

Response to Review Comments

Dear Editor and Referees,

Thank you for offering us an opportunity to improve the quality of our submitted manuscript (**egusphere-2025-2706**). We appreciate very much the referees' constructive and insightful comments, which are helpful for our current submission and future research. In response to the review comments and editorial decision of **major revision**, we have undertaken a comprehensive overhaul of the manuscript.

Key revisions include:

- 1) Restructuring the manuscript from introduction to conclusions, including reorganizing the methodology, splitting the previous content into new *Section 3 Results* and *Section 4 Remote sensing data integration for ARI hotspots*, and newly adding *Section 5 Discussion*.
- 2) Rewriting large parts of the *Section 1 Introduction* to better clarify motivation, research gap, and contribution.
- 3) Improving ARI methodology descriptions in *Sections 2.1-2.4* and adding an analytical workflow diagram in *Section 2.5 Analytical workflow*.
- 4) Strengthening the results through robustness and sensitivity analyses (*Section 3.3*) and enhanced global-to-local interpretation of ARI outcomes (*Section 3.2*).
- 5) Adding an entirely new discussion section responding to referees' scientific concerns and outlining future directions.
- 6) Refining conclusions and writing throughout, improving clarity and coherence.

All modifications are marked using track changes, and related revisions are explicitly referenced throughout this response letter for ease of review. Revised manuscript structure is provided below to assist referees in locating major changes:

1 Introduction

1.1 Background

1.2 Motivation and Objectives

2 Data and Methods

2.1 Disaster event dataset and filtering strategy

2.2 Spatial standardization

2.3 Risk indices

2.3.1 World Risk Index (reference index)

2.3.2 Accumulated Risk Index (proposed index)

2.4 Spatial aggregation for ARI computation

2.5 Analytical workflow

3 Results

3.1 Spatiotemporal characteristics of recent disaster events

3.2 Global ARI distribution and hotspots

3.3 Validation and robustness checks

3.3.1 Sensitivity to the fatality threshold

3.3.2 Sensitivity to grid resolution

4 Remote sensing data integration for ARI hotspots

4.1 Data sources and spatial coverage

4.2 Supplementing data gaps in high-risk grids

In the following, we provide a point-by-point response to all review comments. Referee comments are reproduced in blue text, followed by our responses in black text, and corresponding manuscript changes are highlighted in red text with location references. We hope that the revisions meet the expectations of the editor and referees, and we are grateful for your time and consideration.

➤ Review Comments of Referee #1:

The authors are to be commended for the substantial effort to undertake their study, using an events database from 2013 to 2023. However, this is a narrow time window for a proposed global multi-disaster risk assessment index using 5°×5° latitude-longitude grids.

The authors state that this narrow time window is for preliminary research. Using a threshold of 50 fatalities (Total Deaths), their study identifies 344 major disaster events for detailed statistical analysis. However, death count is an unreliable risk measure, since it fluctuates according to a number of factors such as time of day of event occurrence, and the effectiveness of hazard warnings.

For more meaningful results, the authors should consider a much longer time window, with a lower threshold for fatalities.

We fully acknowledge the referee's concerns regarding the relatively narrow time window and the reliance on a ≥ 50 fatalities threshold, both of which may limit the robustness and representativeness expected of a comprehensive global multi-disaster risk assessment index. We sincerely appreciate this fundamental critique. Importantly, we agree that, from a strict disaster risk theory perspective, longer temporal coverage and more inclusive (multi-dimensional) severity indicators would generally be preferable when constructing a global risk index in the classical sense.

In the revised manuscript, we therefore clarify the scope and methodological intent more explicitly: the ARI is positioned as an exploratory, event-driven indicator designed to capture recent accumulated high-impact disaster activity in combination with country-level risk context, rather than as a definitive replacement for long-term structural risk indices. We also strengthened the manuscript by clarifying limitations, adopting more cautious interpretations, and adding forward-looking directions (temporal weighting, longer windows, and alternative severity indicators). Our detailed responses are provided below. For clarity, we respond from two perspectives: methodological intent and practical data considerations.

1 Response to the issue of a narrow time window

We sincerely thank the referee for raising important concerns regarding the relatively narrow time window adopted in our analysis. We fully understand and respect the referee's view that a limited temporal scope may constrain the representativeness and generalizability of results, particularly for low-frequency extreme events. This concern is valid, and it is one of the reasons we now state more explicitly that the present work is preliminary and exploratory, aiming to establish a transparent framework for near-term accumulated impacts and to provide a basis for future long-window extensions.

First, we recognize the well-established principle in disaster risk research that the frequency of natural hazards tends to decrease as intensity increases, which motivates many studies to adopt longer temporal spans to capture rare extremes. At the same time, the selection of a time window must strike a balance between data reliability/comparability and contemporary relevance, especially for global-scale, grid-based spatial analysis. The core objective of our study is to develop a framework for assessing the recent accumulative impacts of multiple natural hazards across global regions and to provide a data

foundation for future long-term analyses. Our decision to focus on the 2013-2023 period was based on the following considerations:

1) Data consistency and completeness:

Earlier disaster records often suffer from missing entries, inconsistent reporting standards, and definitional ambiguities, with these issues being particularly prevalent in developing countries, which are a primary focus of this study. As explicitly stated on the EM-DAT website, “Pre-2000 data is particularly subject to reporting biases” (<https://www.emdat.be>). Since the early 21st century, EM-DAT has made significant improvements in data standardization, event classification consistency, and geolocation practices (Delforge et al., 2025). Consequently, the 2013-2023 dataset offers higher comparability and reliability for consistent event screening and grid-based spatial allocation. We now make this rationale explicit in *Section 2.1 Disaster event dataset and filtering strategy*:

“In addition, earlier EM-DAT records (particularly pre-2000) are known to suffer from reporting inconsistencies and missing entries, especially in developing countries. Therefore, restricting the temporal scope to the past decade ensures higher data completeness and comparability, consistent with the objective of capturing recent accumulative disaster impacts.”

2) Disaster pattern shifts under climate change and contemporary relevance:

As mentioned in *Section 1.1 Background*, recent years have witnessed noticeable changes in disaster frequency, intensity, and spatial distribution, influenced by climate variability and long-term warming signals. A growing body of literature documents upward trends in the frequency and severity of hazards in recent decades (e.g., Hussain et al., 2023; Wen et al., 2023). We observed that many mainstream global risk indices are not explicitly designed to highlight near-term clustering and its accumulated stress on coping capacity. Using a much longer time window may smooth or dilute signals of recent clustering that are especially relevant for near-term planning. For this reason, we argue that a shorter but higher-quality and more contemporarily relevant time window is suitable for the specific purpose of ARI as a near-term accumulation indicator.

3) Complementarity with existing long-term indices:

Our proposed ARI is not intended to replace existing tools such as the WRI, but rather to provide a complementary perspective that emphasizes the accumulative effects of recent major disaster events, an aspect that may be less visible in long-term, country-level indices. Methodologically, ARI focuses on short-term concentration and accumulation of impacts, which differs from indices like WRI that emphasize long-term exposure and structural vulnerability. We also note that repeated events can generate compound disruptions when new shocks occur before recovery is complete. The World Risk Report has emphasized that vulnerability and coping capacity can be eroded through repeated stressors, not only through singular catastrophic events (Frege et al., 2023). Hence, identifying high-impact accumulation zones within a recent window can be valuable for preparedness and humanitarian prioritization, while remaining complementary to long-term indices. This positioning is now stated explicitly in multiple places in the revised manuscript:

Section 1.2 Motivation and Objectives:

“Despite the substantial progress represented by existing risk indices, two key limitations hinder the identification of recent and spatially concentrated disaster impacts. First, long-term averages of hazard occurrence and socio-economic indicators may obscure the accumulative impacts of recurrent disaster events observed in the most recent decade. However, hazard clusters within short periods can cause compounding disruptions that overwhelm local coping capacities, especially in developing countries (Ridder et al., 2020; Rentschler et al., 2022; Frege et al., 2023). Capturing these near-term accumulative effects therefore requires an index that explicitly emphasizes recent, high-impact disaster events rather than multi-decade averages.”

“The proposed ARI aims to provide a complementary perspective that links recent disaster dynamics with the underlying socio-economic environment.”

“For researchers, ARI offers a complementary lens to existing long-term indices, enabling comparative analyses of

short-term disaster accumulation and supporting studies on emerging hazard patterns under climate change.”

Section 2.3.2 Accumulated Risk Index (proposed index):

“In this study, WRI is not recomputed or modified. Instead, it is adopted as a country-level reference index and contextual weighting factor that encapsulates exposure and vulnerability conditions. The proposed ARI builds on WRI by combining it with the recent occurrence of major disaster events at the grid scale, using WRI as a scaling factor rather than as a quantity to be spatially aggregated on its own.”

“Building on the reference index introduced above, ARI is formulated as a grid-based measure that integrates recent major disaster events with country-level contextual information.”

Section 5 Discussion:

“ARI highlights areas where high-fatality disaster events have clustered within a relatively short period and where underlying country-level exposure and vulnerability conditions amplify their impacts. The index does not aim to replicate or replace long-term structural risk metrics such as WRI, but to provide a complementary perspective focused on recent accumulative shocks. Compared with a grid-based equivalent of the WRI, ARI yields sharper contrasts between hotspots and low-activity areas, indicating that near-term accumulative effects can significantly modify the perceived risk pattern at sub-national and transboundary scales. The results therefore underscore the value of combining event-driven metrics with background socio-economic conditions for identifying areas where resilience may have deteriorated more rapidly than long-term indices would suggest.”

Section 6 Conclusions:

“Compared with the long-term WRI, ARI provides sharper spatial contrasts and highlights sub-national hotspots where repeated shocks may have accelerated resilience loss. While not a replacement for structural risk indices, ARI offers a complementary perspective focused on near-term accumulative impacts.”

That said, we agree with the referee that extending the temporal scope can improve statistical representativeness for rare extremes and enhance alignment with “classical” global risk index expectations. At present, we have not yet arrived at a perfect balance between capturing “recent accumulative effects” and “long-term average risk”. Therefore, rather than claiming that a short window is universally preferable, we now clearly acknowledge the trade-off and treat the present time window as a pragmatic starting point.

To address this tension, we plan to explore the integration of temporal weighting coefficients in future research, which would allow us to differentiate between historical hazard frequency and recent disaster impact intensity, thereby achieving a more nuanced temporal representation of risk. Furthermore, we also intend to extend the temporal coverage to include disaster events from 2000 to 2023 in future studies. We will conduct sensitivity analyses using different time windows (e.g., 2000-2023 vs. 2013-2023) to assess the spatial stability and variability of the ARI across scenarios. These efforts will help enhance the adaptability, transparency, and scientific rigor of the proposed framework.

2 Response to the issue of a ≥ 50 fatalities threshold and the use of death count

We sincerely thank the referee for raising this fundamental issue. The selection of representative and reliable impact indicators remains a key challenge in the field of disaster risk assessment (Yarveysi et al., 2023; Marin et al., 2021). We fully agree that death counts are not a complete or ideal measure of disaster “risk” and that fatalities can be strongly influenced by contextual factors (e.g., warning effectiveness, time of occurrence, infrastructure quality, and response capacity). For this reason, in the revised manuscript we avoid treating fatalities as a comprehensive risk metric; instead, we describe “Total Deaths” as a pragmatic proxy for severe societal impact under current global data constraints, and we position ARI as an event-driven indicator of recent accumulated high-impact shocks, rather than a full multi-dimensional risk model.

In selecting an operational criterion for identifying globally comparable “major” events, we reviewed the full suite of

EM-DAT impact indicators. EM-DAT records 11 impact indicators, broadly grouped into:

- 1) Population-based indicators: Total Deaths, No. Injured, No. Affected, No. Homeless, Total Affected;
- 2) Economic-based indicators: Reconstruction Costs, Reconstruction Costs (Adjusted), Insured Damage, Insured Damage (Adjusted), Total Damage, Total Damage (Adjusted).

However, these indicators are not uniformly complete. In particular, economic indicators have substantial missingness and heterogeneity, especially in less developed countries, which are a core focus of our study. Using such indicators as the primary filtering criterion risks introducing selection bias and undermining comparability (Peduzzi et al., 2009). Similarly, population-based indicators such as “No. Affected” and “No. Homeless” can involve varying definitions and estimation practices across countries, leading to high uncertainty (Newman and Noy, 2023; Peduzzi et al., 2009). By comparison, while imperfect, “Total Deaths” remains one of the most consistently reported indicators across contexts and thus offers relatively higher cross-country comparability at the global scale.

To further improve robustness and reduce noise from small events subject to under-reporting, we adopted a stricter selection threshold (≥ 50 fatalities), above EM-DAT’s minimum entry criterion (e.g., 10 deaths or 100 affected). We emphasize that this threshold does not imply that smaller events are unimportant; rather, it is a pragmatic screening choice to focus on high-impact events that are more consistently recorded and more likely to reflect substantial societal disruption. This rationale is now clarified in *Section 2.1 Disaster event dataset and filtering strategy* and explicitly acknowledged as a limitation in *Section 5 Discussion*:

“Using a threshold of 50 fatalities (Total Deaths), this study identifies 344 major disaster events for detailed statistical analysis. The selection of “Total Deaths” as the primary criterion reflects its superior reliability and objectivity compared to other metrics. Population-based indicators in EM-DAT, such as “No. Affected” and “No. Homeless”, are less reliable due to varying definitions and estimation methods, which often depend heavily on other metrics or overall event assessments (UNDP, 2004; Newman and Noy, 2023; Peduzzi et al., 2009). Similarly, economic indicators like “Total Damage” introduce significant statistical bias due to varying socio-economic capacities across countries with different development levels (UNDP, 2004; Peduzzi et al., 2009). The ≥ 50 fatalities threshold further excludes small-scale events that may be more prone to under-reporting, thereby improving representativeness for global-scale analysis.”

“Nevertheless, several methodological limitations should be acknowledged: 1) Event filtering: Using a ≥ 50 fatality threshold excludes smaller but potentially accumulative events to reduce under-reporting bias and improve international comparability.”

We also recognize that using a ≥ 50 fatalities threshold may reduce sensitivity to moderate but recurrent events in some regions. For this reason, we have added a sensitivity analysis in *Section 3.3.1 Sensitivity to the fatality threshold* to evaluate how ARI patterns change when alternative thresholds are used (40, 60, and 70 fatalities). The results show that although absolute counts and ARI magnitudes change, the primary hotspot configurations remain broadly consistent, supporting the robustness of the spatial conclusions within the intended scope of ARI.

As also stated in the manuscript, the current work is intended as a preliminary study that draws scholarly and policymaker attention to the layered risks faced by developing countries, particularly under conditions of frequent, overlapping natural hazards. We hope this framework can support more equitable and needs-based global humanitarian planning. Looking ahead, we plan to refine the ARI methodology by introducing composite severity indicators to better reflect the multifaceted impacts of disaster events. We also intend to explore distribution-based weighting schemes to dynamically adjust the contribution of different events according to severity, thereby enhancing the model’s precision and flexibility.

Finally, we agree with the referee that more meaningful results could be achieved by incorporating longer time windows and alternative severity measures. Accordingly, in the revised version of the manuscript, we have added an explicit

forward-looking statement at the end of *Section 5 Discussion* to attempt to outline directions for future work:

“Future work could extend ARI by testing composite severity indicators when data quality allows, exploring alternative risk indices alongside WRI, and applying temporal weighting schemes to relate recent accumulative patterns to longer-term trends. Improved geocoding and the use of hazard footprints derived from Earth observation would support refined spatial units in selected regions. Closer coupling between ARI outputs and observed damage or recovery trajectories would also help to evaluate and further develop accumulated risk indicators under changing climate and socio-economic conditions.”

This concludes our response to the referee’s comments. We sincerely appreciate these valuable suggestions. They helped us to sharpen the scope statement, strengthen methodological transparency, and adopt a more cautious and theory-consistent positioning of ARI. We hope the revisions and clarifications address the referee’s concerns and clarify the balance we sought between methodological intent (near-term accumulation focus) and practical constraints of global multi-source data. We remain committed to expanding ARI toward broader temporal and indicator coverage in future work.

➤ **Review Comments of Referee #2:**

Major comments:

1. Clarity of objectives: the manuscript does not sufficiently articulate the purpose of Accumulated Risk Index (ARI). The authors should clearly state what decisions or actions this index is intended to inform. How can policymakers, practitioners, or researchers use ARI in practice? Without a clear statement of utility, the value of the index remains uncertain.

We sincerely thank the referee for this important observation. We agree that the original version of the manuscript did not sufficiently articulate the intended purpose and practical relevance of the ARI. Our intention in developing ARI is not only to provide a statistical tool, but more importantly, to highlight areas where repeated and accumulative disaster impacts create disproportionately high risks that may not be fully captured by traditional long-term risk assessment indices.

In addition, the creation of ARI is intended to complement, rather than replace, the existing global disaster risk assessment framework. As outlined in *Section 1.1 Background*, while representative global indices emphasize different aspects, such as data timeliness, assessment units at the national level, the range of disaster types considered, or the choice of assessment indicators, there remains a clear gap. To our knowledge, no comprehensive global index has yet focused on multiple types of disasters occurring within the past decade while adopting a latitude-longitude grid, rather than national borders, as the minimum unit of assessment. Filling this gap in the existing risk assessment landscape constitutes an important motivation of our study.

Specifically, ARI is designed to support three levels of decision-making:

- 1) For policymakers: ARI can help identify high-risk areas where repeated disaster shocks have eroded resilience, thereby prioritizing these areas for disaster preparedness planning, humanitarian aid allocation, and capacity-building programs.
- 2) For practitioners and humanitarian organizations: ARI provides a spatially explicit signal of recent disaster accumulation, which can guide on-the-ground resource deployment, early recovery planning, and coordination of international assistance.
- 3) For researchers: ARI offers a complementary perspective to existing indices such as the WRI, by focusing on the short-term accumulative effects of disasters. This makes it useful for comparative studies and for evaluating the evolving impacts of climate change on disaster risk patterns.

To address this comment, substantial revisions have been made rather than adding only explanatory sentences. Most importantly, we have reorganized the original introduction into two subsections, *Section 1.1 Background* and the newly developed *Section 1.2 Motivation and Objectives*, where *Section 1.2* is specifically written to clarify the purpose of ARI and its intended users. This structural adjustment directly responds to the referee’s request for clearer articulation of objectives.

In *Section 1.2 Motivation and Objectives*, we now clearly state that:

“Despite the substantial progress represented by existing risk indices, two key limitations hinder the identification of recent and spatially concentrated disaster impacts. First, long-term averages of hazard occurrence and socio-economic indicators may obscure the accumulative impacts of recurrent disaster events observed in the most recent decade. However, hazard clusters within short periods can cause compounding disruptions that overwhelm local coping capacities, especially in developing countries (Ridder et al., 2020; Rentschler et al., 2022; Frege et al., 2023). Capturing these near-term accumulative effects therefore requires an index that explicitly emphasizes recent, high-impact disaster events rather than multi-decade averages.

Second, the prevalent use of country-level assessment units restricts the ability to identify sub-national or transboundary hazard clusters. Natural hazards follow geophysical processes rather than political borders, and their spatial footprint often spans multiple administrative units. These concerns were also highlighted in previous research (Paprotny et al., 2018; Mucke et al., 2011) and reinforce the need for methodological innovation.

Motivated by these gaps, this study introduces a global multi-disaster risk assessment index that focuses specifically on the accumulative effects of major disaster events from 2013 to 2023. The index uses $5^{\circ} \times 5^{\circ}$ latitude-longitude grids as its basic assessment unit and incorporates WRI as a contextual weight coefficient. The proposed ARI aims to provide a complementary perspective that links recent disaster dynamics with the underlying socio-economic environment.

The intended purpose of ARI is threefold: 1) For policymakers, ARI helps identify areas where repeated disaster shocks have eroded resilience and thus require priority in preparedness planning and capacity-building initiatives. 2) For practitioners and humanitarian organizations, ARI provides a spatially explicit indicator of recent accumulative disaster activity that can guide early recovery efforts and international assistance coordination. 3) For researchers, ARI offers a complementary lens to existing long-term indices, enabling comparative analyses of short-term disaster accumulation and supporting studies on emerging hazard patterns under climate change.”

Meanwhile, the applied value of ARI and its complementary role now appear explicitly in multiple other places in the revised manuscript to highlight its potential utility for policy and practice, such as:

Section 5 Discussion:

“ARI highlights areas where high-fatality disaster events have clustered within a relatively short period and where underlying country-level exposure and vulnerability conditions amplify their impacts. The index does not aim to replicate or replace long-term structural risk metrics such as WRI, but to provide a complementary perspective focused on recent accumulative shocks. Compared with a grid-based equivalent of the WRI, ARI yields sharper contrasts between hotspots and low-activity areas, indicating that near-term accumulative effects can significantly modify the perceived risk pattern at sub-national and transboundary scales. The results therefore underscore the value of combining event-driven metrics with background socio-economic conditions for identifying areas where resilience may have deteriorated more rapidly than long-term indices would suggest.”

“Despite these caveats, the combined ARI-remote sensing framework offers practical value. For policy and humanitarian actors, ARI can help flag areas where repeated high-impact events may have eroded resilience and where more detailed assessments and targeted support are warranted.”

Section 6 Conclusions:

“Compared with the long-term WRI, ARI provides sharper spatial contrasts and highlights sub-national hotspots where repeated shocks may have accelerated resilience loss. While not a replacement for structural risk indices, ARI offers a complementary perspective focused on near-term accumulative impacts.”

Overall, by restructuring the introduction and expanding *Section 1.2*, the revised manuscript now explicitly articulates the purpose of ARI and its intended users. Practical use cases are further reinforced in the discussion and conclusions, clarifying

how ARI supports policy prioritization, operational planning, and academic research. We believe this addresses the referee's concern regarding clarity of objectives and improves the manuscript's applicability and interpretability for a broad readership.

2. Methodological rigor and transparency:

2.1. The methodology section is fragmented and lacks coherence. It introduces three elements—identification of 344 major events, the ARI itself, and remote sensing data—without explaining how they are connected. A clear workflow is needed.

We sincerely thank the referee for this valuable suggestion. We fully agree that in the original version, the methodology section was presented in a somewhat fragmented way, which may have made it difficult for readers to clearly understand the logical connection among the identification of major disaster events, the construction of the ARI, and the establishment of the remote sensing dataset. We appreciate this comment, which prompted us to substantially strengthen the structure and transparency of the methodology.

To clarify, the methodological logic of this study can be summarized in three sequential steps:

- 1) Identification of major events: Using the EM-DAT database, we extracted disaster records within the study time window. Events were screened by the threshold of ≥ 50 fatalities to ensure reliability and global comparability, resulting in 344 major disaster events that form the empirical foundation of the analysis.
- 2) Construction of the ARI: Based on the identified events, we aggregated disaster occurrences within each $5^\circ \times 5^\circ$ latitude-longitude grid cell by integrating the number of major events with the weighting factor derived from the WRI, which incorporates exposure and vulnerability dimensions at the country level. The result is the ARI, quantifying accumulative disaster risk over the past decade at the geographic grid scale.
- 3) Development of the dataset: Finally, we linked the high-ARI grids to remote sensing data in order to build an open-source dataset focusing on developing countries. This dataset is intended to facilitate disaster damage detection, risk monitoring, and the training and validation of remote sensing models. This step bridges statistical risk assessment with geospatial applications and is intended to support future research on hazard mapping and disaster response.

To address this concern, we conducted substantive revisions rather than minor textual adjustments, focusing on improving methodological coherence and workflow transparency. Specifically, the following major modifications have been implemented in the revised manuscript:

1) Structural reorganization of the methodology section:

The previous manuscript combined event identification, index construction, and dataset development within a single methods narrative, which may have contributed to the sense of fragmentation highlighted by the referee. In the revised version, we have fully reorganized *Section 2 Data and Methods*, clarified thematic boundaries, and relocated result-oriented content to the newly established *Section 3 Results* and *Section 4 Remote sensing data integration for ARI hotspots*, ensuring that methodological descriptions and analytical outputs are now presented sequentially and non-overlapping. This restructuring creates a clearer narrative flow from data → method → results.

2) Addition of a new subsection dedicated to analytical workflow:

To explicitly bridge the three methodological components (major event identification → ARI construction → remote sensing dataset integration), we created a new *Section 2.5 Analytical workflow*. This subsection synthesizes the technical logic of the study, describes methodological dependencies, and connects EM-DAT event screening, WRI-weighted ARI computation, and high-risk grid dataset development into a unified pipeline, directly addressing the referee's request for clarity.

3) Introduction of a workflow diagram for transparent methodological communication:

Following the referee's suggestion that "a clear workflow is needed", we added a new workflow diagram (*Figure 1*) to

visually illustrate the complete methodological process. The figure clearly shows how target events are filtered, how ARI is calculated from WRI-weighted event aggregation, and how high-risk grids are subsequently linked to remote sensing imagery to form the 3H Dataset. This visual representation substantially improves reader understanding and enables rapid verification of methodological logic.

Together, these revisions significantly enhance the transparency, readability, and continuity of the methodology section. We believe the updated structure, newly added workflow subsection, and visual diagram directly and effectively resolve the referee’s concern regarding fragmentation and methodological coherence.

Revised content now appears as follows:

“To improve methodological coherence and clarify how the different components of this study are integrated, we provide a consolidated workflow summarizing the complete process from data acquisition to index construction and dataset development (Figure 1). The workflow consists of three sequential stages: 1) Identification of major disaster events: Disaster records for 2013-2023 are extracted from EM-DAT and filtered using the ≥ 50 fatalities threshold (Section 2.1). Each selected event is assigned a standardized representative coordinate (Section 2.2), forming the empirical basis for subsequent spatial analysis. 2) Construction of the ARI: For each $5^\circ \times 5^\circ$ grid cell, major disaster events are aggregated and weighted using country-level WRI values to compute ARI (Section 2.3). The choice of grid size and related sensitivity considerations are discussed in Section 2.4 and evaluated in Section 3.3. 3) Development of the remote sensing dataset: Grids with high ARI values, interpreted as hotspots of accumulated risk, are linked to sub-meter satellite imagery. Existing open-source data are compiled and supplemented where necessary to form the proposed dataset (Section 4).

This structured workflow provides a unified view of how statistical disaster records, global risk indices, and remote sensing observations jointly contribute to the assessment of accumulated multi-hazard risks and the development of a complementary geospatial dataset.”

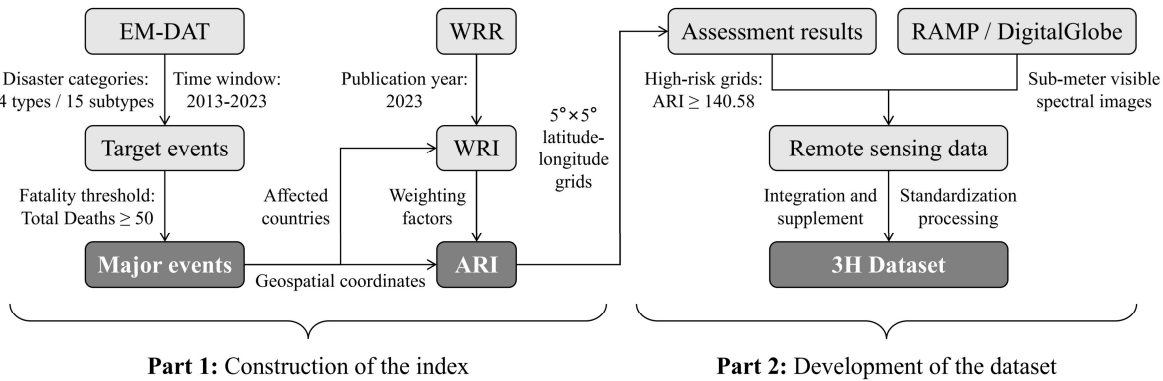


Figure 1: Methodology workflow diagram.

2.2. The threshold of 50 fatalities to identify “major events” appears arbitrary. A sensitivity analysis should be conducted to explore how results change when this threshold is varied.

We sincerely thank the referee for this valuable comment regarding the choice of 50 fatalities as the threshold for defining “major disaster events”. We fully acknowledge that the threshold may appear arbitrary without further justification, and we agree that this decision requires careful clarification.

Our rationale for selecting this threshold is threefold:

1) As discussed in our manuscript, records of disaster impacts in global databases such as EM-DAT often suffer from incompleteness and reporting inconsistencies, especially in developing countries. Among the various impact indicators, “Total Deaths” is relatively more reliable and consistently recorded across regions. By adopting a stricter threshold of 50 fatalities, higher than EM-DAT’s minimum entry criterion of 10 fatalities, we sought to minimize analytical noise from small-scale or less systematically reported events. This also ensures a focus on high-impact disaster events with broad

societal consequences to improve the robustness and representativeness of the assessment results.

2) Our research objective is to emphasize the accumulative effects of severe disasters over the last decade. A higher threshold helps capture those events most likely to cause long-term disruptions and repeated social vulnerability, aligning with the conceptual foundation of the ARI.

3) Our decision to set this threshold was also informed by the need to ensure compatibility with remote sensing applications. The 3H Dataset introduced in our study is constructed based on ARI-identified high-risk grids. In these zones, disasters resulting in large fatalities often correlate strongly with observable physical destruction, such as collapsed buildings, facilitating effective integration with satellite imagery for future disaster detection and classification models.

That said, we recognize the referee's concern about the potential arbitrariness of this choice. This reasoning is now clarified and better structured in the revised manuscript:

Section 2.1 Disaster event dataset and filtering strategy:

“Using a threshold of 50 fatalities (Total Deaths), this study identifies 344 major disaster events for detailed statistical analysis. The selection of “Total Deaths” as the primary criterion reflects its superior reliability and objectivity compared to other metrics. Population-based indicators in EM-DAT, such as “No. Affected” and “No. Homeless”, are less reliable due to varying definitions and estimation methods, which often depend heavily on other metrics or overall event assessments (UNDP, 2004; Newman and Noy, 2023; Peduzzi et al., 2009). Similarly, economic indicators like “Total Damage” introduce significant statistical bias due to varying socio-economic capacities across countries with different development levels (UNDP, 2004; Peduzzi et al., 2009). The ≥ 50 fatalities threshold further excludes small-scale events that may be more prone to under-reporting, thereby improving representativeness for global-scale analysis.”

Section 5 Discussion:

“Nevertheless, several methodological limitations should be acknowledged: 1) Event filtering: Using a ≥ 50 fatality threshold excludes smaller but potentially accumulative events to reduce under-reporting bias and improve international comparability.”

Building upon this foundation, we further conducted a sensitivity analysis to examine how varying the fatality threshold influences ARI outcomes in *Section 3.3.1 Sensitivity to the fatality threshold*. Specifically, we examined how the number of identified “major disaster events” and the resulting ARI patterns change when the threshold is set to 40, 60 and 70 fatalities. As shown in *Table 4*, preliminary results indicate that while the absolute number of events changes significantly with different thresholds, the spatial clustering of high-risk grids remains broadly consistent. This suggests that ARI is robust in identifying key high-risk areas, even though the absolute magnitude of ARI varies with the threshold.

Revised content now appears as follows:

“The baseline ARI configuration uses a threshold of ≥ 50 fatalities (Total Deaths) to select major disaster events (Section 2.1). To test the sensitivity of ARI to this choice, alternative thresholds of 40, 60 and 70 fatalities are considered. For each threshold, the number of major events and the ARI values of grids are recalculated. Table 4 summarizes the total number of major events and the top 10 high-ARI grids identified under each scenario, together with their rankings under the baseline threshold.

As expected, lowering the threshold to 40 fatalities increases the number of major events to 406, whereas raising it to 60 and 70 fatalities reduces the counts to 282 and 242, respectively. Despite these changes in absolute event numbers and ARI magnitudes, the spatial pattern of high-ARI grids remains broadly consistent. In all scenarios, the highest-ranking grids are located in South and Southeast Asia, particularly in regions covering northern India, Bangladesh and the Himalayan belt. Moreover, many of the top 10 grids under the alternative thresholds coincide with those in the baseline configuration, often with only minor changes in rank. This indicates that the identification of key high-risk areas is not strongly dependent on the exact choice of fatality threshold, although the absolute ARI values vary with the number of included events.”

Table 4: Comparison of ARI results under different thresholds (40, 60, and 70 fatalities).

Fatality threshold	Major disaster events	High-risk grids (top 10)					
		ARI ranking	Baseline ranking	Latitude range (°)	Longitude range (°)	Region	ARI
40	406	1	1	25 ~ 30	90 ~ 95	Asia	359.26
		2	2	25 ~ 30	80 ~ 85	Asia	311.2
		3	3	30 ~ 35	70 ~ 75	Asia	306.02
		4	5	25 ~ 30	85 ~ 90	Asia	268.51
		5	10	30 ~ 35	75 ~ 80	Asia	249.12
		6	4	10 ~ 15	120 ~ 125	Asia	234.3
		7	14	15 ~ 20	120 ~ 125	Asia	234.3
		8	6	-10 ~ -5	105 ~ 110	Asia	217.5
		9	7	30 ~ 35	100 ~ 105	Asia	216.8
		10	8	5 ~ 10	120 ~ 125	Asia	187.44
60	282	1	1	25 ~ 30	90 ~ 95	Asia	276.22
		2	2	25 ~ 30	80 ~ 85	Asia	269.68
		3	3	30 ~ 35	70 ~ 75	Asia	253.12
		4	4	10 ~ 15	120 ~ 125	Asia	234.3
		5	9	25 ~ 30	65 ~ 70	Asia	185.15
		6	6	-10 ~ -5	105 ~ 110	Asia	174
		7	10	30 ~ 35	75 ~ 80	Asia	166.08
		8	7	30 ~ 35	100 ~ 105	Asia	162.6
		9	11	20 ~ 25	90 ~ 95	Asia	159.55
		10	5	25 ~ 30	85 ~ 90	Asia	155.61
70	242	1	1	25 ~ 30	90 ~ 95	Asia	234.7
		2	4	10 ~ 15	120 ~ 125	Asia	234.3
		3	2	25 ~ 30	80 ~ 85	Asia	228.16
		4	3	30 ~ 35	70 ~ 75	Asia	200.22
		5	9	25 ~ 30	65 ~ 70	Asia	158.7
		6	5	25 ~ 30	85 ~ 90	Asia	155.61
		7	8	5 ~ 10	120 ~ 125	Asia	140.58
		8	7	30 ~ 35	100 ~ 105	Asia	135.5
		9	17	15 ~ 20	75 ~ 80	Asia	124.56
		10	11	20 ~ 25	90 ~ 95	Asia	118.03

We believe this additional analysis strengthens the methodological transparency of the study and demonstrates that our main conclusions are not overly sensitive to the specific threshold chosen. We thank the referee again for this important suggestion, which has helped us to improve the rigor and credibility of our work.

2.3. Similarly, the choice of 5°×5° spatial grids requires justification. Alternative grid sizes should be tested to assess robustness.

We sincerely thank the referee for this insightful comment regarding the spatial resolution of 5°×5° grids. We agree that the choice of spatial units is critical in disaster risk assessment, as it directly affects the interpretation of results and their potential applications.

Our decision to adopt a 5°×5° latitude-longitude grid was based on several methodological considerations, which we further clarify below:

1) As discussed in the manuscript, data availability and consistency remain a major challenge when working with EM-DAT,

as disaster events are often recorded without precise geocoded information, especially in developing regions. While some records, such as earthquakes, include specific latitude-longitude coordinates, many others are documented only through textual descriptions of affected places (e.g., cities or villages), and large-scale hazards like floods and storms frequently span extensive areas without precise boundaries. This lack of standardization makes finer spatial resolutions (e.g., $1^\circ \times 1^\circ$ or smaller) impractical, as they would introduce excessive uncertainty and fragmentation due to missing or imprecise georeferencing. By contrast, the $5^\circ \times 5^\circ$ grid offers a pragmatic compromise, enabling systematic allocation of events to standardized representative coordinates while ensuring a balance between spatial detail and global data reliability.

2) As also noted in the manuscript, an excessively large grid size would dilute spatial contrasts, while an overly small grid would lead to two main issues: difficulty in adequately characterizing large-scale hazards that extend across wide regions, thereby compromising the statistical independence of grid-level events; and amplification of relative errors, as changes in exposure (e.g., population distribution) and event allocation become increasingly sensitive to minor variations at finer scales. The $5^\circ \times 5^\circ$ grid therefore represents a compromise that ensures statistical robustness without sacrificing spatial interpretability.

3) The chosen $5^\circ \times 5^\circ$ resolution is consistent with common practice in global disaster risk research, where results based on finer grids can be systematically aggregated into $5^\circ \times 5^\circ$ units. This ensures comparability with existing multi-hazard indices and facilitates integration with remote sensing applications, which are typically designed for regional to continental scales. Moreover, the primary aim of ARI is to highlight broad regional disparities and accumulative disaster risks rather than provide localized hazard mapping. A $5^\circ \times 5^\circ$ resolution is therefore well-suited for revealing clustering patterns across continents and large transboundary regions, aligning with the scope and purpose of this preliminary study.

That said, we recognize the referee's concern about the potential arbitrariness of this choice. This reasoning is now clarified and better structured in *Section 2.4 Spatial aggregation for ARI computation* of the revised manuscript:

“Specific statistical calculations for ARI depend on the selection of region j , which enables for the possibility of using sub-national regions as basic assessment units. However, the choice of region size is critical. An excessively large region size diminishes the advantages of using latitude-longitude regions as assessment units. Conversely, an overly small size adversely affects the results in two ways: 1) Due to the extensive impact range of large-scale hazards, which can extend over thousands of square kilometers, it becomes challenging to adequately characterize hazard types (e.g. floods and storms) that affect large areas. This limitation inhibits an effective representation of the independence of regional disaster events. 2) Relative errors will be amplified, leading to a decrease in the overall usability and reliability of data. Additionally, changes in exposure due to variations in population distribution and hazard zones over a given period will become more pronounced (Dilley, 2005), significantly affecting the timeliness of related data. Moreover, as disaster risk assessments become more refined at smaller scales (Feng et al., 2017; Smith et al., 2019), the pursuit of excessively small assessment unit sizes on a global scale clearly no longer meets practical application needs. Therefore, this study ultimately selects $5^\circ \times 5^\circ$ latitude-longitude grid as the minimum unit for regional division. This resolution provides a balance between global coverage, data reliability and compatibility with remote sensing applications. Robustness checks using alternative grid sizes ($2.5^\circ \times 2.5^\circ$ and $10^\circ \times 10^\circ$) are presented in Section 3.3.”

Building upon this foundation, we further conducted a sensitivity analysis to examine how varying the grid resolution influences ARI outcomes in *Section 3.3.2 Sensitivity to grid resolution*. In response to the referee's suggestion, we conducted a robustness check using alternative grid sizes ($2.5^\circ \times 2.5^\circ$ and $10^\circ \times 10^\circ$), as shown in *Figure 7*. Results show that while the number of events allocated to each grid naturally changes with resolution, the overall spatial distribution of high-ARI regions, such as South Asia, Southeast Asia, and Central America, remains consistent across scales. This indicates that the observed geographic clustering of accumulative disaster risks is not an artifact of the chosen grid size, but rather a robust feature of the underlying disaster data.

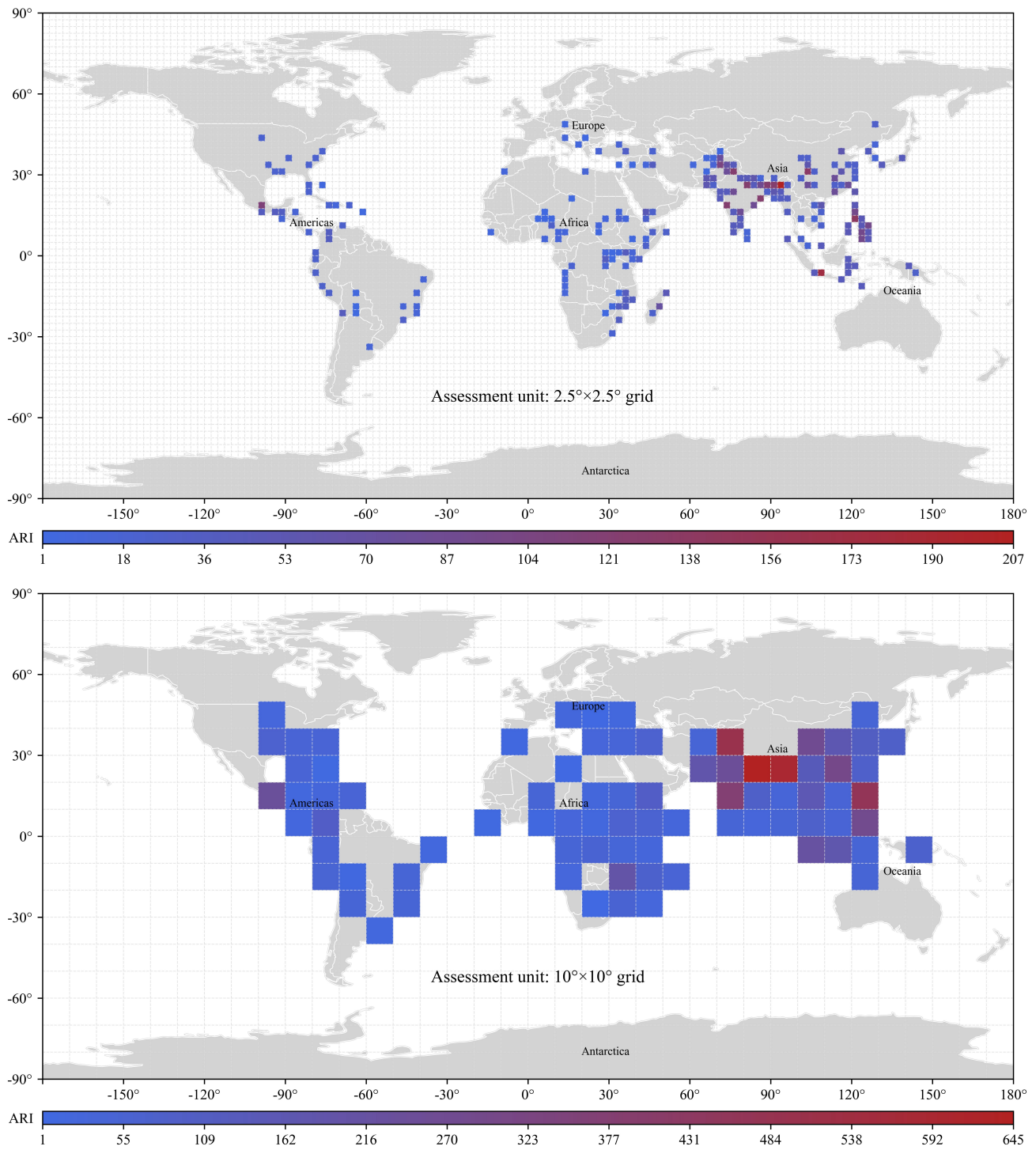


Figure 7: Comparison of ARI results under different spatial grids (2.5°×2.5° and 10°×10°).

Revised content now appears as follows:

“The baseline ARI computation adopts 5°×5° latitude-longitude grids as the minimum spatial unit (Section 2.4). To evaluate the influence of spatial resolution, additional ARI maps are generated using finer (2.5°×2.5°) and coarser (10°×10°) grids. Figure 7 compares the resulting ARI distributions for the two resolutions.

When a finer 2.5°×2.5° grid is used, the number of non-zero ARI cells increases, and localized variations within large countries become more pronounced. Conversely, the 10°×10° grid leads to fewer non-zero cells and smoother spatial gradients. However, across all three resolutions, the main clusters of high ARI consistently appear in the same broad regions, namely South and Southeast Asia and parts of Central America. The locations of the most prominent hotspots remain stable, even though their exact shapes and ARI magnitudes change with grid size. This suggests that ARI is robust in identifying the principal geographic concentrations of accumulated disaster risk at the global scale.”

We are grateful to the referee for this constructive suggestion, which has allowed us to enhance the methodological transparency and credibility of the study. To further address the referee’s concerns regarding parameter choices, we have implemented both fatality-threshold and grid-resolution sensitivity tests and integrated them into a unified robustness framework (*Section 3.3 Validation and robustness checks*). The results consistently show that although absolute ARI values vary with parameter settings, the spatial configuration of high-risk clusters remains stable across reasonable variations. Taken together, these analyses confirm that the identification of global ARI hotspots is not an artifact of a specific parameter choice, but a persistent pattern of recent disaster accumulation. This provides stronger methodological credibility for ARI as a global comparative indicator.

This argument now appears explicitly in multiple places in the revised manuscript, such as:

Section 3.3 Validation and robustness checks:

“Given that ARI is a newly proposed index, it is important to assess the robustness of its spatial patterns with respect to key methodological choices. This section examines the sensitivity of ARI to the fatality threshold used to define major disaster events and the spatial resolution of the latitude-longitude grids. These analyses are intended to evaluate whether the main spatial patterns identified in Sections 3.1 and 3.2 are stable under reasonable variations.”

“Taken together, the sensitivity analyses on the fatality threshold and grid resolution indicate that the main spatial features of ARI, especially the identification of high-risk clusters in developing regions, are relatively insensitive to reasonable variations in key methodological parameters. ARI is therefore suitable for comparative hotspot identification at the global scale, while detailed local interpretation should consider the uncertainties discussed in Section 5.”

Section 5 Discussion:

“Sensitivity analyses show that ARI is relatively robust to reasonable changes in both the fatality threshold and grid resolution. While absolute ARI values vary, the spatial configuration of high-risk grids remains stable, confirming that ARI is more suitable for hotspot identification than for precise quantitative ranking.”

3. Use of WRI and ARI definitions: the ARI is described as a simple aggregation of the WRI within fixed spatial units. Since the WRI methodology is not fully transparent in terms of input data and weighting, the implications of spatial aggregation are difficult to interpret.

We sincerely thank the referee for raising this important methodological concern. We fully agree that the WRI has inherent limitations, particularly regarding the transparency of its weighting schemes and the heterogeneity of its input datasets. Our intention in introducing the ARI is not to replicate or directly aggregate WRI values, but rather to integrate them as a weighting factor in order to highlight the accumulative effects of repeated major disaster events.

To avoid misunderstanding, we emphasize that ARI is not the spatial aggregation of WRI, and WRI does not contribute directly to the quantitative risk value. Instead, ARI is primarily driven by disaster event frequency, with WRI only acting as a contextual weighting modifier. Specifically, ARI is constructed from two complementary components:

- 1) The frequency of major disaster events (defined by our threshold criterion), which captures the recent and recurrent hazard impacts at the latitude-longitude grid scale.
- 2) The country-level WRI value, which serves as a scaling factor to adjust for differences in exposure and vulnerability across regions. In this way, WRI does not determine the event frequency itself but modulates its relative impact, ensuring that the same number of events in two different countries will not be treated as equivalent if their social and institutional resilience differs.

We acknowledge the referee’s concern that the opacity of WRI’s internal weighting may introduce interpretive challenges. To address this, we have revised the manuscript to clarify the conceptual role of WRI within ARI: it functions not as a direct additive component, but as a contextual adjustment that reflects underlying socio-economic conditions. Furthermore, ARI is

explicitly designed as a complementary measure to WRI rather than a replacement. Whereas WRI provides a holistic but relatively static long-term assessment, ARI emphasizes the short-term accumulative stress induced by multiple disasters, which WRI alone cannot capture.

To improve transparency, we have added a methodological note in the revised manuscript, explicitly stating that: ARI results are primarily driven by event data from EM-DAT, while WRI values serve as modifiers of relative risk. The aggregation to $5^{\circ} \times 5^{\circ}$ grids is applied only after this adjustment, allowing consistent spatial comparisons. The interpretive focus of ARI should therefore be placed on the patterns of accumulative disaster occurrence, with WRI serving only as a secondary factor contextualizing exposure and vulnerability.

The conceptual clarification and methodological separation of WRI and ARI have been explicitly incorporated into the revised manuscript, such as:

Section 2.3.2 Accumulated Risk Index (proposed index):

“In this study, WRI is not recomputed or modified. Instead, it is adopted as a country-level reference index and contextual weighting factor that encapsulates exposure and vulnerability conditions. The proposed ARI builds on WRI by combining it with the recent occurrence of major disaster events at the grid scale, using WRI as a scaling factor rather than as a quantity to be spatially aggregated on its own.”

“Building on the reference index introduced above, ARI is formulated as a grid-based measure that integrates recent major disaster events with country-level contextual information. Conceptually, ARI integrates two components: 1) the frequency of major disaster events (defined by the filtering strategy in Section 2.1) occurring within each latitude-longitude grid cell over the period 2013-2023; and 2) the country-level WRI value associated with each event, which adjusts for differences in exposure and vulnerability among affected regions. In this way, ARI emphasizes the short-term accumulation of high-impact shocks in contexts with varying underlying risk conditions.”

“In ARI, these advantages are retained while shifting the focus from long-term structural risk to the accumulative impacts of recent major events.”

“For grid cells intersecting multiple countries, each event is linked to the WRI of the country recorded in EM-DAT as part of the affected area. This event-level association allows ARI to reflect both the spatial distribution of recent major events and the country-specific risk context.”

Section 5 Discussion:

“ARI highlights areas where high-fatality disaster events have clustered within a relatively short period and where underlying country-level exposure and vulnerability conditions amplify their impacts. The index does not aim to replicate or replace long-term structural risk metrics such as WRI, but to provide a complementary perspective focused on recent accumulative shocks.”

“(3) Data dependency: ARI inherits inherent uncertainties from EM-DAT (e.g., reporting heterogeneity) and WRI (e.g., indicator transparency). Thus, ARI should be interpreted as a relative index of recent accumulative risk, best complemented by local datasets or expert analysis in operational decision-making.”

“Both EM-DAT and WRI provide harmonised global datasets, yet gaps and inconsistencies remain unavoidable.”

“Likewise, while WRI provides a consolidated view of exposure and vulnerability, some underlying datasets are updated asynchronously and are not fully documented. Because ARI combines event occurrence with WRI only as a contextual weighting factor, rather than spatially interpolating WRI, its results should be read as relative indications of where recent high-fatality events intersect with risk contexts, not as precise estimates of absolute risk levels.”

Section 6 Conclusions:

“While not a replacement for structural risk indices, ARI offers a complementary perspective focused on near-term accumulative impacts.”

We believe this clarification will help distinguish the distinct functions of ARI and WRI and avoid the misunderstanding that ARI is merely a spatial aggregation of WRI values. In future work, we also plan to test the sensitivity of ARI to alternative risk indices, such as INFORM or HDI-based indicators, to further assess the robustness of results under different socio-economic adjustment schemes. In *Section 5 Discussion*, we now clearly state that:

“Future work could extend ARI by testing composite severity indicators when data quality allows, exploring alternative risk indices alongside WRI, and applying temporal weighting schemes to relate recent accumulative patterns to longer-term trends.”

4. Results and interpretation:

4.1. The results are descriptive and primarily spatial distributions of event types. They do not sufficiently engage with the implications of the findings.

We sincerely thank the referee for this constructive comment. We agree that in the original manuscript, the results were largely descriptive, with an emphasis on mapping the spatial distributions of disaster events and ARI values. While these visualizations are informative, we acknowledge that the interpretive depth regarding their broader implications was limited.

In response, we have implemented substantial structural and content-level revisions, rather than making minor additions. The revised manuscript now includes a fully reconstructed *Section 3 Results*, separating analytical outcomes from methodology and enabling clearer interpretation. This section is newly divided into:

3.1 Spatiotemporal characteristics of recent disaster events

3.2 Global ARI distribution and hotspot identification

3.3 Validation and robustness checks

This restructuring clarifies the analytical logic and places ARI results in a dedicated space, improving narrative coherence and visibility of key findings. Additionally, we added a new *Section 5 Discussion*, which moves beyond descriptive reporting and synthesizes broader implications regarding vulnerability, policy relevance, and scientific insight.

Importantly, beyond narrative reorganization, we have strengthened the interpretive content of the results themselves, particularly in *Section 3.2 Global ARI distribution and hotspot identification*. In addition to the global comparison between ARI and Equivalent WRI (*Figure 3*), we introduced a new regional zoom-in analysis (*Figure 4*) that explicitly demonstrates how ARI moves beyond descriptive spatial distributions.

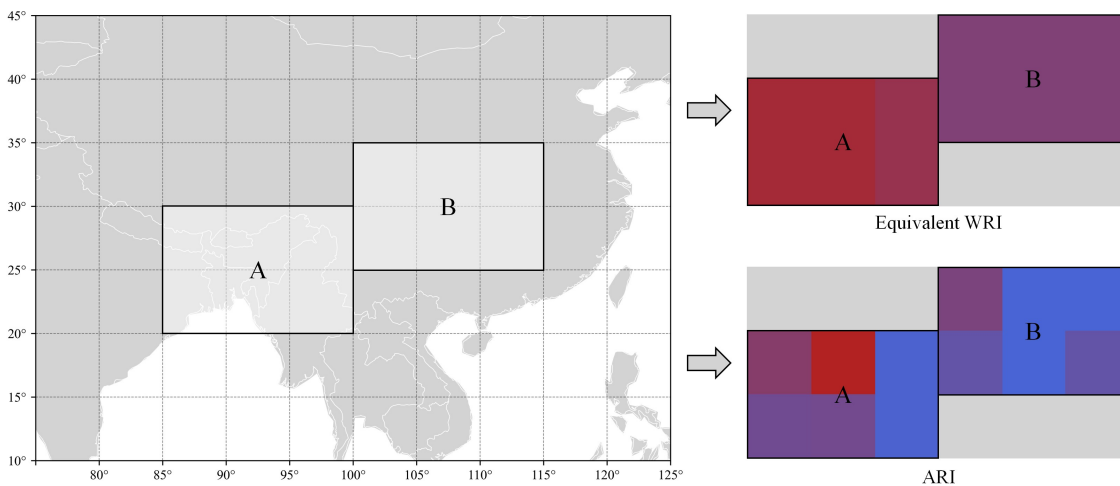


Figure 4: Local zoom-in comparison of ARI and Equivalent WRI in 5°×5° latitude-longitude grids.

Specifically, *Figure 4* provides a grid-aligned comparison for a representative high-risk region encompassing the Himalayan belt and South-Southeast Asia. Two contrasting situations are illustrated:

1) A transboundary corridor (Window A), where Equivalent WRI primarily reflects country-level background differences,

whereas ARI further differentiates neighbouring $5^{\circ} \times 5^{\circ}$ grid cells according to the spatial concentration of recent major disaster events within the same hazard corridor.

2) A large-country case (Window B) entirely within China, where Equivalent WRI becomes spatially uniform by construction, while ARI reveals marked sub-national heterogeneity, identifying localized hotspots associated with recent event accumulation.

This zoom-in comparison demonstrates that ARI does not merely reproduce spatial patterns at a finer resolution, but rather adds interpretive value by revealing sub-national and transboundary contrasts that are obscured in long-term, country-level indices. As such, the results now engage directly with the implications of accumulated disaster risk, rather than remaining purely descriptive.

The revised results and discussion therefore engage explicitly with the significance of the findings from four complementary perspectives:

1) Link ARI patterns to regional socio-economic contexts: We emphasize that the clustering of high ARI grids in Asia, particularly South and Southeast Asia, reflects not only physical hazard exposure but also structural vulnerabilities such as high population density, rapid urbanization, and limited disaster preparedness. This highlights regions where accumulative impacts may overwhelm local coping capacity.

2) Highlight policy relevance and decision-making implications: We discuss how ARI can inform priority setting for international humanitarian aid and disaster risk reduction planning. For instance, regions repeatedly identified as ARI hotspots (e.g., the Himalayan belt and Southeast Asia) may warrant greater investment in resilience-building and recovery assistance, as these areas face persistent accumulative stress rather than isolated events.

3) Situate ARI in relation to existing indices: We point out that traditional long-term, country-level indices such as WRI may obscure localized accumulative impacts in areas where recent disaster clustering has significantly eroded resilience. This limitation becomes particularly pronounced within large countries such as China and India, where disaster exposure vary widely across different subnational regions. By comparing ARI with WRI, we reveal cases where discrepancies highlight blind spots in current global risk assessment practices, especially in capturing short-term accumulative impacts at finer spatial scales.

4) Explore scientific implications: We discuss how ARI contributes to ongoing debates about cascading and compounding risks under climate change. The observed geographic clustering suggests that certain regions are entering feedback loops where repeated disasters exacerbate vulnerability, leading to non-linear risk accumulation.

A representative excerpt from the revised manuscript illustrating this enhanced interpretation is included below:

Section 3.1 Spatiotemporal characteristics of recent disaster events:

“Overall, major disaster events exhibit marked spatial clustering rather than a random global distribution. High concentrations of events are observed along tectonically active belts and in monsoon-affected regions, particularly in and around South and Southeast Asia, as well as in parts of Central America. These clusters often extend across national borders, reflecting the fact that many large-scale hazards, such as riverine floods and tropical cyclones, inherently transcend administrative boundaries. The observed patterns imply that assessments based solely on national units may obscure important cross-border and subnational variations in disaster occurrence. This provides empirical motivation for adopting a latitude-longitude grid as the basic assessment unit in the ARI, thereby aligning the spatial representation of risk more closely with the geographic footprint of major disaster events.”

Section 3.2 Global ARI distribution and hotspots:

“While both ARI and Equivalent WRI highlight elevated risk in parts of Asia, their spatial contrasts differ: ARI produces a more pronounced differentiation between high- and low-risk grids, particularly in areas where multiple major disasters have clustered in recent years. This suggests that the incorporation of recent event accumulation can sharpen spatial contrasts

relative to long-term, country-level indices.”

“Two contrasting situations are illustrated. In Window A, which spans multiple countries, Equivalent WRI mainly reflects country-level background differences, whereas ARI further differentiates neighbouring 5°×5° cells according to the spatial concentration of recent major events within the same transboundary corridor. In Window B, which lies entirely within China, Equivalent WRI becomes spatially uniform by construction, while ARI still exhibits marked intra-national heterogeneity, indicating sub-national hotspots associated with recent event accumulation. Together, the zoom-in comparison is consistent with the interpretation drawn from Figure 3 that incorporating near-term event accumulation can sharpen spatial contrasts relative to long-term, country-level reference indices, particularly along transboundary hazard corridors and within large countries where sub-national variability is important.”

“Asia clearly dominates in terms of both the number of non-zero grids and the magnitude of ARI values, reflecting the combined influence of high hazard exposure and large populations in many developing countries. In other continents, non-zero ARI grids are fewer and generally of lower magnitude, although localized clusters of high ARI can still be identified, for instance in parts of Central America.”

“Based on their geographic coordinates, these high-risk grids involve ten countries: India, China, Pakistan, the Philippines, Nepal, Bhutan, Indonesia, Myanmar, Bangladesh and Mexico. All of these are developing countries characterized by high population density, relatively weak infrastructure and unequal distribution of social resources (Koks et al., 2019; Rentschler et al., 2022; Berlemann et al., 2018).”

Section 4 Remote sensing data integration for ARI hotspots:

“The ARI results presented in Section 3.2 identify high-risk latitude-longitude grids where recent major disaster events have accumulated and where local resilience may have been significantly eroded. These grids provide spatial focal points for detailed risk assessment and management at finer scales.”

Section 5 Discussion:

“Compared with a grid-based equivalent of the WRI, ARI yields sharper contrasts between hotspots and low-activity areas, indicating that near-term accumulative effects can significantly modify the perceived risk pattern at sub-national and transboundary scales. The results therefore underscore the value of combining event-driven metrics with background socio-economic conditions for identifying areas where resilience may have deteriorated more rapidly than long-term indices would suggest.”

“Sensitivity analyses show that ARI is relatively robust to reasonable changes in both the fatality threshold and grid resolution. While absolute ARI values vary, the spatial configuration of high-risk grids remains stable, confirming that ARI is more suitable for hotspot identification than for precise quantitative ranking.”

“Thus, ARI should be interpreted as a relative index of recent accumulative risk, best complemented by local datasets or expert analysis in operational decision-making.”

“Despite these caveats, the combined ARI-remote sensing framework offers practical value. For policy and humanitarian actors, ARI can help flag areas where repeated high-impact events may have eroded resilience and where more detailed assessments and targeted support are warranted. For the scientific community, the 3H Dataset provides standardized, sub-metre imagery for ARI hotspots in developing countries, lowering the barrier for remote sensing-based studies on damage detection, exposure mapping and recovery monitoring.”

In summary, while our study remains preliminary, we believe the revised manuscript now better engages with the broader implications of the findings. By framing ARI not only as a descriptive spatial tool but also as an instrument to highlight neglected dimensions of accumulated risk, we aim to strengthen its scientific and policy relevance.

4.2. The conclusions (Line 188) are too strong relative to the evidence provided. The statement that “traditional risk assessment methods may underestimate risks in high-risk areas and overestimate them in low-risk areas” is not convincingly

supported by the presented results. A more cautious and better-argued interpretation is required.

We sincerely thank the referee for pointing out this important issue. We agree that the original formulation of our conclusion, that “traditional risk assessment methods may underestimate risks in high-risk areas and overestimate them in low-risk areas”, was expressed too strongly relative to the evidence provided in the manuscript. Our results indeed indicate that the ARI amplifies the contrast between high- and low-risk areas compared with the WRI, but we acknowledge that this does not constitute definitive proof of systematic under- or overestimation by traditional indices.

In the revised manuscript, we have therefore adopted a more cautious and better contextualized interpretation. Specifically, we now emphasize that: ARI highlights sharper disparities between regions than WRI, particularly in areas where repeated disaster clustering has significantly eroded resilience. Such discrepancies are especially pronounced in large countries such as China and India, where disaster exposure vary substantially across subnational regions. This suggests that country-level, long-term indices like WRI may sometimes smooth over these localized accumulative impacts of recent disasters, potentially creating blind spots in global comparative risk assessments. We explicitly note that these findings are preliminary and require further validation through comparative studies with other risk indices and datasets.

In particular, the revised results in *Section 3.2 Global ARI distribution and hotspots*, including a newly added regional zoom-in comparison (*Figure 4*), provide concrete evidence that ARI highlights spatial contrasts associated with recent disaster accumulation that are not resolved by country-level reference indices. These comparisons demonstrate that ARI does not imply that traditional indices are incorrect, but rather that long-term, country-level indices may smooth over sub-national and transboundary contrasts when recent event clustering is strong, whereas ARI is more sensitive to near-term accumulative effects at finer spatial scales.

Accordingly, in the revised manuscript, we have replaced the previous strong statement with a more cautious and balanced formulation:

Section 3.1 Spatiotemporal characteristics of recent disaster events:

“These clusters often extend across national borders, reflecting the fact that many large-scale hazards, such as riverine floods and tropical cyclones, inherently transcend administrative boundaries. The observed patterns imply that assessments based solely on national units may obscure important cross-border and subnational variations in disaster occurrence. This provides empirical motivation for adopting a latitude-longitude grid as the basic assessment unit in the ARI, thereby aligning the spatial representation of risk more closely with the geographic footprint of major disaster events.”

Section 3.2 Global ARI distribution and hotspots:

“While both ARI and Equivalent WRI highlight elevated risk in parts of Asia, their spatial contrasts differ: ARI produces a more pronounced differentiation between high- and low-risk grids, particularly in areas where multiple major disasters have clustered in recent years. This suggests that the incorporation of recent event accumulation can sharpen spatial contrasts relative to long-term, country-level indices.”

“Together, the zoom-in comparison is consistent with the interpretation drawn from *Figure 3* that incorporating near-term event accumulation can sharpen spatial contrasts relative to long-term, country-level reference indices, particularly along transboundary hazard corridors and within large countries where sub-national variability is important.”

Section 5 Discussion:

“Compared with a grid-based equivalent of the WRI, ARI yields sharper contrasts between hotspots and low-activity areas, indicating that near-term accumulative effects can significantly modify the perceived risk pattern at sub-national and transboundary scales.”

Section 6 Conclusions:

“Compared with the long-term WRI, ARI provides sharper spatial contrasts and highlights sub-national hotspots where repeated shocks may have accelerated resilience loss. While not a replacement for structural risk indices, ARI offers a

complementary perspective focused on near-term accumulative impacts.”

We believe that these revisions bring the conclusions into closer alignment with the evidence presented, while still underscoring the conceptual contribution of ARI. Future work will be required to assess whether such contrasts correspond to localized risk underestimation, overestimation, or simply different sensitivities across indices, using time-series resilience indicators, sub-national exposure data, and alternative risk frameworks (e.g., INFORM, HDI-based metrics).

4.3. A dedicated Discussion section should be added, critically evaluating the robustness of the findings, uncertainties in the data, and implications for disaster risk management.

We sincerely thank the referee for this constructive suggestion. We fully agree that the manuscript in its original form did not sufficiently provide a critical discussion of the robustness of our findings, the uncertainties associated with the dataset, and the broader implications for disaster risk management.

In response, we have added a dedicated discussion section (*Section 5 Discussion*) in the revised manuscript to allow deeper interpretation beyond descriptive spatial patterns. In addition to the previously mentioned content, this new section mainly aims to explicitly address the following aspects:

- 1) Robustness of findings: We discuss how the key spatial patterns identified by the ARI remain consistent even when alternative thresholds (e.g., 40, 60, and 70 fatalities) or different spatial units (e.g., $2.5^{\circ} \times 2.5^{\circ}$ and $10^{\circ} \times 10^{\circ}$ latitude-longitude grids) are considered. We also highlight that ARI's comparative role, rather than its absolute values, is the most robust outcome of this exploratory study.
- 2) Uncertainties in the data: We acknowledge the limitations of the EM-DAT database, including missing or incomplete economic loss data, non-standardized georeferencing of disaster locations, and potential reporting biases in earlier years and in developing countries. In addition, we recognize the uncertainties inherent in the WRI, which integrates multiple dimensions, yet the underlying data quality and weighting scheme are not always fully transparent. Taken together, these uncertainties highlight the need for cautious interpretation of ARI results. They also reinforce our intention that ARI should be viewed not as a definitive measure of disaster risk, but as a complementary tool to existing indices, designed to foreground recent accumulative impacts and to encourage further refinement in global risk assessment practices.
- 3) Implications for disaster risk management: We emphasize that ARI is not intended to replace existing global indices (such as WRI), but rather to complement them by highlighting accumulative disaster impacts at finer geographic scales. This has practical relevance for policymakers and humanitarian organizations, particularly in identifying “hotspot” regions where repeated disasters strain resilience. The integration of ARI outputs with open-source remote sensing datasets (the 3H Dataset) is also discussed as a way to support applied monitoring and early response systems.

Revised content now appears as follows:

“ARI highlights areas where high-fatality disaster events have clustered within a relatively short period and where underlying country-level exposure and vulnerability conditions amplify their impacts. The index does not aim to replicate or replace long-term structural risk metrics such as WRI, but to provide a complementary perspective focused on recent accumulative shocks. Compared with a grid-based equivalent of the WRI, ARI yields sharper contrasts between hotspots and low-activity areas, indicating that near-term accumulative effects can significantly modify the perceived risk pattern at sub-national and transboundary scales. The results therefore underscore the value of combining event-driven metrics with background socio-economic conditions for identifying areas where resilience may have deteriorated more rapidly than long-term indices would suggest.

Sensitivity analyses show that ARI is relatively robust to reasonable changes in both the fatality threshold and grid resolution. While absolute ARI values vary, the spatial configuration of high-risk grids remains stable, confirming that ARI is more suitable for hotspot identification than for precise quantitative ranking. Nevertheless, several methodological limitations

should be acknowledged: 1) Event filtering: Using a ≥ 50 fatality threshold excludes smaller but potentially accumulative events to reduce under-reporting bias and improve international comparability. 2) Spatial simplification: The assignment of a single representative coordinate per event inevitably simplifies complex hazard footprints, particularly for floods and storms with broad spatial extents. 3) Data dependency: ARI inherits inherent uncertainties from EM-DAT (e.g., reporting heterogeneity) and WRI (e.g., indicator transparency). Thus, ARI should be interpreted as a relative index of recent accumulative risk, best complemented by local datasets or expert analysis in operational decision-making.

Both EM-DAT and WRI provide harmonised global datasets, yet gaps and inconsistencies remain unavoidable. The standardization of textual location descriptions into point coordinates reduces heterogeneity but cannot eliminate spatial uncertainty in extended-impact events. Likewise, while WRI provides a consolidated view of exposure and vulnerability, some underlying datasets are updated asynchronously and are not fully documented. Because ARI combines event occurrence with WRI only as a contextual weighting factor, rather than spatially interpolating WRI, its results should be read as relative indications of where recent high-fatality events intersect with risk contexts, not as precise estimates of absolute risk levels.

Despite these caveats, the combined ARI-remote sensing framework offers practical value. For policy and humanitarian actors, ARI can help flag areas where repeated high-impact events may have eroded resilience and where more detailed assessments and targeted support are warranted. For the scientific community, the 3H Dataset provides standardized, sub-metre imagery for ARI hotspots in developing countries, lowering the barrier for remote sensing-based studies on damage detection, exposure mapping and recovery monitoring.

Future work could extend ARI by testing composite severity indicators when data quality allows, exploring alternative risk indices alongside WRI, and applying temporal weighting schemes to relate recent accumulative patterns to longer-term trends. Improved geocoding and the use of hazard footprints derived from Earth observation would support refined spatial units in selected regions. Closer coupling between ARI outputs and observed damage or recovery trajectories would also help to evaluate and further develop accumulated risk indicators under changing climate and socio-economic conditions.”

This section is written not only to interpret the results, but also to critically examine their robustness, limitations, and practical significance, responding directly to the request for a more analytical and reflective discussion rather than descriptive reporting. By explicitly structuring this new discussion section, we believe the revised manuscript now provides a more balanced and critical evaluation of both the strengths and limitations of our approach, while clearly outlining its added value and potential applications.

5. Integration of remote sensing data: the role of remote sensing imagery in the analysis is unclear. While the data collection effort is commendable, the manuscript does not explain how these datasets contribute to or validate the ARI results. This section currently feels disconnected from the core analysis.

We sincerely thank the referee for highlighting this important issue. We acknowledge that in the submitted version, the role of remote sensing data was not sufficiently articulated, which may have led to the impression that this component was disconnected from the ARI analysis. We appreciate this insightful comment and have made substantial revisions to clarify the conceptual linkage and analytical logic between ARI and remote sensing integration.

Our intention in introducing the remote sensing datasets, compiled into the open-source “3H Dataset”, is not to directly validate the ARI at this preliminary stage, but rather to establish a foundation for future methodological integration. Specifically, the ARI is designed to identify spatial clusters of accumulated disaster risks, while the 3H Dataset provides normalized and accessible remote sensing data for those high-risk grids, with a focus on developing countries. By combining these two elements, we aim to create a bridge between macro-level risk assessment and micro-level hazard detection. In practice, the ARI can help to prioritize where remote sensing efforts should be focused (e.g., monitoring urban growth in hazard-prone regions, or tracking recovery in repeatedly affected areas), while the 3H Dataset can support the development

and training of remote sensing-based models to detect damage patterns and validate risk assumptions at finer spatial scales.

To address the referee's concern directly, we have implemented structural revisions rather than textual additions alone:

1) Reorganized remote sensing content into a dedicated *Section 4 Remote sensing data integration for ARI hotspots* after the ARI results, ensuring that remote sensing now functions as an application extension of ARI rather than an isolated dataset module. The previous single section has been restructured into a three-part logical framework:

4.1 Data sources and spatial coverage

4.2 Supplementing data gaps in high-risk grids

4.3 Standardization and construction of the 3H Dataset

This adjustment explicitly positions remote sensing as post-ARI integration for hotspot grids, directly addressing the referee's "disconnection" concern.

2) Added explanatory text to clarify how ARI outputs guide the remote sensing dataset design and future validation pathways. Key statements now appear in the revised manuscript, including:

Section 1.2 Motivation and Objectives:

"In addition, in order to facilitate further remote sensing-based research, this study systematically integrates and supplements existing sub-meter visible spectrum remote sensing imagery for ARI-identified high-risk grids, and publishes them as an open-source dataset after standardization. This dataset provides essential support for future studies at finer spatial scales."

Section 2.5 Analytical workflow:

"3) Development of the remote sensing dataset: Grids with high ARI values, interpreted as hotspots of accumulated risk, are linked to sub-meter satellite imagery. Existing open-source data are compiled and supplemented where necessary to form the proposed dataset (Section 4).

This structured workflow provides a unified view of how statistical disaster records, global risk indices, and remote sensing observations jointly contribute to the assessment of accumulated multi-hazard risks and the development of a complementary geospatial dataset."

Section 4 Remote sensing data integration for ARI hotspots:

"The ARI results presented in Section 3.2 identify high-risk latitude-longitude grids where recent major disaster events have accumulated and where local resilience may have been significantly eroded. These grids provide spatial focal points for detailed risk assessment and management at finer scales. At such scales, remote sensing technology is widely recognized as a key tool for mapping exposed assets and post-disaster damage.

However, high-resolution remote sensing data, particularly sub-meter visible spectral imagery, are still predominantly obtained through commercial purchases and remain costly for many institutions, especially in developing countries. Existing open-source datasets tend to focus on large cities or specific crisis events, often in developed regions, rather than on systematically identified global risk hotspots. In addition, open-source data from different providers vary in image type, size, and other attributes, complicating direct integration and application. Overall, the lack of specific, consistent and readily accessible high-resolution imagery in high-risk but data-scarce regions constitutes a major bottleneck for advancing disaster risk assessment and management.

In this context, the role of remote sensing in the present study is complementary rather than strictly validating: ARI is used to highlight high-risk grids, and remote sensing data are then integrated and supplemented for these grids to create an open-source dataset that can support future damage detection, risk monitoring and model development. This approach establishes a bridge between macro-scale statistical risk assessment and micro-scale Earth observation applications."

Section 4.1 Data sources and spatial coverage:

"Based on the ARI-derived high-risk grids (Section 3.2), an extensive review of existing open-source remote sensing datasets was conducted to identify sub-metre visible spectral imagery covering these grids."

Section 4.3 Standardization and construction of the 3H Dataset:

“By focusing explicitly on ARI-identified hotspots in developing countries and providing standardized imagery, the 3H Dataset is expected to help mitigate several key challenges faced by these regions, including the scarcity of high-quality data, the high cost of commercial imagery and limited access to consistent data sources. In turn, it is intended to support future research and to encourage wider application of remote sensing technologies in multi-disaster risk assessment and management.”

3) Strengthened conceptual integration in the *Section 5 Discussion*, highlighting that the ARI-remote sensing coupling forms a workflow from risk identification → hotspot localization → observation-based assessment:

“Despite these caveats, the combined ARI-remote sensing framework offers practical value. For policy and humanitarian actors, ARI can help flag areas where repeated high-impact events may have eroded resilience and where more detailed assessments and targeted support are warranted. For the scientific community, the 3H Dataset provides standardized, sub-metre imagery for ARI hotspots in developing countries, lowering the barrier for remote sensing-based studies on damage detection, exposure mapping and recovery monitoring.”

Taken together, these revisions ensure that the role of remote sensing is now clearly positioned as an operational extension and potential validation pathway of the ARI framework, rather than a disconnected data contribution. ARI identifies where risk accumulates, while the 3H Dataset provides how Earth observation can support detailed evaluation in those locations, particularly in data-scarce developing regions. We believe this restructuring significantly improves conceptual coherence and addresses the referee’s concern.

Minor comments:

1. Line 107: “percentages reflect approximate contributions to global damage and losses from 2013 to 2023” → add a reference.

We appreciate the referee’s careful reading and valuable suggestion. We would like to clarify that the percentages reported in Line 107 are not directly derived from a published reference, but are instead calculated by the authors based on the subset of major disaster events identified in EM-DAT for the period 2013-2023. To avoid any ambiguity, we have revised the sentence in the manuscript to explicitly state that these percentages are approximate values derived from our own calculations using EM-DAT data.

The revised sentence now reads:

“Within the 2013-2023 study period, EM-DAT records 3217 target disaster events across four disaster types and fifteen disaster subtypes (Table 2). Applying the ≥ 50 fatalities criterion results in 344 major disaster events that form the empirical basis of this study. Among these, earthquakes, storms, floods and mass movements account for 9.9%, 22.7%, 58.1% and 9.3% of events, respectively. As estimated from the selected major events, these percentages provide an approximate indication of the relative contribution of different disaster types to recent high-fatality impacts at the global scale, indicating that floods dominate recent high-fatality disasters, followed by storms, while earthquakes and mass movements contribute smaller but non-negligible shares.” (*Section 3.1 Spatiotemporal characteristics of recent disaster events*)

We believe this clarification makes the source and nature of these values more transparent.

2. Line 115: provide a table with examples of how textual descriptions of event locations were translated into latitude/longitude coordinates. This is crucial for validating spatial accuracy, especially when events span multiple grid cells. Clarify how multi-cell events were handled.

We sincerely thank the referee for this helpful suggestion. We agree that illustrating the procedure by which textual descriptions of disaster locations were translated into latitude-longitude coordinates will improve the transparency and

reproducibility of our approach. In the revised manuscript, we have added a supplementary table (*Table 3*) that provides several representative examples. This table shows the original EM-DAT textual description, the selected representative location (e.g., city or administrative unit), and the corresponding latitude-longitude coordinates derived from administrative boundary data.

Regarding events that span multiple grid cells, we clarify that our procedure assigns each event to a single representative coordinate. In selecting this coordinate, we prioritize locations that lie near the center of the reported affected area and where the most severe impacts were observed. While this introduces some approximation, it ensures a consistent and standardized allocation across all events, which is essential for global-scale analysis. We explicitly note this limitation in the revised manuscript, and we discuss its implications in the new discussion section.

Revised content now appears as follows:

Section 2.2 Spatial standardization:

“For events spanning multiple administrative units, one representative coordinate is selected to avoid double-counting across neighbouring grid cells. This approach maintains consistency across the global dataset while acknowledging inherent spatial uncertainty, which is further discussed in Section 5, and forms the basis for the grid-based aggregation used in the ARI. To enhance transparency and reproducibility, several representative examples of this conversion are provided in Table 3, which documents the original location descriptions, selected representative sites, and corresponding coordinates.”

Table 3: Example of translating textual descriptions of event locations into latitude-longitude coordinates.

Event date	Disaster type	Disaster subtype	Country	Location (textual description & geographic center)	Latitude (°)	Longitude (°)
2021-8	Earthquake	Ground movement	Haiti	Nippes Department (18.40°N, 73.00°W)	18.41*	-73.48*
				Grand'Anse Department (18.55°N, 74.00°W)		
				Sud Department (18.35°N, 73.55°W)		
2019-9	Storm	Tropical cyclone	Bahamas	Great Abaco Island (26.39°N, 77.09°W)	26.47	-77.18
				Grand Bahama Island (26.66°N, 78.32°W)		
2020-8	Flood	Flood (General)	Pakistan	Sindh Province (25.50°N, 69.00°E)	26.35	68.85
				Khyber Pakhtunkhwa Province (34.95°N, 72.33°E)		
				Balochistan Province (28.00°N, 67.00°E)		
2014-8	Mass movement	Landslide	Nepal	Jure Village (27.77°N, 85.87°E)	27.77	85.84
				Mankha Village (27.76°N, 85.83°E)		

*: Precise latitude-longitude coordinates are available for all earthquake events in EM-DAT.

Section 5 Discussion:

“2) Spatial simplification: The assignment of a single representative coordinate per event inevitably simplifies complex hazard footprints, particularly for floods and storms with broad spatial extents.”

“Both EM-DAT and WRI provide harmonised global datasets, yet gaps and inconsistencies remain unavoidable. The standardization of textual location descriptions into point coordinates reduces heterogeneity but cannot eliminate spatial uncertainty in extended-impact events.”

3. Equation 1: the risk formulation omits the exposure component, which limits its alignment with established risk theory (as WRI dose). This should be acknowledged and discussed.

We thank the referee for this constructive comment. We acknowledge that *Equation 1* in the manuscript, expressed as “Risk = Hazard × Vulnerability”, does not explicitly include the component of exposure. As the referee points out, this omission may appear inconsistent with established formulations of disaster risk, such as the WRI framework.

In fact, as we clarified in the surrounding text, the equation was intended only as a simplified schematic representation to introduce the conceptual basis of disaster risk assessment. We emphasize that *Equation 1* is not intended to serve as the

formal risk model of the ARI, but only as a conceptual entry point to situate the proposed index within established risk theory. We also noted that this basic formulation has evolved in different indices and frameworks, such as the DRI, which incorporates hazard, population, and vulnerability, and the WRI, which defines risk as the geometric mean of exposure and vulnerability. The purpose of *Equation 1* was therefore illustrative rather than prescriptive.

At the same time, we agree with the referee that exposure is indispensable in capturing the realities of disaster risk, particularly when hazard intensity and vulnerability vary spatially but exposure levels determine the magnitude of consequences. To address this concern, we have revised the manuscript to clarify that although *Equation 1* does not explicitly contain exposure, the operational framework of the ARI does incorporate it through the use of WRI as a weighting factor. Since the WRI integrates both exposure and vulnerability dimensions, this ensures that ARI captures the combined influence of hazard occurrence and the socio-spatial context of affected regions.

In the revised manuscript, we now explicitly acknowledge in the related section that *Equation 1* is a simplified schematic and not a full risk equation, and add a brief discussion noting that exposure is indirectly embedded in the ARI framework via the WRI component, while also recognizing that this indirect representation may limit interpretability compared to approaches where exposure is modeled explicitly.

Revised content now appears as follows:

Section 2.3 Risk indices:

“Equation (1) is a simplified schematic representation that highlights the joint influence of physical processes and socio-economic fragility. This basic formula has evolved into various forms to accommodate different dimensional requirements and methodological approaches. In many operational frameworks, a third component (Exposure) is explicitly included to reflect the population and assets at risk. For example, DRI expands the model to define risk as the product of hazard, population, and vulnerability (UNDP, 2004), while WRI characterizes the risk index as the geometric mean of exposure and vulnerability (Frege et al., 2023).”

Section 2.3.2 Accumulated Risk Index (proposed index):

“In this study, WRI is not recomputed or modified. Instead, it is adopted as a country-level reference index and contextual weighting factor that encapsulates exposure and vulnerability conditions. The proposed ARI builds on WRI by combining it with the recent occurrence of major disaster events at the grid scale, using WRI as a scaling factor rather than as a quantity to be spatially aggregated on its own.”

“Building on the reference index introduced above, ARI is formulated as a grid-based measure that integrates recent major disaster events with country-level contextual information. Conceptually, ARI integrates two components: 1) the frequency of major disaster events (defined by the filtering strategy in Section 2.1) occurring within each latitude-longitude grid cell over the period 2013-2023; and 2) the country-level WRI value associated with each event, which adjusts for differences in exposure and vulnerability among affected regions. In this way, ARI emphasizes the short-term accumulation of high-impact shocks in contexts with varying underlying risk conditions.”

4. Chapter 3 (supplement of remote sensing data): clarify its role in the study. does it serve as validation, contextual information, or a complementary dataset? Currently, its integration is not explained.

We thank the referee for raising this important point. We agree that the role of the “supplement of remote sensing data” in Chapter 3 was not sufficiently explained in the original version. The remote sensing component serves not as ARI validation, but as a contextual and complementary dataset that supports future micro-scale analyses.

To clarify, the remote sensing component of our study consists of two complementary parts, as reflected in the original chapter title “Integration and supplement of remote sensing data”:

1) Integration of remote sensing data: We conducted an extensive review of existing open-source datasets, which provide

valuable sub-meter satellite images, but their coverage is uneven. They tend to prioritize major cities or certain regions, while many high-risk grids identified by ARI, particularly in developing countries, remain underrepresented. Moreover, available datasets differ in resolution, image type, and preprocessing standards, which complicates their direct application.

2) Supplement of remote sensing data: To address these gaps, we supplemented high-resolution visible spectral imagery specifically for two high-risk grids, where no relevant open-source data was available. For each of these grids, we acquired and standardized sub-meter imagery covering a 50 km² area (100 km² in total), centered on the coordinates of the disaster event with the maximum “Total Deaths”. The supplemented imagery was standardized and then integrated with the existing open-source datasets. This supplementation ensures that the most data-scarce yet high-risk grids are adequately represented.

Importantly, this remote sensing dataset, compiled as the 3H Dataset, is not intended to directly validate the ARI in this preliminary stage. Rather, it complements the ARI framework by providing a foundation for future methodological integration. Specifically, the ARI identifies spatial clusters of accumulated disaster risk, while the 3H Dataset provides accessible remote sensing resources for those high-risk grids. Together, they establish a bridge between macro-level risk assessment and micro-level hazard detection, supporting subsequent and independent validation efforts using Earth observation data, as well as applications such as post-disaster recovery monitoring and the training of machine learning models for damage detection.

In the revised manuscript, we have expanded the related section (*Section 4 Remote sensing data integration for ARI hotspots*) to emphasize this complementary role and to clarify how integration and supplement jointly contribute to building a usable dataset that supports the operationalization of ARI findings..

Revised content now appears as follows:

Section 1.2 Motivation and Objectives:

“In addition, in order to facilitate further remote sensing-based research, this study systematically integrates and supplements existing sub-meter visible spectrum remote sensing imagery for ARI-identified high-risk grids, and publishes them as an open-source dataset after standardization. This dataset provides essential support for future studies at finer spatial scales.”

Section 4 Remote sensing data integration for ARI hotspots:

“In this context, the role of remote sensing in the present study is complementary rather than strictly validating: ARI is used to highlight high-risk grids, and remote sensing data are then integrated and supplemented for these grids to create an open-source dataset that can support future damage detection, risk monitoring and model development. This approach establishes a bridge between macro-scale statistical risk assessment and micro-scale Earth observation applications.”

Section 4.2 Supplementing data gaps in high-risk grids:

“Despite the relatively broad coverage provided by RAMP and DigitalGlobe, two high-risk grids (Grids 3 and 7 in Table 5) lack suitable open-source imagery. To address these gaps, additional visible spectral satellite images were acquired and processed specifically for these grids. For each of the two grids, imagery covering a rectangular area of approximately 50 km² was obtained, resulting in 100 km² of supplementary data in total. The centre of each acquisition area was chosen to coincide with the geographic location of the major disaster event with the maximum “Total Deaths” in that grid, thereby ensuring that the supplemented imagery is directly relevant to the most severe recorded impacts. The supplemented images were selected to match, as closely as possible, the spatial resolution and spectral characteristics of the existing open-source datasets, so that they could be seamlessly integrated into a unified dataset.”

Section 4.3 Standardization and construction of the 3H Dataset:

“By focusing explicitly on ARI-identified hotspots in developing countries and providing standardized imagery, the 3H Dataset is expected to help mitigate several key challenges faced by these regions, including the scarcity of high-quality data, the high cost of commercial imagery and limited access to consistent data sources. In turn, it is intended to support future research and to encourage wider application of remote sensing technologies in multi-disaster risk assessment and

management.”

Once again, we sincerely appreciate the referees’ constructive comments and the editor’s careful handling of our submission. We have carefully addressed every point raised, undertaking substantial structural revisions, clarifying methodological assumptions, expanding analytical depth, and improving the conceptual integration between the ARI framework and the remote sensing dataset. These revisions, in our view, have significantly strengthened the scientific rigor, transparency, and overall coherence of the manuscript.

At the same time, we fully acknowledge that some of the methodological choices adopted in this study may limit the suitability of the proposed ARI for certain applications, particularly those requiring long-term or fully multidimensional risk representations. We therefore welcome continued critical discussion and future refinement of the framework, both by ourselves and by the broader research community.

We are grateful for the opportunity to improve our work and genuinely hope that the revised version now meets the standards of **Natural Hazards and Earth System Sciences**. We would be pleased to further revise the manuscript should any additional clarification or adjustment be required.

Kind Regards,

Qingzhao Kong and Erqi Zhu

15th Dec. 2025

References

- Ceferino, L., Merino, Y., Pizarro, S., et al.: Placing engineering in the earthquake response and the survival chain, *Nature Communications*, 15, 4298, <https://doi.org/10.1038/s41467-024-48624-3>, 2024.
- Delforge, D., Wathelet, V., Below, R., et al.: EM-DAT: the emergency events database. *International Journal of Disaster Risk Reduction*, 105509, <https://doi.org/10.1016/j.ijdr.2025.105509>, 2025.
- Frege, I. A., Radtke, K., Kienzl, P., et al.: *World Risk Report 2023 Focus: Diversity, Bündnis Entwicklung Hilft / IFHV*, ISBN 978-3-946785-16-3, 2023.
- Hussain, M. A., Shuai, Z., Moawwez, M. A., et al.: A Review of Spatial Variations of Multiple Natural Hazards and Risk Management Strategies in Pakistan, *Water*, 15(3), 407, <https://doi.org/10.3390/w15030407>, 2023.
- Marin, G., Modica, M., Paleari, S., et al.: Assessing disaster risk by integrating natural and socio-economic dimensions: A decision-support tool, *Socio-Economic Planning Sciences*, 77, <https://doi.org/10.1016/j.seps.2021.101032>, 2021.
- Newman, R., and Noy, I.: The global costs of extreme weather that are attributable to climate change, *Nature Communications*, 14, 6103, <https://doi.org/10.1038/s41467-023-41888-1>, 2023.
- Peduzzi, P., Dao, H., Herold, C., et al.: Assessing global exposure and vulnerability towards natural hazards: the Disaster Risk Index, *Natural hazards and earth system sciences*, 9, 1149-1159, <https://doi.org/10.5194/nhess-9-1149-2009>, 2009.
- Wen, J., Wan, C., Ye, Q., et al.: Disaster Risk Reduction, Climate Change Adaptation and Their Linkages with Sustainable Development over the Past 30 Years: A Review, *International Journal of Disaster Risk Science*, 14, 1-13, <https://doi.org/10.1007/s13753-023-00472-3>, 2023.
- Yarveysi, F., Alipour, A., Moftakhari, H., et al.: Block-level vulnerability assessment reveals disproportionate impacts of natural hazards across the conterminous United States, *Nature Communications*, 14, 4222, <https://doi.org/10.1038/s41467-023-39853-z>, 2023.