Referee #1: Anonymous

Paper investigates female authorship of peer-reviewed publications in seismology. The topic of the paper is of high significance as authorship of scientific peer-reviewed papers remains an important criterion for assessing researchers' performance and consequent career advancement. Therefore, any biases or underrepresentation of any demographic groups may lead to lower chances of recognition (job opportunities, career progressions, funding, etc.).

I will not comment on the probabilistic approach of determining the gender of authors based on the first name nor on the statistical method. However, the size of the sample used is sound in terms of statistical significance. The overall reasoning and justifications of various decisions is very appropriate.

The results are quite relevant and of interest. The representation by journal is very useful and (potentially) an eye-opener to both female and male authors.

Q 1.1 In the discussion, the authors compared their results to those of the European Commission. They observed a correlation between 24% of authorship by women and 30% women representation in natural sciences. I am not sure it adds much to the discussion as this comparison is very difficult and thus any correlation is oversimplified.

We thank the reviewer for pointing out this sentence. Actually, it was poorly formulated: The observations of Lariviere et al. (2013) and also the "She Figures" report from the European Commission refer to women authorship, not to representation in general. However, going back to the original data, we also noticed that the measures we were comparing were not entirely comparable (we had been comparing overall women's authorship in our data to the average proportion of women authors *per article* in the She Figures report). We now changed this part of the discussion so as to include a slightly more detailed comparison to the values in the She Figures report and a rationale for performing this comparison.

Added text in lines 276 – 283:

"In terms of unique authors, we found that these are women with a probability of 26% (any author position), 29% (first author position), and 22% (last author position). These numbers correlate well with the proportion of unique active women authors of peer-reviewed publications in natural sciences compiled by the She Figures report based on bibliographic data from Scopus (European

Commission, Directorate-General for Research and Innovation, 2021, p. 221). Among these active authors, there are 29% of women in the early to mid-career segment and 23% of women in the established segment. Although our study considers a specific subfield of natural sciences and a different genderizing strategy, the comparison provides an important indicator that our automated procedure of retrieving and genderizing author names is reliable because it produces results close to those expected for a larger, more general dataset."

Q1.2 In the same line, there are attempts to draw correlations between different fields (i.e. life sciences), which may oversimplify the unique dynamics of each scientific field. However, the lack of more data specific to geosciences in general and seismology in particular explains the comparison with other fields – even if I would prefer to see some more cautiousness in the next stage. In fact, perhaps this limitation of the interpretation / comparisons would deserved to be better highlighted.

Thank you for this comment. It is indeed true that the lack of studies that address gender inequality in the field of seismology makes it harder to draw field-specific comparisons. Increase in similar studies would be ideal to rigorously address the reasons for gender disparities. We consider that the academic careers in many, if not all, fields follow similar rules and assessment procedures (number of publications, h-index, networking leading to collaborations and more publications, etc). This allows us to make comparisons to different scientific fields. We would also like to emphasise that comparing our results to general trends in various academic fields is helpful. This is because there potentially exist systemic problems that impede the progression of the academic career of women researchers, and gender inequality studies can help tackle them. We have added the following text in lines 260 - 265 to better explain our motivation to compare our results to those of similar studies in other academic fields:

"Throughout this discussion, we compare our findings to the literature on gender gaps in authorship in STEM and other fields of research. Although direct comparisons to other academic fields could be simplifying, they provide an indication of the general trend of women's under-representation in authorship and its subsequent effects in academic career progression. Moreover, they allow us to illustrate some of the possible consequences of authorship gender gaps that have not been studied in geosciences yet. The comparisons should be read as a motivation to conduct further studies that aim to address and quantify gender inequities in the field of seismology or geophysics, rather than as a quantitative discussion."

Q1.3 Correlations with and extrapolation from EGU data is an important asset of the paper, as it represents an important reliable data specific to geosciences and seismology. However, EGU data is rather complex. There are membership datasets, registration at General Assembly (GA)

datasets, and there have been changes on how data is collected (i.e. gender changed from optional or mandatory field a few times…). The last two years of GA were also severely impacted by COVID-19 restrictions (online vs. onsite vs hybrid) which adds other layers of complexity to its data. In conclusion, I would strongly encourage the authors to pursue comparisons with EGU data but either in a dedicated chapter or even paper.

We thank the reviewer for this comment, which actually reveals some shortcomings that we did not discuss in enough detail. We now separated the description of high-end productivity from the description of comparing EGU demographics with our data. This allows us to highlight the limitations of the comparison more clearly. We think that the comparison is important to show that our approach produces reasonable results when compared to self-declaration. However, ultimately, both datasets measure slightly different things (society membership vs. authorship). Because of that, we have also adapted the first recommendation in the conclusion section, adding:

"As representation in professional societies and conference attendance are not equal to publication productivity, journal publishers should also consider collecting inclusive and self-declared demographic data." (Lines 403 - 305)

Once the EGU dataset has been collected consistently for 5 or 10 years, it will definitely merit a more detailed investigation. It will provide more comprehensive gender diversity data than what is currently available and what we are able to provide, particularly as it is self-declared and more inclusive than automated genderization. The reviewer also alerted us to an inaccuracy here; we are actually considering membership data, which we confused with GA registration data (because attendees often renew their membership during abstract submission or registration). This inaccuracy has now been corrected. The new section is (lines 218 – 234):

"3.6 Towards comparing demographic and bibliographic data

Since 2016, the European Geosciences Union (EGU) systematically collects self-declared demographic information from participants upon membership registration. In 2016–2017, response rates to the question about gender were low, around 50 % for overall seismology section attendees and around 40 % for early career scientists (ECS). The response rates increased in 2018–2019 to around 60 % (ECS: around 50 %) before reaching close to 100 % in 2020.

Between 2018 and 2021, 29–33 % of all seismology section members and 35–38 % of the early career members identified as women. For the years with higher response rates, and for both levels of seniority, EGU seismology members have a larger probability of being female than manuscript

authors in our dataset. The EGU demographics refer to unique members of the seismology community, while our overall probabilities are computed for authorships, i.e. one person may appear repeatedly. The discrepancy in women's participation might indicate a gender gap in authorship. However, several points prevent a direct comparison of both datasets: (i) the results of automatic genderization are less reliable than self-declaration; (ii) while we expect a large overlap of the researchers represented in both datasets, section members may be more frequently working in European countries than article authors; (iii) our dataset can distinguish first authors (who are often, but not necessarily early-career researchers) from other authors, while EGU considers self-declared ECS; these two populations are not directly comparable. Considering these limitations, we find some consistency between our results and the section membership data, namely that ECS members / first authors are more likely to be women than overall members / authors, and that the rate of women membership / authorship up to 2020 is between 20 and 30 % for all members / authors. We cannot conclude that a gender gap in publication productivity exists due to the mentioned caveats, but based on our results such a gap is possible."

Q 1.4 *The paper conclusions are interesting and the conclusions are of interest for all other fields . Especially, the last bullet "Those evaluating research performance should remain aware that there are, as of now, gender gaps in high-productivity, solo, and high-impact authorship in seismology. "point that deserves a large dissemination. The authors could go even further in terms of ambition in the set of recommendations.*

We thank the reviewer for this positive comment. We moved the recommendation further up, so that we can refer to it in the next point that urges readers to find the causes of these gaps. Also, we extended the recommendation to point out more clearly what the consequences could be:

"Those evaluating research performance should remain aware that there are, as of now, gender gaps in high-productivity, solo, and high-impact authorship in seismology. If this is not taken into consideration in funding and hiring decisions, it may contribute to perpetuating the leaky pipeline problem." (Lines 408 – 409).