Schneiderwind and Johnson, 2021)"

Lines 349-358: Discussion of women first authors having more women co-authors. Issues/Suggestions: The authors state, "This may not be the result of them being more equitable but could rather mirror a preference for working with people similar to oneself," but there is no longitudinal data to establish temporal precedence, no controls for confounders (field subspecialty, institution, career stage), or alternative explanations not tested (e.g., mentorship networks, discrimination by male-led teams).

A/ We agree with the reviewer that other factors could be at play, but we consider this analysis impossible with the bibliometric data that we have at our disposal and thus think that the suggested analysis is out of the scope of the present study. Note that we have consciously used careful wording ("may", "could") to indicate that ours is only a possible reasoning.

Lines 379-388: "The present data shows that in addition to the low number of studies published by tropical coauthors, on average, their indexed studies get published in lower-impact journals (impact factor IF=3.91, 95% confidence interval (CI): 3.72, 4.10 (n=287)) compared to their non-tropical peers (IF=4.58, 95% CI: 4.28, 4.88 (n=6638))." Issues/Suggestions: As stated above (Lines 157-171), there is potential

for significant biases introduced to these data by the exclusion of papers from the Global South and non-English language journals. Furthermore, there are limitations and biases in the concept of Impact Factor making it problematic for use in research (Seglen, P. O. (1998). Citation rates and journal impact factors are not suitable for evaluation of research. *Acta Orthopaedica Scandinavica*, 69(3), 224-229.)

A/ We agree that the exclusion of non-English papers could add a bias. However, as addressed above, their exclusion would not have a major effect on our conclusions as they represent 1% of the total number of papers. In case that including these papers was feasible, it is very likely that the shown differences will be even stronger as non-English journals have a limited international audience and therefore lower impact factors.

We agree with the reviewer that impact factor is not a perfect metric. Therefore, we would like to clarify that in this study such metric is not used to evaluate the quality of the research, but to determine the type of journals where the different groups publish their work.

Lines 444-460 (Figure 6 and discussion): The analysis of "years until last registered paper" as proxy for leaving the field. Issues/Suggestions: This assumes the last publication is equivalent to leaving field, but it could also be career pauses/shifts (Researcher to Admin), database lag, non-Scopus publications. There is a bias toward younger researchers where those who have entered the field recently (2015-2020) have been publishing dissertations and for tenure review, meaning they are a maximum productivity and have had less opportunity to leave. Include a survival analysis or time-to-event test given the temporal data structure and compare retention across time periods with different follow-up durations.

A/ We thank the reviewer for this suggestion. Regarding caveats, we agree that database lag and non-Scopus publications present inaccuracies for absolute numbers, and also that young researchers may be publishing more. However, none of this affects the comparison between men and women. Career pauses present no problem for the analysis, as long as the person has ended their pause and published again, and shifts are considered leaving the scientific field. Since the focus of the analysis is the comparison between men and women, we consider the statistical test that we added in response to your next comment sufficient.

Line 453-455: "However, in recent years, after 1995, we see no difference in years between the genders in how long they remain in the cloud physics field judging by publication activity." Issues/Suggestions: This conclusion is unsupported. A visual inspection of Figure 6 indicates no statistical test. Also, this does not account for different entry cohort sizes.

A/ We have added a statistical test for the difference in gender distributions to the Figure, which we explain in the caption. This test takes the different entry cohort sizes into account. Note that the y-axis gives the relative frequency, so that also here the difference in entry cohort sizes is considered.

Editor comments:

Synthesis of **reviews** of, and **revision** suggestions for, egusphere-2025-**4499**, Ladino et al, **The state of diversity, equity, and inclusion in the cloud physics community**, by David Crookall, Handling Editor for *Geoscience Communication*.

All four reviewers state that this is an important article, but one that needs to be revised. My view is that your article is extremely important, and that it needs some considerable revision, even more so as the topic is so crucial, not just for the cloud phys community (CPC), but for all of geoscience. Indeed, I would argue that no science, particularly geoscience, can achieve its rightful fulfilment and do its work satisfactorily if DEI is not also a foundational pillar of the science. Fully-implemented DEI needs to be achieved if the scientific aspirations are to be achieved.

Before I get to the more specific reviewer comments, I would like to interject some suggestions. None of the following needs to take more than a short paragraph each, unless you think that it would need more space. I suggest that you situate your DEI work within:

- the broader areas of DEI in the geosciences more broadly (eg, has DEI research been conducted in other-than-CPC fields?);
- the broad area of DEI research in general;
- the general failure of society to uphold DEI in any meaningful manner; and
- the literature and knowledge of geoethics, injustice, human rights and sociology. You probably know that the EGU now has a Working Group on Geoethics – see <https://www.geoethics.org>.

A/ Thank you for these suggestions. We addressed the four points by adding/modifying the following paragraphs in the revised manuscript.

Lines 30-32: “Rather we should consider an equity-centered approach by understanding our ethical responsibilities to benefit research of the climate system, as it is promoted by the International Association for Promoting Geoethics (IAPG) in Geosciences.”

Lines 47-53: “Although various programs, policies, and inclusive environments are being developed to address issues such as gender inequality in academia (Ranganathan et al., 2021; Bernstein, 2024), many gaps in other historically excluded groups such as Hispanic, black, indigenous communities, scientists with disabilities, and LGBTQIA+ scientists still remain (Bernard and Cooperdock, 2018; NCSES, 2019; Gewin, 2020; Odekunle, 2020; Berhe et al., 2022; Marin-Spiotta et al., 2023). Thus, we must recognize that developing a diverse, equitable, and inclusive workplace is a very challenging task (e.g., Apaari, 2021; Reisch and Jani, 2025) and requires considerable reform of systemic deeply rooted institutionalized practices.”

Lines 54-60: “Many fields lack diversity in gender, ethnicity, disability, and language including health (Wenneras and Wold, 1997; Landman and Dandolu, 2009; McDermott et al., 2018), science, technology, engineering and mathematics (STEM)(Nadis, 1999; Watt,

2005; Haak, 2002; Barber et al., 2020; Ceci and Williams, 2011; Leslie et al., 2015), as well as humanities, in particular philosophy (Nielsen, 2016; Botts et al., 2014; Herfeld et al., 2022). The lack of diversity, equity, and inclusion in academia is one of the manifestations of a system originally structured to privilege dominant groups and exclude marginalized populations, a pattern that has historically characterized most fields in science and that continues to influence their institutional practices”

Lines 89-91: “Even though DEI has been evaluated in geosciences as a whole, there are very few studies focusing on specific geosciences subfields such as atmospheric, oceanic, solid Earth, aquatic, planetary, and environmental sciences, among others (Dyess and Teplitzky, 2022; D’Elia, et al., 2023; Pruitt et al., 2023, de Vos et al., 2023).”

Possibly the greatest damage of no or low DEI is to be found in gender discrimination and sexual harassment; about which much has been published, but for which little improvement seems to have been achieved. In many ways, discrimination and lack of DEI ‘begin life’ in the fleeting verbal and non-verbal exchange in everyday settings, which is rarely recorded, except for research purposes. Underlying this are attitudes, identities and attributions derived from intergroup behaviour – see the groundbreaking work by Tajfel, Turner, Billig, Hewstone, Giles and colleagues. More recent references include those placed at the end of this synthesis. Towards the end, it would be useful to have a section discussion directions for future research.

A/ The discussion for future directions can be found in Lines 623-658: “Increasing inclusion and diversity is essential to promote fairness and equal chances for all authors.

Furthermore, expanding data coverage in the tropics is critical for a better understanding of the Earth’s climate. As a result, the scientific community must collaborate to address these gaps and encourage fair collaboration and opportunities for all scientists in order to increase the quality of scientific research. Finally, fostering a more inclusive and welcoming environment within the cloud physics community will benefit not only individuals who are currently underrepresented, but it will also help to foster a more productive and collaborative research culture that can better address the complex challenges facing the field. Lastly, the following recommendations could provide a guideline on how to achieve a more diverse, equitable, and inclusive community:

- Provide institutional resources for diversity and inclusion training that align with the ethical responsibilities to be adhered to at all levels, from undergraduate students to senior professors. This should extend beyond just recruiting, but be available to facilitate the careers after recruiting.
- Promote, facilitate, and secure the participation of graduate students and early careers scientists from the Global South in the most important cloud physics academic events that take place annually, especially those organized by the ICCP (i.e., workshops, conferences, etc.)

- Persons in reviewer roles for conferences, funding applications, and peer-reviewed articles should receive training on how to avoid unconscious bias in treating submissions from women, scientists from historically excluded groups and from tropical countries fairly, solely focusing on the science.
- Seek a gender balance and pursue individuals from historically excluded groups for faculty positions, editorial boards, scientific committees, awards, and review panels.
- Foster a culture of inclusivity in which scientists of all backgrounds experience a sense of belonging (instead of bullying or harassment). This can be achieved by promoting respect for diverse opinions and increasing the visibility and retention rate of diverse opinions and ideas.
- Journals should implement policies to avoid the practice of parachute or helicopter science and promote fair collaborations where support for colleagues from the Global South prevails.
- Increase the number of submitting proposals led by scientists working in a tropical country in collaboration with Global North research groups. Building partnerships and collaborations across institutions and regions can help break down silos in science and promote diversity of thought.
- Funding agencies should prioritize projects conceived and led by researchers working in a tropical country and must provide partnership, mentorship and supervision programs that encourage underrepresented groups in science, which aim to build resources in economically disadvantaged countries.
- Foster a sense of community which can be a powerful motivator for collaboration and engagement, which can be achieved through events and activities that bring different scientists together, such as workshops, seminars or social events.
- Support and promote DEI initiatives such as the Latin America Early Career Earth System Scientist Network (LAECESS, Yáñez-Serrano et al 2022) or the *Directions for future research and applications on clouds, precipitation, and extreme weather in Africa* workshop (<https://www.cloudsfor.africa/>)."

In your rebuttals, you have already agreed with many of the reviewers' suggestions. I appreciate very much your effort to take reviewers' comments on board. As a result, what you propose as changes, will likely make a great improvement to the article as a whole. All four reviewers produced excellent reports that provide you with some great help to improve your MS in a significant manner. It would be good, if you felt like it, to thank the reviewers for their tremendous help, even though they chose to be anonymous. (If you also wish to pass on your thanks in a more personal way, I am happy to do that for you. Just send me an email.)

A/ This is a great suggestion. We have added the following text into the Acknowledgements section: Lines 669-670: "We thank the four anonymous reviewers and the Editor (Prof. David Crookall) for the time spent in reading and reviewing our study. Their suggestions and comments helped us to significantly improved the manuscript"

Finally, congratulations on tackling such an important and complex area. If you work hard on revising the article, it could become a major reference in the area (good for you) and contribute strongly to making the situation better for many 1000s of (geo)scientists suffering from a lack of DEI and aspiring for a better world (good for others). After all, they know what it is like, they have been for far too long on the receiving end of injustice that they do not deserve. In that sense, your article is more important than a piece of research; it could and should be a catalyst for positive change.

A/ Thank you for these encouraging words.

Below, I will attempt to synthesize the reviewers' comments in the expectation that this will also help you to revise your article. In the table below, the first two columns are self-explanatory. Col 3 indicates which reviewer, 1, 2, 3 and 4, as indicated here <https://egusphere.copernicus.org/preprints/2025/egusphere-2025-4499/#discussion>. I assume that the numbers will be the same as yours. The last column is a rough indication of how 'sever' the problem is and thus how important it is to revise. 1 is low, 3 is high.

My comments below are based mostly on the four reviews received, but also include a few more personal views. Some revisions are very important and others less so; I have tried to indicate what I consider the importance of each suggested revision. Also, some revisions do not appear to be possible at this stage, and they should be noted in a section titled 'limitations' and/or discussed in a section on 'implications for future research'.

A/ Thank you for this suggestion. As we mentioned in the answer to Reviewer #4 several statements were added to the revised manuscript. However, in addition to this, we added Section "3.6 Limitations" where we summarized the most important limitations from the present study:

Lines 578-590: **"3.6 Limitations.** Although the present study reached important conclusions to understand how diverse, equitable, and inclusive the cloud physics community is, there are different limitations that inhibit us from going deeper into our analysis and that did not allow us to reach more robust conclusions. For example, the number of the survey respondents is small (n=198), and hence, some of the conclusions need to be taken with caution. On the other hand, in the metadata analysis, non-English publications (n=97, 1.18% of the publications) were included, which could have biased the conclusions related to parachute science as well as those for authors with tropical affiliations. It is also important to note that the usage of genderize.io to assign the gender of the authors is only a good approximation, it is binary, and it has some bias. Additionally, as IF was the only used journal metric to evaluate the access of the authors to high profile journals, and as it has been highlighted as a flawed metric of the performance of individual researchers (Editorial, 2013), the obtained conclusions could be misleading. Finally, as the authors could have taken career pauses and/or shifts, and this may be different between men and women, the analysis shown in Section 3.5 could be slightly bias."

Problem or issue	Possible remedy, but you decide if and how much to implement	Reviewers	Severity	Answer
Small sub-samples. $N=198$ seems rather small for reliable sub-group analysis (gender/ethnicity/region).	You cannot change the numbers, so maybe include a cautionary note in your text and/or discuss in limitations.	1, 3, 4	2	A cautionary note was added to the revised manuscript. Lines 581-582: “the number of the survey respondents is small ($n=198$), and hence, some of the conclusions need to be taken with caution. ”
Overstated claims. Disagreement with claims of "global representativeness" (30–47% of the community).	Reword.	3, 4	2	The sentence was modified as follows: Lines 159-161: “This brings us to a percentage coverage of 30-47% normalized to the attendee numbers from the three previous conferences and 21.1% as the lower limit of the response rate, calculated as a percentage of the listserv which is 198/938.”
Vague questioning. Terms like "feel included" or "sensitive to gender issues" are poorly defined and open to interpretation.	Reword.		1	We agree that this is an open-ended question and not well defined. However, what it feels to be included means different things to different people, in different situations and we wanted to encompass the individual feeling of inclusion within the context being the cloud physics community, regardless of the locations being the workplace, education or social situations. We wanted respondents to answer based on their perceptions of feeling

				included and not define different spheres of inclusion a priori. We believe it is a valuable finding as it speaks to the sense of belonging (or lack thereof) from the respondents. We acknowledge this limitation in the following statement. Lines 392-393: "It is important to note that the feeling of inclusion means different things to different people. Therefore, these results must be interpreted with caution". Regarding "sensitivity to gender issues", we did not want to specify specific gender issues because everyone's perception of sensitivity and gender related issues are subjective. In general, it is meant to give an overview if work places of the respondents are supportive with regards to different genders working in the same space. So, we believe that this question provides a good overview of gender related issues in the workplace without delving deeper into specific details.
Response rates. Denominators are missing; unclear how many actually received the survey vs. responded.	Indicate if at all possible, or elaborate on why it is not possible.	3	2	The survey was sent to a maximum of 938 email addresses (as of 2023); however, there is no way of verifying that all these email addresses were active

			because of many retired past members, and many people who have passed away. In addition, we do not have the exact numbers for 2021, since the listserv was migrated to a different server so the earliest numbers we have are for 2023. Thus, we report that the upper limit of the listserv numbers is 938. We have now added this information to Lines 160-161 of the revised manuscript. For the 2021 meeting in Pune there were 570 attendees which we have already indicated in the original manuscript (previously Line 137, now Line 159). We believe comparing the response rates to the past three meeting attendee numbers is also useful because it captures the active participation of respondents in the survey. However, it does bias against those who could not attend due to schedule conflicts, distance, pricing. This is why we do not use only the 2021 numbers to calculate the response rates, but also use numbers from the 2012 and 2016 meetings which were not in Asia. This brings us to the % coverage of 30-47%. This information is now included in addition to the lower limit of the response rate calculated as a
--	--	--	--

				percentage of the listserv which is 198/938 = 21.1% (Lines 160-161 of the revised manuscript).
Anonymity anomaly. You say that you know specific regional trends (e.g., Indian respondents), but also that your data is anonymous.	Not sure about this one. Maybe a quick note to explain, eg, you may have collected names, but they are kept secret?	3	1	We were able to reach this conclusion because we have the data on ethnicities in the survey, so we can calculate the percentage of ethnically Indian respondents which was 14% (n=27). Furthermore, when we compare the 2021 meeting figures (28% n=160) to the 2012 (3% n=20) and 2016 (1%, n=4) meetings it becomes clear that the Indian respondents had a higher percentage representation than their percentage presence at previous conferences. This is explained by the fact that the conference was indeed held in India in 2021 making it more accessible to the local participants.
Stat Assumptions. Chi-square tests used where cell counts are <5; lack of Fisher's exact tests or power calculations. "A standard rule of thumb credited to Cochran advises against using the Chi-square test when more than 20% of cells have expected frequencies below 5, or when any cell has an expected frequency below 1". https://psychology.town/statistics/key-assumptions-chi-square-test/	Best option: Redo the stats. Use Fisher's exact test for small N; report effect sizes and power calculations. It may well be that the indication turns out to be (almost) the same, in which case not too much to change in your text. If the exact test shows very different results, then you will have to change your text.	4	3	We started with broad descriptives to set the scene, show overall trends, and highlight groups that might be worth a deeper look. For example, finding that 30% of respondents don't feel included prompted us to break the data down further by age, gender, and identity. We then tested meaningful comparisons, but only when group differences were large enough to warrant inference (for instance, running a χ^2

				test on gender vs. career impact where the result was significant. Finally, we qualified our limitations by explicitly stating when a small sample size prevented formal testing. For example, with only 15 LGBTQIA+ respondents, no statistical test was run, though the finding was still noted descriptively and considered alongside external context. To elaborate a bit more: - Descriptive reporting (e.g., Lines 230–237 in the original manuscript): When presenting subgroup proportions without formal hypothesis testing (e.g., “35% of those <40 did not feel included”), we were highlighting observed patterns for qualitative or exploratory context, not making inferential claims. These were not intended as statistical tests and were explicitly framed as descriptive. - Statistical testing (e.g., Lines 244–245, 277–278, 291–293 in the original manuscript): We applied chi-square tests with effect sizes (Cramér’s V) only when comparing two or more groups on a categorical outcome where a meaningful association
--	--	--	--	--

				was hypothesized (e.g., gender vs. perceived career impact). This was done to assess whether observed differences were likely due to chance. - Explicit “not statistically significant” statements (e.g., Line 257–258 in the original manuscript): These were included only when a formal test was conducted and the result did not reach conventional thresholds: to avoid overinterpreting non-significant trends. - No testing for small subgroups (e.g., LGBTQIA+ N=15): We intentionally avoided statistical testing for very small groups (e.g., $n < 20$) because chi-square tests are unreliable with low cell counts. Instead, we presented these as descriptive observations, often with reference to external literature (e.g., Le Bras, 2021) to contextualize them. Therefore, we consider that there was no arbitrary “decision rule” rather, a principled distinction between descriptive illustration (for exploratory or small-sample observations) and inferential testing (for
--	--	--	--	--

				comparisons with adequate power).
Metadata exclusion. Excluding 97 non-English papers silences the Global South (or a part of) the very population being studied. It may well be that non-EN authors suffer more from lack of DEI than their EN counterparts, simply because lack of EN is a disadvantage in science and publishing. If your topic was different (eg, cloud formation, not DEI), then this would matter less.	The most satisfactory option would be to include these and redo your stats, maybe comparing EN and non-EN authors, but this may entail a large and/or lengthy process, which would be difficult at this stage. The most realistic option would be to include an explanation. In my view, claiming that translation is not possible would be seen as a feeble excuse, esp given that translation tools are easily available on the web.	4	3	We agree with the reviewer and the Editor that including those 97 papers would have been ideal; however, this would require hiring several people to either translate the documents or to corroborate that the online translation software did a good job, which was unfeasible for our study that had no dedicated funding. Since these papers present about 1% of the total number of papers, their analysis would not have a major effect on our conclusions. The following sentence was added to the revised manuscript. Lines 582-584: “On the other hand, in the metadata analysis, non-English publications (n=97, 1.18% of the publications) were included, which could have biased the conclusions related to parachute science as well as those for authors working in a tropical country”
Attrition proxy. "Years since last paper" will be seen a weak proxy for leaving the field; it does not account for career pauses (more prevalent for women), long projects, different pressures/needs to publish, resources for ACPs or database lag.	This issue certainly weakens any statistical inference that you can draw, simply because the initial data are unrepresentative. If you have no way in which you can tighten up this indicator, then it would be a good idea to explain this and its impact on your results in a limitations section?	4	2	The following text was added to the revised manuscript. Lines 588-590: “Finally, as the authors could have taken career pauses and/or shifts, and this may be different between men and women, the analysis shown in Section 3.5

				could be slightly biased”
Inconsistent reporting. Missing effect sizes, ($n=xx$) values, and clarity on descriptive vs. inferential stats.	This seems important to specify if you possibly can.	3, 4	3	n values and the missing answers from the survey are now provided in Table S2
Data source bias. Scopus-only search inherently limits Global South visibility. Impact factor (IF) use is biased.	Limiting visibility by data source could skew your results and analysis. If you can bring in other sources, that would be good. Otherwise mention in limitations? IF is strongly contested (see, eg, https://occamstypewriter.org/scurry/2012/08/13/sick-of-impact-factors/). Is there any way in which you could bring in one or two additional indicators? If not, then a discussion about how the use of (solely) IF, is only partially representative but provides an idea of tendency.	4	2	We agree with the reviewer and the Editor that impact factor is not a perfect metric. We added a text to acknowledge the IF limitations. Lines 586-588: “Additionally, as IF was the only used journal metric to evaluate the access of the authors to high profile journals, and as it has been highlighted as a flawed metric of the performance of individual researchers (Editorial, 2013), the obtained conclusions could be misleading.”
“Tropical”. Conflates geography with socio-economics; described as exoticizing and a “false dichotomy.”	You suggest “scientist working in a tropical country”. It might be less unwieldy to use an expression like “scientist working in the tropics”. In any case, it would be important to explain clearly what this category includes. Does it include both westerners working in the tropics and tropical locals working there? My question then is: Would it not make sense to distinguish between these two groups, as they probably come from two different cultures or cultural backgrounds (and culture plays a major role in our live (see, eg, E T Hall, F Trompenaars, Y Y Kim, G Hofstede, T Toomey), and thus in our DEI activity.	1, 2, 3, 4	3	The Editor is correct that cultural backgrounds play a big role in our lives, and hence, in the way we could see DEI. However, these differences also stand out between tropical countries (e.g., Brazil vs. Kenya vs. Thailand vs. Bangladesh, etc.). Therefore, we are not 100% sure if making a distinction of the background of a scientist working in the tropics could add important information. The term “tropical scientist” was changed to “scientist working in the tropics (anyone from whatever country

				who lives and works in a tropical country).” Lines 103-104
"Minority". Imprecise term; reviewers prefer "marginalised", "underrepresented".	You say that you will use the term "historically excluded". However, the presupposition in that term is 'not included any more', 'outside'. I would think that 'marginalised' would be more accurate. However, if you wish to use your term, then it would be important to specify that the exclusion is still ongoing today. Maybe a term like "historically and currently excluded", but then maybe simply "excluded", but that is stronger than "marginalised". Excluded implies that they have been banished to outside, whereas marginalised implies they have been pushed to the side, but are still in to a certain extent. Maybe a short lit rev of these terms is warranted? Maybe use the two terms for two distinct groups?	1, 2, 3, 4	3	"historically excluded" was suggested by reviewer #3 ("another alternative would be to say "historically excluded" groups. This is a better descriptor that is commonly used and recognizes the fact that these groups have been intentionally excluded by the system, as opposed to just individual actions"). In the revised version we used "historically and currently excluded" the first time and then we kept "excluded".
White-default binary. "White vs. non-White" grouping obscures intra-group diversity and centres whiteness, reinforcing white supremacy frameworks.	I have had misgivings about the term 'non-white' for quite a while now. What do Africans say? 'Non-black'? Here are some suggestions: dark-skinned and clear skinned, "People of Colour (POC)", 'Black, Indigenous and People of Colour (BIPOC)', "Racialized groups". It might be worth asking a DEI specialist. Maybe best to use groupings like East Asian, South Asian, Indigenous / Native American, Hispanic/Latino, Middle Eastern/North African, White.	4	2	Instead of non-white, in the revised manuscript we use the term people of global majority (POGM) or simply global majority (GM) coined by Campbell-Stevens (Campbell-Stevens 2009; 2021a), which acts to de-center White as the reference and encompasses Black, Asian, Brown, Indigenous, and mixed-heritage people, comprising approximately 80–85% of the world's population (Campbell-Stevens 2021b) highlighting their status as the global majority. We have now changed

				all figures and text to change non-White to (PO)GM.
Geographic error. Misclassifies Australia (Global North) as "Tropical" and ignores subtropical Egypt/South Africa.	I should think that some small tweaks would correct that.	2	1 or 2	The following text was added to the revised ms, Lines 218-220: "For a country to be considered tropical in this study, at least 50% of its inhabitants need to live in the tropics, i.e., between 23.5° N and 23.5° S and be considered a low-income country, with the exception of Taiwan, Singapore, and Hong Kong (See Table 1)."
Gender essentialism. Strictly binary (men/women) erases trans and non-binary individuals. Bias of genderize.io.	I am not sure what you can do about this, short of redoing all your data and stats. Probably best to provide an explanation in a limitations section.	1, 2, 3, 4	2	The following text was added to the revised ms. Lines 210-215: "It is important to note that genderize.io assigns a binary (female or male) gender based on the person's first name. This can only be a good approximation, as names cannot always be used to infer the gender (e.g., Sebo, 2021; Lockhart et al., 2023; Marty et al., 2023). Critically, the binary gender assumption inherent in this methodology reinforces cisgender normativity and erases non-binary and possibly also trans identities. While being aware of these important limitations, the superficial approach necessitated by the quantitative analysis of the large amount of data cannot avoid them."

Systemic framing. Lack of clarity concerning your use of ‘bias’.	Maybe you can reframe "bias" as intentional structural design, rather than just skewed individual cognition. But best to look at some of the extensive literature on bias, especially intergroup bias, eg, Fiske, S. T. (2015). Intergroup Biases: A Focus on Stereotype Content. <i>Current Opinion in Behavioral Sciences</i>, 3, 45–50. https://doi.org/10.1016/j.cobeha.2015.01.010. Grigoryan, L., et al (2022). Multiple categorization and intergroup bias: Examining the generalizability of three theories of intergroup relations. <i>Journal of Personality and Social Psychology</i>, 122(1), 34–52. https://doi.org/10.1037/pspi0000342. Tajfel, H., et al (1971). Social Categorization and Inter-Group Behavior. <i>European Journal of Social Psychology</i>, 1, 149–178. https://doi.org/10.1002/ejsp.2420010202. In-group favoritism. (2026). In <i>Wikipedia</i>. https://en.wikipedia.org/w/index.php?title=In-group_favoritism&oldid=1354643693. Plenty of others too.	3, 4	2	The following text was added to the revised ms. Lines 210-215: “It is important to note that genderize.io assigns a binary (female or male) gender based on the person’s first name. This can only be a good approximation, as names cannot always be used to infer the gender (e.g., Sebo, 2021; Lockhart et al., 2023; Marty et al., 2023). Critically, the binary gender assumption inherent in this methodology reinforces cisgender normativity and erases non-binary and possibly also trans identities. While being aware of these important limitations, the superficial approach necessitated by the quantitative analysis of the large amount of data cannot avoid them.” and Lines 314-320: “Research practices in which scientists from wealthier countries conduct studies in lower-income countries, often without meaningful collaboration or involvement with local researchers or communities and then publish the results without authorship acknowledgement or benefit to the locals, have been termed helicopter or parachute science. These practices have a colonial and
---	---	-------------	----------	---

				extractivist legacy and may be perpetuated by the belief that Global North institutions have knowledge to be transferred to partners in the Global South (Gewin, 2023; Stefanoudis et al., 2021; Odeny and Bosurgi, 2022; de Vos and Schwartz, 2022, and references therein). Our results suggest that parachute science is unfortunately common in the cloud physics community”.
Post-hoc intersectionality. Intersectionality is mentioned in passing rather than serving as the study’s foundational framework.	At this stage, I doubt that you could change this to make intersectionality the main framework of your study. If you had strong reasons for avoiding intersectionality, then they should be mentioned. Maybe (for future researchers) you can speculate briefly on what difference it would have made if you had done your study fully with an intersectional approach.	4	2	The following text was added to the revised ms. Lines 455-457: “The reference to intersectionality is interpretive, it offers a possible lens for understanding, not a statistically proven explanation. Therefore, the reported percentages are to be taken with a caveat because of the small numbers.”
Ability/disability. This fails to address the overrepresentation of able-bodied individuals or systemic barriers for disabled peers.	I am not sure how important this is. Maybe it does not change any of your results/stats?	3	1	The following text was added to the revised ms. Lines 54-57: “Many fields lack diversity in gender, ethnicity, disability, and language including health (Wenneras and Wold, 1997; Landman and Dandolu, 2009; McDermott et al., 2018), science, technology, engineering and mathematics (STEM)(Nadis, 1999; Watt, 2005; Haak, 2002; Barber et al., 2020; Ceci and Williams, 2011; Leslie et al., 2015), as well as

				humanities, in particular philosophy (Nielsen, 2016; Botts et al., 2014; Herfeld et al., 2022).”
Parachute science.	You should to explicitly link "parachute science" to colonialism and cite FAIR data practices. Is this possible in your ms; I imagine yes.	4	2	We added the following text. Lines 314-320: “Research practices in which scientists from wealthier countries conduct studies in lower-income countries, often without meaningful collaboration or involvement with local researchers or communities and then publish the results without authorship acknowledgement or benefit to the locals, have been termed helicopter or parachute science. These practices have a colonial and extractivist legacy and may be perpetuated by the belief that Global North institutions have knowledge to be transferred to partners in the Global South (Gewin, 2023; Stefanoudis et al., 2021; Odeny and Bosurgi, 2022; de Vos and Schwartz, 2022, and references therein). Our results suggest that parachute science is unfortunately common in the cloud physics community” and Lines 331-333: “Although there is not a clear way to eliminate these practices, as highlighted by Ramírez-Castañeda et al. (2022) the following four

				principles should stand out when performing field work: be collaborative, be respectful, be legal, and be safe.”
Historical context. This ignores the disciplinary origins of cloud physics and the geopolitical shifts over the 50-year metadata period.	I would be really good if you can include something on this. After all, all our social interactions and decisions take place in the context of historical developments.	3, 4	2	The following text was added to the revised manuscript. Lines 114-117: “It is important to note that the cloud physics community is a multidisciplinary field where meteorologists, physicists, biologists, oceanographers, chemists, environmental engineers, and chemical engineers, among others, work together in cloud and climate modelling, field, theoretical, and laboratory studies which implies variabilities in issues related to DEI”
Literature review	Both 3 & 4 suggest that you update with sources from the last five” years, and include disability and Berhe et al. (2024). See also the biobibliography that I included below.	3, 4	2	The following 22 recent studies were added to the revised manuscript: Ranganathan et al. (2021), Schneiderwind and Johnson (2021), Stefanoudis et al. (2021), Berhe et al. (2022), Hall et al. (2022), Castro et al. (2022), de Vos and Schwartz (2022), Odeny and Bosurgi (2022), Ramírez- Castañeda et al. (2022), Son and Bell (2022), Marin-Spiotta et al. (2023), Bernstein (2024), Almouzi (2024), Feijó and Orre (2024), Rubbia (2025), Brown and Laihonen (2025), Dovchin (2025), Godrie (2025), Metcalfe et al. (2025), Titirici et al. (2025), Arosoaei et al.

				(2026), and Atilas et al. (2026).
Visual errors (Fig 1)	Fix the colour palette – do not use the same colour (purple) for two different data categories.		2	Original Figure 1 (now Figure 4) was fixed.

Some additional comments as I looked through your ms.

Graphics. You should spruce up your graphics. Here are a few suggestions.

You will find a range of excellent graphics in this article: Leisner, M., de Paula, D., Vasconcelos, Y. et al. Coastal cliff research trends and future directions. Discov Geosci 4, 6 (2026).

<https://doi.org/10.1007/s44288-025-00370-9>

A/ We have used brighter colors for the graphics that are still colorblind safe. However, we keep the current graphics largely the same, because we find the way we have prepared the figures to convey the messages we discuss in the paper. We looked at the graphics in the paper and found some useful information, however, the data from our metadata analysis and survey does not allow us to produce similar graphics to the Leisner et al. (2026) paper.

Your fig 1. A). Brighten up the colours. It would be clearer, in my view, if you grouped by type of person, eg, three histograms of women, then men, then all. Each of the 4v or 5 -not clear which) graphs should have a title. Did you discuss the reasons for the prominence of the four 'big' countries? Put women before men in col sequence! Try and show to a colleague to see what they say.

A/ We have attempted to brighten up the colors, but yet are restricted by color blind safe options as required by Copernicus publications. We need to follow the color-blind safe options available from our plotting software, given this is a paper on DEI.

Having separated histograms for man, woman, and all, stratified by age, as was done in Figure 1a is less preferred. The reason is that in the way we plotted it now it is easy to compare the numbers in each age group of men vs. women and what fraction of the total respondents the binary genders form. If we separate into different histograms the comparison across gender for each age category becomes harder to appreciate. In the current version one can easily compare women across age categories, men across age categories as well as men and women across age categories. In addition, one can compare the total respondents across age categories to easily see that the age demographic of the entire survey is. Lastly, we would like to point out based on the editor's comments we are including tables summarizing the numbers in the figures and the data from all figures will be uploaded to a persistent DOI - as such the interested reader can indeed replot the data with different styles as they see fit. We acknowledge that there are many ways of plotting the data that favor highlighting different aspects, (gender, age, ethnicity etc). We have primarily focused on the gender quality and differences in ethnicities which is why we have chosen the current form of display.

These plots have been presented internally at various group meetings and presentations and have evolved to the current stage through such feedback and review.

In Fig. 1 (now Figure 3), the graphs do not have a title - they are biographic data as such are self-explanatory. At most the X-axis have titles which have been indicated in the now Figure 3a/b.

We have already added in the original version of the manuscript why there is a big presence from India. This is covered in the manuscript in section 2.1 (Lines 161-163). Here we explain that there was a higher participation rate from Indian scientists as the conference where we distributed the survey was held in Pune, India.

We now added a discussion in the revised manuscript based on the editor's remarks on the prominence of participants from the "4 big countries". This has been added to Lines 164-166: "Germany, UK, and USA also have a noticeably higher participation rate because there is a large research presence of cloud physics historically from these countries, with established cloud chamber facilities."

Fig 2. A) Same as above. If you used the term 'feel included' in your questionnaire, then put in quote marks in the graph. B) C) & D) Same – group by women (yes & no in one column), then by men (in a second col). Easier to compare visually. Title should be **your**, not you. It is unclear to what graph each heading belongs. Best to have no more than two related graphs in a figure.

A/ Regarding the colors: We have attempted to brighten up the colors, but yet are restricted by color blind safe options as required by Copernicus publications. We need to follow the color-blind safe options available from our plotting software, given this is a paper on DEI.

Again, we would like to place focus on the comparison between man and woman, as such we stratified each of the "yes" and "no" responses by man and woman. If we plot the woman and men and stratify by yes and no, we focus more on how women responded to yes and no, rather than how man vs. woman responded to yes/no. So, we prefer to keep the bar chart as is in old Figure 2 (Revised version Figure 4). In addition, given that the editor has asked us to summaries all the data from the figure in tables (which is going to be included as a persistent DOI upon acceptance of the paper, in the data availability section), this data can be easily plotted differently for a specific reader's interest. We have focused on the gender inequality here within ethnicity and therefore prefer keeping the plots as they are.

The entire question as presented in the survey is given on the graphs. We have now added quotation marks to indicate that the questions presented on each graph are indeed quoted from the survey. The positions of the titles are explained in the figure caption. This is necessary so we can indicate the answers to the sub questions as well, which are important for the reader of the paper. We have repositioned them slightly to make it clear visually which graph the questions belong to. But again, the position of the questions relative to the figure is clearly mentioned in the figure caption.

The title should in fact be “you”, not “your”

For E) did you include – check on – how respondents’ answers stand up to reality? Eg, Was their, whoever (‘they’ is), report of low citations supported in the data and stats? Ie, did they over or under estimate the effect on citations?

A/ From the metadata analysis, we know that authors affiliated with a tropical country overwhelmingly publish in lower impact factor journals as discussed in section 3.1 (Lines 298-301). This implies that they also face lower citation indices, which supports the survey data, where people who identified as minorities (POGM) felt that one negative effect of being a person of global majority is lower citations (figure 2e).

Fig 3. Again I would regroup by white women, non-white women, white men, non-white men. It might be that some categories bully more than others, eg I would suggest: biggest bullies → white men – non-white men – white women – non-white women → least bullies. Cut you need to check the literature to see if that is correct. Then that raises that question of whether your sample reflects the wider population. In any case, you should change the category names to better ones than white/non-white (as mentioned by the reviewers).

A/ We agree and have changed non-white to “people of global majority” and supported this change by literature on this topic in the revised manuscript. Regarding the grouping, we do not have information on bullies as the editor suggests but people who experienced bullying.

We have included quotation marks around each of the 4 questions addressed in this figure to indicate that the question is identical to that from the survey.

The grouping as is now, shows how within the “white” community women and men fare and same for the POGM community. The comparison across the same gender but different ethnicity is much easier seen in the current format of the plots. As the focus is the comparison between men and women for a given ethnicity we prefer to keep the plots as is.

As the data will be made available through a persistent DOI for all the plots presented in this paper, it would be very easy for anyone interested to replot this in a different way. However, we have designed our discussion with a focus on gender within ethnicities which was a big push to call for gender quality in the cloud physics field. We therefore think this approach of plotting directly benefits the readers of the very relevant gender equality questions. And does not preclude anyone from visualizing this differently given the data will be open access.

Fig 4. Do you have figures for ‘first author woman tropical’, ‘first author man tropical’, etc? Would they be worth showing on a second graph? Do you have such figures for geoscience as a whole, and for science as a whole? It would be good to compare in a series of graphs these various lines.

A/ We have modified the figure by adding a new panel with the information for first authors working in a tropical country. We could not find comparable literature data for geosciences or science as a whole.

The following text was added to the revised manuscript. Lines 285-287: “Although there are differences in the gender first authorship (Fig. 1b) in the articles published by author working in a tropical country, the gender gap is smaller when no geographic distinction is applied (Fig. 1b).”

Fig 5. I understand the general idea of what you are doing here. However, it is not crystal clear for me. Can you see if you can make these clearer. I would separate c) and put it in another Fig.

A/ We apologize for this, but we do not know how we can make the figure clearer. We have divided the figure in two (Figure 2 and Figure 3), following your suggestion.

Fig 6 needs to be made much clearer. Maybe separate women and men, and juxtapose the two graphs. X axis is labelled “Years until last registered paper”. I do not know what this means. Years from when? What is meant by “registered”? Rewrite the caption, maybe include an example, eg, take one or two data points and say what they really indicate.

A/ We have changed the X-axis to “Years between the first and last published paper of a given author”. Additionally, an example was added to the revised figure and the following text was added to the figure caption. Lines 574-577: “An example is given in the fifth panel, highlighted with a dark green edge: of the 30 women co-authoring their first publication, i.e. entering the field between 1990 and 1995, 10% stayed in the field for 17 years (the years until their last co-authored publication).”

Put your final recommendations in a whole new section. Maybe from time to time indicate the place or result in your study which gives rise to a particular recommendation. In other words, justify your recommendations, especially by your results. They sound excellent, but the reader will want to know what motivates you to make, and what underlies, the recommendation.

A few more – sorry.

L 31. You mention ethical. It would be excellent to bring in the geoethical dimension or implications of your already excellent work. See eg <https://www.geoethics.org/>.

A/ Thank you for the suggestion. The following text was added to the revised manuscript. Line 32: “as it is promoted by the International Association for Promoting Geoethics (IAPG) in Geosciences.”

L 33. “Diverse teams improve the quality of scientific research” – this could be strengthened by a couple of refs supporting the notion.

A/ The following references were added to the revised manuscript: Swartz et al. (2019), Almouzni (2024), Perez-Sepulveda et al. (2025), and Titirici et al. (2025).

L 70. It would be good if you might speculate (or even report literature) o, why “geosciences are among the least diverse fields, in which women and underrepresented groups are exposed to systemic bias that prevents them from equal progress and recognition”. What are the root causes?

A/ The following text was added to the revised manuscript. Lines 74-77: “Although the reasons for the low diversity in geosciences are not fully understood, Marín-Spiotta et al. (2020) listed historical legacies of exclusion, hierarchical power dynamics, culture of impunity, and the discriminatory and unsafe practices during fieldwork as some of the important barriers that goes again diversity.”

L 165. in a language different **than** English → different **from**. Please check for other instances.

A/ This was fixed here and in the entire revised manuscript.

L 299 onwards – next whole paras. It would be good to summarise the salient figures in tables.

All the data from all figures in the paper will be uploaded to a persistent DOI that will be made available in the paper’s “data availability” section. As such summarizing the numbers in tables would be a repeat of the data availability. We prefer to reduce this redundancy by having the same tables in the paper and the data availability statement. In the cloud physics community, having all the figures presented in tables is a necessity to publish and this is always done through a persistent DOI upon acceptance of the publication.

L 325. “Although there is a consensus (67% of total 326 respondents) that collaborating with colleagues and institutions from developing countries (no latitudinal definition was given for this question) is key for the development of the cloud physics field, a large fraction of the survey participants (60%) are not planning to do so.”

Avoid ‘there us/are’ structures. The above may be reworded something like:

We find a consensus (67% of total 326 respondents) that collaborating with colleagues and institutions from developing countries (no latitudinal definition was given for this question) is key for the development of the cloud physics field. However, a large fraction of the survey participants (60%) are not planning to avail themselves of the opportunity.

A/ Thank you for the suggestion. The sentence was modified accordingly in Lines 503-506.

Throughout your article, keep your sentences short: if necessary, split a long one into two and use a connector. The easier you make it for readers to understand, the more your work will be appreciated and cited. In addition to the Copernicus guides, look at this <https://oceansclimate.wixsite.com/oceansclimate/writing-guide>.

A/ Thank you for the suggestion. We have revised the manuscript to avoid long sentences.

L. 350. led by women are better balanced – balanced in what?

A/ To clarify this, sentence was modified as follows. Lines 250-252: “The metadata analysis clearly shows that the papers led by women are better balanced (i.e., the number of women co-authorship and the number of total coauthors is closer), yet still not balanced in the gender of co-authors”

L. 354. **Always** a , before but.

A/ This was fixed here and in the entire revised manuscript.

L. 355. it implies that one route to increase women co-authorships is to have women authors in the lead. Yes, but **how???** How do you put women as lead author? Who does it? Based on what criteria? Etc?

A/ Although we do not have a clear answer for this, we think that increasing the number of women senior scientists and Professor positions could be a way to have more women as lead authors. The following text was added to the revised manuscript. Lines 257-258:

“Nevertheless, it implies that one route to increase women co-authorships is to have women authors in the lead. This could happen if more women get senior scientists and professorship positions.”

L 382. “to note that the lower rates of English proficiency and the apparent “narrow impact” of some tropical studies (based on e.g., desk rejections by Editors), could pressure authors to publish some of their studies in local (mostly non-indexed) journals with low visibility” – It would be important to clarify language skill and marginalisation. How do the two intersect? What is the relationship? The literature in applied linguistics and sociolinguistics will help here.

A/ Thank you for the suggestion. The following text was added to the revised manuscript.

Lines 304-306: “Also, the lack of English skills leads to linguistic discrimination, exacerbating the marginalization of non-native English authors, as demonstrated in different sociolinguistic studies (e.g., Dovchin, 2025; Brown and Laihonon, 2025).”

L 447. “It was found that” → We found

A/ Modified

L 451. “Again, it is obvious that...” Why is it ‘obvious’?

A/ The sentence was modified as follows. Line 556: “As shown in Fig. 1, women scientists started appearing in the community only after 1985”

L 484. “The absence of leadership and engagement from tropical academics” – do you have some refs to support this?

A/ The sentence was based on the metadata analysis results. To clarify this, the sentences was modified as follows. Lines 607-609: “The absence of leadership and engagement from academics working in a tropical country, as shown in Fig. 1, is an illustration of the cloud physics community’s geographically biased worldwide collaboration network.”

L 566. “collabo- ration”. Check for other instances?

A/ This was fixed and the references were double-checked.

L 628. K r e i d e n w e i s , S . M . , P e t t e r s , M ☐ extra spaces

A/ This was fixed.

Decision. With all the revisions that you are being asked to do, I will ask you to submit a new version of your MS and for reviewers to take another look. Please make two copies, as follows:

- **File 1:** Track all the changes in Word. Use line numbering. Make a pdf. This is to comply with the publisher requirement to submit revisions showing changes.
- **File 2:** Make a clean copy of your text – ie, accept all your revisions? Make sure that the clean copy is what you want. Make a pdf. This is so that reviewers can see the revised version more easily.
- **Combined file:** Combine both above files into one pdf, in the above order. You can use ilovepdf.com. At the very top, include a note saying something like *this file contains two versions of the revised text: The first with revision mark ups; the second clean, starting on page nn**. Then submit to the GC website. * Use the pdf page number (somewhere around the centre page of the pdf file), not the text page number.

I will then ask the reviewers to take a relatively quick look to see if they see (I hope mostly minor) things that they think should be revised further. If any questions, email me.

Some fairly recent publications on gender issues and (geo)science.

A/ Thank you for these suggestions. We have included 22 new recent studies in the revised manuscript.

Adams, R., & editor, R. A. E. (2026, May 8). Sexual harassment more than twice as prevalent at England's top universities, analysis finds. *The Guardian*. <https://www.theguardian.com/education/2026/may/08/sexual-harassment-statistics-england-top-universities-analysis>

Batty, D., & Davis, N. (2018, July 7). Why science breeds a culture of sexism. *The Guardian*. <https://www.theguardian.com/science/2018/jul/07/why-science-is-breeding-ground-for-sexism>

Belot, H. (2023, April 3). Australia's Antarctic program beset by frequent sexual harassment, second report finds. *The Guardian*. <https://www.theguardian.com/australia-news/2023/apr/03/australias-antarctic-program-beset-by-frequent-sexual-harassment-second-report-finds>

Defranza, D., Mishra, H., & Mishra, A. (2020). *How Language Shapes Prejudice Against Women: An Examination Across 45 World Languages*. <https://doi.org/10.31234/osf.io/mrbcf>

Diazon, S., Duran, A., Mantes, D., & Monteza, A. M. (2025). Unveiling language-based gender discrimination in fostering gender equality through discourse analysis. *European Journal of Literature Language and Linguistics Studies*, 8, 66–86. <https://doi.org/10.46827/ejll.v8i3.571>

Holmes, J. (2014). Language and Gender in the Workplace. In *The Handbook of Language, Gender, and Sexuality* (pp. 431–451). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118584248.ch22>

Hudson, S. T. J., Myer, A., & Berney, E. C. (2024). Stereotyping, prejudice, and discrimination at the intersection of race and gender: An intersectional theory primer. *Social and Personality Psychology Compass*, 18(2), e12939. <https://doi.org/10.1111/spc3.12939>

Menegatti, M., & Rubini, M. (2017). Gender Bias and Sexism in Language. In M. Menegatti & M. Rubini, *Oxford Research Encyclopedia of Communication*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780190228613.013.470>

Remsö, A., Bäck, H., & Renström, E. A. (2025). A Sexist Language? Gender Differences in Attitudes and Use of Gender-Inclusive Pronouns. *Journal of Language and Social Psychology*, 44(5), 682–703. <https://doi.org/10.1177/0261927X251335305>

Ryan, M. K. (2023). Addressing workplace gender inequality: Using the evidence to avoid common pitfalls. *British Journal of Social Psychology*, 62(1), 1–11. <https://doi.org/10.1111/bjso.12606>

- Sharma, M. (2025). Language as a Mirror of Inequality: A Psychological Analysis of Gendered Language and Sociolinguistic Harm in Nepalese Society. *Open Journal of Social Sciences*, 13(09), 330–345. <https://doi.org/10.4236/jss.2025.139020>
- Stroi, O. (2020). Gender-Biased Language of the Workplace. *Discourse*, 5, 120–131. <https://doi.org/10.32603/2412-8562-2019-5-6-120-131>
- UN Women. (2025, April 21). *How gender inequality and climate change are interconnected*. UN Women – Headquarters. <https://www.unwomen.org/en/articles/explainer/how-gender-inequality-and-climate-change-are-interconnected>
- Xiao, H. (n.d.). *Gender in language and gender in the social mind*.