

We would like to thank the reviewers and the editor for their helpful and constructive comments on the manuscript “The State of Global Catastrophic Risk Research: A Bibliometric Review”.

We found all of the feedback to be useful in improving and sharpening the research. We have now updated the manuscript to address the comments. We believe it has been significantly strengthened and is now suitable for publication. Below we have listed the reviewer and editor comments in black, along with our responses in light green. Text that has been added to the manuscript is in *italics and darker green*.

Reviewer 1

For brevity we only highlight here those remarks of reviewer 1 which call for changes in the paper. Overall, we thank the reviewer for their very positive assessment of our paper.

Are substantial conclusions reached?

Yes, mainly in concluding that GCR can indeed be considered a subfield and in documenting it's different aspects and inter-connections. However, I do think more could be said about ways to improve the output, balance (in gender, regional origin of authors as well as topic spread), and take-up of GCR ideas. Namely, expanding section 7 with more detail about what is currently being lost or not taken full advantage of due to the limitations of the field described. Perhaps also adding some discussion of other sub-fields that were able to grow and become established disciplines would be valuable

To address these suggestions, we have rewritten the sections 7.2, 7.3 and 7.5. The rewritten sections are as follows:

7.2 More research on a wider range of topics

Another point which clearly comes up is the dominance of the AI topic within GCR research. Usually around a quarter of publications each year focus on AI, and essentially all the other clusters also discuss AI at least in part. In contrast, broader perceptions of global risk likely focus much more on climate change and nuclear war. Other parts of the GCR landscape, for example ecosystem collapse, are under-researched in comparison. The focus on AI is likely related to the history of the field: while the origin of GCR thinking can be traced back to concerns around nuclear war (Beard and Bronson, 2023) and AI was only mentioned in passing in the paper that coined the term ER (Bostrom 2002), much of the early work on GCR/ER under that specific name gave major attention to AI risk, and this has intensified over time. While we are not arguing that there should be less research on AI risk (better understanding of GCR in all domains is good), more research into non-AI topics might be fruitful, while also making it easier for the GCR/ER field as a whole to find common ground with other fields. The missing focus on nuclear war in the GCR/ER literature is especially striking, given its very present threat and long history.

To better connect research clusters, regular cross-cluster workshops could foster collaboration, particularly between technically-focused clusters (like AI) and policy-oriented ones (like Governance). Shared methodological frameworks could help bridge gaps - for instance, the risk assessment approaches from Climate Change research could inform AI risk analysis. The Global

Resilience cluster already demonstrates successful integration, drawing insights from multiple domains to address food security challenges.

Adjacent fields offer valuable insights for GCR research. Insights from system dynamics and ecology could improve understanding of risk cascades. International relations theories could strengthen analysis of global cooperation challenges. Public health approaches could inform response strategies.

7.3 Connection to other fields and diversity

This connection with other fields might also help with another major problem: the skewed gender balance. This is likely due to the roots of GCR/ER research in fields with highly unequal gender balances, such as some STEM fields and philosophy. This imbalance should be addressed. Beyond the moral and ethical reasons, there are also practical reasons: more diverse perspectives typically allow us to come to more solid conclusions, and a diverse field is generally more inviting for others. Things that could be done here include targeted mentoring programs or connecting GCR/ER research to fields with a better gender balance as well as working towards achieving better balance across the board. Other possible approaches include: funding more regional research hubs, specific funding for researchers outside of the most active GCR research countries and dedicated research projects to explore the lack of diversity in GCR research.

When it comes to possible lessons to take from other fields, there are ample examples. Climate Science demonstrates how a field can build credibility and policy influence through structured assessment processes. The IPCC model shows how regular, comprehensive reviews can establish scientific consensus and inform policy (Beck and Mahony, 2018). For GCR research, similar assessment reports could help synthesize findings across clusters and highlight key areas needing attention.

The field of Disaster Risk Reduction shows how to bridge research and practice through regular conferences that combine researchers and practitioners, standardized frameworks that guide both research and policy, and strong connections to national agencies (Dickinson et al., 2016; Vasileiou et al., 2022). GCR research could adapt these approaches, particularly in connecting theoretical work to practical risk reduction.

7.5 Policy uptake of GCR/ER ideas

What this article clearly shows is that there is now an established field of GCR/ER research and consequently a wide variety of knowledge available. However, many countries do not yet take this research into account for their risk assessments, because GCR/ER prevention seldom rises anywhere near the top of the list of pressing issues. Those at the top often seem more immediately important to everyday life. Also, current risk matrices and visualisation approaches can hide the difference of magnitude between small scale disasters and GCR, and the uncertainties and complexity around GCR make them difficult to deal with for policy makers (Boyd and Wilson, 2023).

Several mechanisms could strengthen the research-policy connection. The concentration of research in the US and UK (see Figure S7) means many countries lack direct access to GCR expertise. The field's heavy focus on artificial intelligence (Figure 2) leaves other risks like ecological collapse underexplored in policy contexts. Additionally, the lack of standardized risk assessment frameworks makes it difficult for policymakers to compare and prioritize different risks. The abstract nature of

many GCRs makes them difficult to communicate compared to immediate threats. Institutional constraints also play a role - most government agencies are structured around specific hazards rather than catastrophic risks as a category. The uneven gender and geographic representation in the field (Figure 5) further limits engagement with diverse policy perspectives.

Boyd and Wilson (2023) specifically argue that it is of utmost importance that we include GCR/ER in national risk assessment. This could be done by extending the time horizons and include a broader engagement of the public and experts, moving away from the opaque conditions under which national risk assessments are often conducted today. Also, global coordination here would be very helpful, as the threats themselves are global as well.

A good starting point for this is the United Nations Common Agenda, as it argues for more global cooperation to tackle global challenges. There has been further movement around in the context of the Summit of the Future in 2024. Here the Pact for the Future was drafted, which explicitly mentions existential risks and argues for the global governance of long-term risks. While this is a good first step, much of it is still in the planning phase and not yet implemented in a meaningful way. Pushing this forward could be an important step for being more prepared for GCR and ER.

Another important step in the right direction was the Global Catastrophic Risk Management Act in the United States. This piece of legislation resulted in RAND creating a Global Catastrophic Risk Assessment with the focus on the United States (Willis et al., 2024). The assessment looked at artificial intelligence; asteroid and comet impacts; sudden and severe changes to Earth's climate; nuclear war; severe pandemics, whether resulting from naturally occurring events or from synthetic biology; and supervolcanoes. It also highlights that it is essential to include such risks in risk assessments. This means this report can be used as an argument for policy makers to introduce a broader risk assessment in their countries. Also, just as the Pact for the Future, the assessment argues that international collaboration and preparation is crucial.

However, there is also research that makes very concrete suggestions on what countries could do today to increase resilience, especially against risks to the food system (Garcia Martinez et al., 2024). For greater resilience, maintaining international food trade and cooperation is paramount, as is developing robust food storage and distribution systems. Critical infrastructure preparation involves creating national food security strategies, establishing task forces to address extreme food system risks, and investing in research on rapid deployment of resilient food technologies. These broad interventions, adaptable to various catastrophe scenarios, could significantly mitigate famine risk while requiring relatively modest investments compared to the potential humanitarian benefits. To overcome these barriers, the field could develop more concrete scenarios and metrics and create regional research hubs to better connect with national policymakers. The successful integration of climate science into policy frameworks provides a model for how GCR research could achieve similar influence.

Another approach is public engagement. GCR can trace its origins back to the concerns around nuclear weapons. Here new scientific insights around fallout and nuclear winter gave the disarmament movement powerful arguments to use in public discourse around these threats. In addition, an important insight from the debates around nuclear winter was that many people did not even know about this threat, but were gravely concerned once they realized the danger. This realization in the public realm then gave the disarmament movement the opportunity to push for

change (Beard and Bronson, 2023). It seems plausible that many people today are similarly not aware of the dangers we face and would push for mitigation if they knew about them.

Is the amount and quality of supplementary material appropriate?

Yes, though I think a bit more text to provide context for each of the figures presented in the SM could be helpful for readers

We have extended the figure descriptions.

Reviewer 2

We thank reviewer 2 for their positive assessment of our paper.

In terms of novelty, this is the first attempt ever at an academic review of the field of GCR and ER, as far as I know. I remember thinking about doing this a couple of years ago and I'm glad someone did it. The conclusions are overall not groundbreaking, but are adequate. It was interesting to see the different relative size of the literature clusters and the implications of it. Other than a few minor comments, I think the paper clearly warrants publication.

Section 7.2 makes an interesting, potentially misleading claim that the current focus on AI is due to the origin of the field. While it is true that the modern ER field as conceptualized by Bostrom did arise in part from AI risk concerns, its history comes from further back, and is more related to nuclear risk. Even in the original "Global Catastrophic Risks" book there was only one chapter focused on AI risk, while there was much more material on natural risks and nuclear risk, and the 2002 paper that coined the term existential risk only has one paragraph on AI risk. I think the focus on AI came later, after studying the topic more, and is very unlikely to change, particularly now in the light of recent developments in the field of AI.

We have rephrased this section to better account for how the focus on AI developed:

Another point which clearly comes up is the dominance of the AI topic within GCR research. Usually around a quarter of publications each year focus on AI, and essentially all the other clusters also discuss AI at least in part. In contrast, broader perceptions of global risk likely focus much more on climate change and nuclear war. Other parts of the GCR landscape, for example ecosystem collapse, are under-researched in comparison. The focus on AI is likely related to the history of the field: while the origin of GCR thinking can be traced back to concerns around nuclear war (Beard and Bronson, 2023) and AI was only mentioned in passing in the paper that coined the term ER (Bostrom 2002), much of the early work on GCR/ER under that specific name gave major attention to AI risk, and this has intensified over time. While we are not arguing that there should be less research on AI risk (better understanding of GCR in all domains is good), more research into non-AI topics might be fruitful, while also making it easier for the GCR/ER field as a whole to find common ground with other fields. The missing focus on nuclear war in the GCR/ER literature is especially striking, given its very present threat and long history.

I think the paper would benefit from providing a more comprehensive history of GCR and ER, which could mention post-World War II movements that catalyzed such concerns: the Atomic Scientists' seminal warnings, Bertrand Russell and the related widespread anti-nuclear demonstrations spearheaded by organizations like the Campaign for Nuclear Disarmament and the Committee of 100; as well as Isaac Asimov's "a choice of catastrophes" which was the first nonfiction work in the topic—even if these people were not using the terms GCR/ER because they hadn't been invented yet. I recommend adding a few sentences or a paragraph on this. As a reference for this they can use this book chapter they already cite: Beard, S. J., & Bronson, R. (2023). 1. A Brief History of Existential Risk and the People Who Worked to Mitigate It. *SJ Beard et al*, 1-26.
<https://doi.org/10.11647/OBP.0336.01>

We have rewritten the second paragraph in the introduction to better account for the history of the field:

A subset of global risk research focuses on possible worst-case outcomes. This is the study of global catastrophic risk (GCR) and existential risk (ER). There are varying definitions of what constitutes a GCR or an ER. Common definitions are that GCR is the risk of the death of a significant fraction of all humans or a significant loss of well-being on a global scale, while ER is the risk of human extinction or catastrophes judged to be of a similar magnitude (for example, a permanent global collapse or a long-lasting global dictatorship). For such concepts to take hold, first the general realization was needed that it was possible for humanity to go extinct (Moynihan, 2019, 2020b). A key part of this intellectual history involved the realization of the inherent danger of the first nuclear weapons, and post World War II movements which catalyzed these concerns: the Atomic Scientists' seminal warnings, Bertrand Russell and the related widespread anti-nuclear demonstrations spearheaded by organizations like the Campaign for Nuclear Disarmament and the Committee of 100; as well as Isaac Asimov's "A choice of catastrophes" (Asimov, 1981). Similarly, the concern around climate change brought the idea of global threats to humanity into many people's minds (Beard and Bronson, 2023). Building on these realizations of global threat GCR and ER became distinct concepts in the 2000s. ER was introduced as a term by Bostrom (2002), while GCR was brought to prominence in an edited volume by Bostrom and Cirkovic (2008). Over the 2000s and 2010s the ideas around this were further refined and led to the founding of several research institutes like the Future of Humanity Institute, the Future of Life Institute, the Global Catastrophic Risk Institute and the Centre for the Study of Existential Risk (Beard and Bronson, 2023). In parallel, the concepts and ideas from this field reached a broader audience via popular books like "Superintelligence" (Bostrom, 2014) and "The Precipice" (Ord, 2020).

I would also have liked to see some indication of what the most influential papers or books in the field are. Would it be Bostrom and Cirkovic's seminal "Global Catastrophic Risks" book? Maybe a popular book like Toby Ord's "The Precipice" or Bostrom's "Superintelligence"? Carl Sagan's warnings about nuclear winter that reached the media significantly and could be said to still echo to this day? Or something else?. I find the question of how influential the writings are to be more relevant than who is most prolific, which was studied in the analysis (Fig. 4).

To address this, we added a paragraph to section 3.3

While the authors in Figure 4 are the most prolific ones, this does not necessarily mean that they are the most impactful ones in the field. However, impact is much more difficult to assess, as there is no

unified measure of impact. Still, we can interpolate this by looking at what comes up using different measures. Possible metrics here are citation counts or the OpenAlex relevance score. For these two metrics the most prominent works are the paper which introduced the term ER (Bostrom, 2002), the book which introduced the term GCR (Bostrom and Cirkovic 2008) and the book “The Precipice” (Ord, 2020), which tries to evaluate which hazards entail the highest risk for humanity. The importance of these works is also reflected by the fact that they are often featured prominently in discussions around the history of the field (Beard and Torres, 2020) or summaries of GCR/ER relevant literature (Shackelford et al., 2019). Given the recognition of their works, this would make Nick Bostrom, Toby Ord and Milan Cirkovic the most impactful GCR/ER researchers to date. In addition to the works mentioned here, Nick Bostrom also contributed to the general recognition of existential risk by publishing the book “Superintelligence (Bostrom, 2014) about the dangers of AI, which is more highly cited than all the other examples here.

I would have been interested to see a broader treatment of the policy recommendations the field has generated and its limited policy uptake, beyond the brief discussion of section 7.5. I think doing an analysis of how much GCR documents have been cited in policy would be too much and reach beyond the scope, but it could benefit from citing some relevant examples such as: the “Global Catastrophic Risk Management Act” law of the United States, the UN discussions around existential risk in “Our Common Agenda” and “What kind of institution is needed for existential security”.

The following reference has more discussion on the potential reasons why the policy uptake of GCR policy has been low, and proposes an intervention to address this: Boyd, M., & Wilson, N. (2023). Assumptions, uncertainty, and catastrophic/existential risk: National risk assessments need improved methods and stakeholder engagement. *Risk analysis*, 43(12), 2486-2502.
<https://doi.org/10.1111/risa.14123>

You can find many examples of policy proposals to address GCR and their academic sources here:
<https://airtable.com/appik0mTEiTni2SgE/shr8XPqTaoiyAfEcL/tblqS8BUzkkpK5Err>

This review paper has some discussion on efforts to implement GCR resilience work into policy, particularly against nuclear war: García Martínez, J. B., Behr, J., Pearce, J., & Denkenberger, D. (2024). Resilient foods for preventing global famine: a review of food supply interventions for global catastrophic food shocks including nuclear winter and infrastructure collapse. *Critical Reviews in Food Science and Nutrition*, 1-27. <https://doi.org/10.1080/10408398.2024.2431207>

We have significantly extended the discussion around policy with the suggestions:

What this article clearly shows is that there is now an established field of GCR/ER research and consequently a wide variety of knowledge available. However, many countries do not yet take this research into account for their risk assessments, because GCR/ER prevention seldom rises anywhere near the top of the list of pressing issues. Those at the top often seem more immediately important to everyday life. Also, current risk matrices and visualisation approaches can hide the difference of magnitude between small scale disasters and GCR, and the uncertainties and complexity around GCR make them difficult to deal with for policy makers (Boyd and Wilson, 2023).

Several mechanisms could strengthen the research-policy connection. The concentration of research in the US and UK (see Figure S7) means many countries lack direct access to GCR expertise. The

field's heavy focus on artificial intelligence (Figure 2) leaves other risks like ecological collapse underexplored in policy contexts. Additionally, the lack of standardized risk assessment frameworks makes it difficult for policymakers to compare and prioritize different risks. The abstract nature of many GCRs makes them difficult to communicate compared to immediate threats. Institutional constraints also play a role - most government agencies are structured around specific hazards rather than catastrophic risks as a category. The uneven gender and geographic representation in the field (Figure 5) further limits engagement with diverse policy perspectives.

Boyd and Wilson (2023) specifically argue that it is of utmost importance that we include GCR/ER in national risk assessment. This could be done by extending the time horizons and include a broader engagement of the public and experts, moving away from the opaque conditions under which national risk assessments are often conducted today. Also, global coordination here would be very helpful, as the threats themselves are global as well.

A good starting point for this is the United Nations Common Agenda, as it argues for more global cooperation to tackle global challenges. There has been further movement around in the context of the Summit of the Future in 2024. Here the Pact for the Future was drafted, which explicitly mentions existential risks and argues for the global governance of long-term risks. While this is a good first step, much of it is still in the planning phase and not yet implemented in a meaningful way. Pushing this forward could be an important step for being more prepared for GCR and ER.

Another important step in the right direction was the Global Catastrophic Risk Management Act in the United States. This piece of legislation resulted in RAND creating a Global Catastrophic Risk Assessment with the focus on the United States (Willis et al., 2024). The assessment looked at artificial intelligence; asteroid and comet impacts; sudden and severe changes to Earth's climate; nuclear war; severe pandemics, whether resulting from naturally occurring events or from synthetic biology; and supervolcanoes. It also highlights that it is essential to include such risks in risk assessments. This means this report can be used as an argument for policy makers to introduce a broader risk assessment in their countries. Also, just as the Pact for the Future, the assessment argues that international collaboration and preparation is crucial.

However, there is also research that makes very concrete suggestions on what countries could do today to increase resilience, especially against risks to the food system (Garcia Martinez et al., 2024). For greater resilience, maintaining international food trade and cooperation is paramount, as is developing robust food storage and distribution systems. Critical infrastructure preparation involves creating national food security strategies, establishing task forces to address extreme food system risks, and investing in research on rapid deployment of resilient food technologies. These broad interventions, adaptable to various catastrophe scenarios, could significantly mitigate famine risk while requiring relatively modest investments compared to the potential humanitarian benefits.

To overcome these barriers, the field could develop more concrete scenarios and metrics and create regional research hubs to better connect with national policymakers. The successful integration of climate science into policy frameworks provides a model for how GCR research could achieve similar influence.

Another approach is public engagement. GCR can trace its origins back to the concerns around nuclear weapons. Here new scientific insights around fallout and nuclear winter gave the disarmament movement powerful arguments to use in public discourse around these threats. In addition, an important insight from the debates around nuclear winter was that many people did not

even know about this threat, but were gravely concerned once they realized the danger. This realization in the public realm then gave the disarmament movement the opportunity to push for change (Beard and Bronson, 2023). It seems plausible that many people today are similarly not aware of the dangers we face and would push for mitigation if they knew about them.

Other than that, I think the scientific methods and assumptions valid and clearly outlined, the results are sufficient to support the interpretations and conclusions, the authors give proper credit to related work and clearly indicate their own contribution, the title clearly reflects the contents of the paper, the abstract provides a concise and complete summary, the overall presentation well-structured and clear, and the number and quality of references are appropriate.