

Dear Editorial Board,

First of all, we would like to thank the constructive feedback provided by the two reviewers and the opportunity given to submit a new revised manuscript.

We have now uploaded the revised manuscript, the track changes document, the supplementary material and this Author's response letter. This document aims to give a point-by-point answer to the reviewers' comments after incorporating all the requested changes into the manuscript. Our answers are written in blue.

To the editors, please note that the title, the order of the authors and the affiliations have been slightly modified.

We hope that the revisions meet the expected standards and look forward to receiving the reviews of the newly uploaded manuscript.

Please do not hesitate to contact us for any further information or queries regarding the manuscript.

We look forward to hearing from you soon. Thank you very much in advance.

Sincerely,

Iris Schneider Pérez, on behalf of the team

Reviewer #1

This paper presents the development of a historical database of six natural hazard events (i.e., Earthquakes, Floods, Ground subsidence, Landslides, Rock falls and Wildfires) in the Garrotxa region of Catalonia, Spain, covering the period 1900 - 2023. The authors aim to address challenges of data fragmentation, inconsistency, and lack of accessibility by compiling existing records into a centralised and standardised database for the case study.

The idea of developing a multi-hazard dataset is timely, especially in the context of the growing global concern that hazards no longer occur in isolation but often interact, creating more complex risk scenarios. Understanding such interactions is essential for improving forecasting, preparedness, and risk reduction.

However:

While the paper claims to present a multi-hazard database, the methodology currently treats each hazard type separately, with little evidence of integration or analysis of their potential interactions. For the database to be considered truly “multi-hazard,” the authors should explain whether and how different hazard datasets are compared, overlaid, or analysed together - for example, through spatial correlation, cascading event identification, shared exposure layers, or temporal overlaps. This clarification would significantly strengthen the paper’s contribution.

This issue also ties into the lack of a clear definition of “multi-hazard.” The manuscript uses terms like “multi-hazard” and “multi-risk” interchangeably, and at times reverts to simply “natural hazards.” It would be helpful to clearly define the scope of the database early on: Are these hazards considered individually or in terms of interrelationships? If they are treated separately, as seems to be the case, the term “multi-hazard” may need to be refined to reflect that - perhaps described instead as a “multiple single-hazard” dataset - and more justification should be given regarding how this contributes to future integrated risk analysis.

The dataset has been redefined as a “single-layer multi-hazard dataset” (Line 77) or simply as a “natural hazard dataset” (title) because the natural hazard events were compiled separately. Cascading events were not specifically captured (Lines 404–405), although notes from citizen science workshops indicate evidence of interrelationships between hazards (Table 4).

The manuscript explains that the dataset is designed as a first step toward multi-hazard risk assessment. This aim has guided its construction at every stage: defining the temporal framework (Line 184); selecting the natural hazards (Line 196); reviewing variables that may influence the triggering and development of hazard processes (Line 201, Table 1, Table S1); including meteorological variables to enable future multi-hazard analysis using graph theory (Table 2); and outlining specific future steps for multi-hazard risk assessment based on the data collected in this study (Section 5.1).

The manuscript has been revised to ensure consistent and accurate use of key terms. The term *multi-hazard* is defined in line 28 following UNDRR (2016) and Gill et al. (2022). The term *multi-hazard risk* is used when referring to the potential impacts of natural hazards, taking into account its dynamic nature as a function of hazard, exposure, vulnerability, and mitigation strategies (lines 32–33). The terms *natural hazard* and *single-layer multi-hazard* are used when hazards are considered individually.

In the following, I outline the major points of this research:

1) Title

1. The title feels somewhat broad and could benefit from more specificity. Terms like "methodological approach" and "multi-hazard analysis" are quite generic. It would be helpful to clarify the type of methodology used and the paper's main contribution or novelty.

The title has been changed to *Development of a Natural Hazard Dataset: A Foundational Step Toward Multi-Hazard Risk Assessment. Application to the Garrotxa Region (Catalonia, Spain)*. This emphasizes that the main result is a natural hazard dataset (single-layer approach), which can serve as a basis for future data-driven multi-hazard risk assessments. It also specifies the study region.

2) Abstract

1. As noted in the general comments, the abstract claims to introduce a methodology for the creation of a multi-hazard database, but the terminology shifts inconsistently throughout, alternating between *multi-risk*, *natural hazard database*, and *multi-hazard*. This inconsistency should be addressed. *This issue has been addressed, and the terms are now used properly.*
2. The abstract would benefit from a clearer structure. *The structure has been modified according to the points given. Below, we mention in which line the information can be found.* It should briefly:
 - Explain the importance of multi-hazard assessments, *(Lines 11-12)*
 - Summarise the data sources and methodology, *(Lines 13-17)*
 - Present the key findings, and *(Lines 17-18)*
 - Emphasise the broader significance of the work. *(Lines 18-20 and 22-23)*

Key limitations and challenges have also been outlined in lines 20–22.

3) Introduction

The contribution (state of the art) is not clear. Since the paper is about multi-hazard data collection, it is expected that the introduction discuss more about the importance of data collecting and the relevant existing literature both locally/nationally and globally. According to the paper, *it is the first attempt of collecting this data in the region*, but the readers also would like to know other attempts either nationally or globally. The state of the art needs to be strongly supported by literature. Therefore, this manuscript lacks the scientific backup. The introduction outlines an important initiative, but several key claims would benefit from further clarification or supporting evidence. For example, the claim that this is the first multi-risk database for the Garrotxa region should be supported by a brief overview of existing datasets and what sets this work apart.

The state of the art (lines 25–62) has been rewritten and now follows this structure: (1) the importance of an overall multi-hazard risk perspective; (2) the definition of multi-hazard; (3) the main characteristics of multi-hazard risk; (4) scientific research on methodologies for multi-hazard and multi-hazard risk assessment; and (5) the role of databases in supporting these

methodologies. Examples are provided, and the main challenges in database construction are discussed. Relevant scientific literature is cited to support these points.

The main contribution of this study is that the dataset unifies information previously scattered across various sources (line 78). This is further detailed in the methodology (Section 3.2), where the sources reviewed are described. In this sense, it represents the first single-layer multi-hazard dataset of the region, which standardizes the data and facilitates future data-driven multi-hazard risk assessments.

Additionally, terms like “intrinsic, extrinsic, and evolutionary variables” are not defined and may be unfamiliar to readers.

The terms *intrinsic*, *extrinsic*, and *evolutionary variables* are defined in the methodology section (lines 217–221). The terminology has been slightly adjusted to improve clarity: *not influential; internal or constitutive; external or environmental; and process development variables*.

A new section titled *3.1 Data identification* has been added. It provides further information on the theoretical approach used to identify which variables influence each natural hazard considered. The search methodology is explained and summarized in Table 1, while the resulting variables are presented in Table S1 (supplement).

Finally, while the broader benefits of the database are valuable, they could be more directly linked to specific features or applications.

The broader benefits of the dataset are described in relation to the characteristics of the region (line 67) and its role as a regional risk management tool within the GarMultiRisk project (line 73).

4) The entire section 3 “Natural hazards in the Garrotxa region”, lacks citations. Given the descriptive detail, it is important to support this information with references to scientific studies, local hazard reports, or official datasets. This will strengthen the credibility of the section and allow readers to verify the information presented.

Citations have been added, and the corresponding references have been included at the end of the manuscript.

5) Methodology

1) For a paper centered on a data product, it lacks detail on how the data were processed, structured, validated, and made usable. The structure of the database is not deeply described or visualised. The methodology would benefit from a more detailed explanation of how heterogeneous datasets were standardised and cleaned. For example, how were missing values addressed, or duplicate/conflicting reports reconciled?

Several tasks have been carried out to address these concerns:

- The methodology has been restructured into three sections: *3.1 Data identification*, *3.2 Data collection strategy*, and *3.3 Data processing*. *Data identification* describes the criteria used to define the time span, select the natural hazards, and identify the main variables influencing each hazard. *Data collection strategy* details the sources reviewed (inventoried, written, and oral). *Data processing* explains, for each hazard and source, how the data were incorporated into the final dataset. Tables have been added to summarize the information and improve readability.

- The final dataset has been uploaded to an official repository and assigned a DOI. It is publicly available at <https://digital.csic.es/handle/10261/394924> (Lagresa et al., 2025). The *Data availability* section has been updated accordingly, and the link is also provided in the *Results* section (line 297).
- Table S2 (supplement) has been created. It provides a detailed definition of the codes, descriptions, units (where applicable), and examples for each variable collected across the natural hazards included in the dataset.

2) The mention of citizen science workshops is a valuable component of the methodology. However, the description is too brief to evaluate the scientific validity of this approach. I recommend providing more detail on the workshop structure, the nature of the participation, the type of data collected, and how this information was validated and integrated into the database. As this is a peer-reviewed paper, reproducibility and transparency are essential.

More information has been provided on this behalf in Lines 234-237, Table 4 and Figures 2 and 3. Section 3.3 explains how information collected from oral sources were used during the data processing step (e.g. line 260).

I would like to emphasise that for a paper centered on developing a database, there is surprisingly little about: The data model or structure. There is lack of clarity on data accessibility. The paper mentions several data sources - institutional databases, press archives, citizen science input - but it does not explain how accessible these are. Are the datasets open-access? Can others replicate this approach in another region? Were any data restricted or available only by request?

This has been carefully addressed in Tables 2, 3, and 4. Replicability is specifically discussed in Section 5.3.

To improve transparency and reproducibility, I recommend expanding on the workflow used to compile and validate the data, including how the data was structured, what tools or software were used, and which sources are publicly accessible. Additionally, including a schematic or diagram of the database architecture would make it easier for other researchers to apply this approach to other regions.

As mentioned above, the information has been summarized in tables, and the dataset is now publicly available. The software used for dataset management was Excel (line 298), and the final structure is briefly described in lines 297–303.

3) The manuscript references several institutions and data sources (e.g., ICGC, AEMET, METEOCAT) without providing background or context for the reader. For an international audience unfamiliar with Catalan or Spanish institutions, it would be helpful to briefly introduce each one at first mention - including its role, type (e.g., national agency, academic unit), and why it was relevant for the data used. This would improve clarity and accessibility of the paper.

Descriptions of each institution used for the inventoried data collection are provided in Table 2. Table 3 also gives details on the two digitized newspapers.

4) The manuscript states that ground subsidence data were obtained through direct contact with researchers who had previously studied the phenomenon. While collaborative data sharing can

enrich a study, from a scientific perspective, it raises concerns about the validity, traceability, and reproducibility of the database. For future improvement, it would be important to clearly describe how such data were validated, whether they are documented in peer-reviewed sources, and if they are accessible to other researchers. Ensuring that all data sources are verifiable and citable is essential to strengthen the credibility and scientific rigour of the database.

The main data on ground subsidences were obtained from previously published sources, in particular the PhD thesis of Fábregat, I. (2020). A personal request was made to the author to obtain the Excel table. This is outlined in Table 3.

6) Discussion and challenges

There is a clear structural issue between the “Discussion” and “Challenges and Future Steps” sections. Many of the key challenges are currently placed in the Discussion, while the dedicated “Challenges” section focuses almost entirely on future plans. This creates confusion for the reader. The Discussion does not fully engage in critical interpretation of the findings. A stronger discussion would:

- Reflect on the implications of the database for multi-hazard risk analysis,
- Compare the work to similar efforts in other regions,
- And consider its potential value for scientific and policy communities.

The discussion section has been fully restructured and rewritten. It is now organized into five subsections, each addressing a key point:

- Section 5.1 *Multi-hazard risk assessment framework* outlines future steps for conducting a multi-hazard risk assessment based on the dataset.
- Section 5.2 *Implications for regional risk management strategies* explains how the study contributes to the development of regional strategies through the GarMultiRisk Project and the resulting Natural Risk Reduction and Management Plan. Even at the dataset stage, relevant insights are provided.
- Section 5.3 *Scalability* discusses how the dataset could be replicated in other regions.
- Section 5.4 identifies the main challenges encountered during data collection and processing.
- Section 5.5 presents the main limitations of the study.

As much of the current content is descriptive, the paper would benefit from a more analytical tone and a clearer articulation of its broader relevance.

The content of the paper has been thoroughly revised to adopt a more analytical tone and a clearer structure. Lines 440–444 of the conclusion highlight the broader relevance of the study.

Minor comments

I have not included minor comments, as the issues outlined above affect the overall framing, methodology, and scientific positioning of the paper. In my view, substantial restructuring and clarification would be needed for the manuscript to meet publication standards.

Reviewer #2

The manuscript “Methodological approach to multi-hazard analysis: the case of the Garrotxa region (Catalonia, Spain)” introduces a multi-hazard events database for the Garrotxa region. This contribution is certainly valuable in itself and supports the advancement of comprehensive risk assessments and improved preparedness and response strategies, aligning with the objectives of international initiatives such as the Sendai Framework.

Nevertheless, **the presentation quality of the work requires significant improvement before publication.** The three main issues are: (i) the lack of adequate references to support the reported information and findings, and (ii) the lack of a detailed explanation of the database design procedure and the database structure, (iii) the lack of a critical discussion on the applicability of the database to analyse hazard interactions and advance multi-hazard disaster risk management in general.

Please find my detailed comments below:

The **title** does not reflect the content of the paper. Why is the keyword “database” not mentioned explicitly? I suggest providing a more suitable title for the work. *The title has been changed to **Development of a Natural Hazard Dataset: A Foundational Step Toward Multi-Hazard Risk Assessment. Application to the Garrotxa Region (Catalonia, Spain)**. This emphasizes that the main result is a natural hazard dataset (single-layer approach), which can serve as a basis for future data-driven multi-hazard risk assessments. It also specifies the study region.*

The abstract does not follow a logical structure, and it is quite difficult to immediately capture the original contribution of the work to the state-of-the-art. I invite the authors to review the abstract to *The abstract has been rewritten and it now addresses the following comments (specific lines in parentheses) : i) state the problem or gap to address (lines 11-13); ii) clearly what the research seeks to address (lines 18-19) ; iii) summarize the methodology, data, and analysis techniques used (lines 15-18); iv) highlight the key results or outcomes of the study (lines 13-15); v) summarize the significance of the findings and their broader implications (lines 18-19 and lines 22-23).*

At the end of the abstract (lines 23-24), the possibility of confronting “similar potential hazard scenarios” is mentioned. This statement is not clear. What does it mean? Moreover, what do you mean exactly by the expression “hazard scenarios”? This expression is not mentioned at all throughout the whole manuscript, except for two brief mentions in chapter 7.

*These lines have been modified. The term **hazard scenario** is now used in lines 12, 28, and 338. The term **scenario** refers to a specific development of a natural hazard event or its impacts.*

The expressions “multi-hazard” and “multirisk” (or “multi-risk”) are used interchangeably in the manuscript. Nevertheless, they refer to two different concepts. See e.g. the classification introduced by Zschau. (2017) or the definitions reported in the Disaster Risk Gateway - <https://disasterriskgateway.net/index.php/Definitions>). I invite the authors to check their use carefully. I think that in the majority of cases in the manuscript, the term “multi-hazard” would be more appropriate.

*The manuscript has been revised to ensure consistent and accurate use of key terms. The term **multi-hazard** is defined in line 28 following UNDRR (2016) and Gill and Malamud (2014). The term **multi-hazard risk** is used when referring to the potential impacts of natural hazards, taking into account its dynamic nature as a function of hazard, exposure, vulnerability, and mitigation*

strategies (lines 32–33). The terms *natural hazard* and *single-layer multi-hazard* are used when hazards are considered individually.

Introduction, line 35: “In this context, scientific knowledge about the hazards that affect a region, both temporally and spatially, is crucial.” This sentence lacks references. There has been a significant amount of literature in the last fifteen years that has addressed the role of spatial and temporal overlapping of hazards. See e.g. Gill and Malamud (2014) or De Angeli et al. (2022), among others. Similarly, the sentence reported at lines 37-39 and 44-46 of the Introduction would benefit from the inclusion of additional references to those of López-Saavedra and Martí (2023).

Citations have been added, and the corresponding references have been included at the end of the manuscript.

Introduction, lines 64-65: What does “intrinsic, extrinsic, and evolutionary variables” mean? These terms are not subsequently mentioned in the section (4.3) where the variables are supposed to be described.

The terms *intrinsic*, *extrinsic*, and *evolutionary variables* are defined in the methodology section (lines 217–221). The terminology has been slightly adjusted to improve clarity: *not influential*; *internal or constitutive*; *external or environmental*; and *process development variables*.

A new section titled *Data identification* has been added. It provides further information on the theoretical approach used to identify which variables influence each natural hazard considered. The search methodology is explained and summarized in Table 1, while the resulting variables are presented in Table S1 (supplement).

Section 3 (Natural hazards in the Garrotxa region) completely lacks references. I invite the authors to provide the sources of this information, even in the form of grey literature (e.g., technical reports, official websites) when peer-reviewed literature is not available.

Citations have been added, and the corresponding references have been included at the end of the manuscript.

Section 4 (Methodology), lines 175-181: What has been the criterion for the selection of these hazards? This selection needs to be explained and supported better.

This is now explained in Lines 186-200.

Section 4 (Methodology), lines 185-188: Which is the source of this information?

As the methodology has been rewritten, references are now provided for earthquakes, landslides, and rockfalls in Section 2.2, and specifically for ground subsidences in line 166.

Section 4 (Methodology), lines 254-257: Which stakeholders have been involved? How have the workshops been organised? I invite the authors to provide more details.

Table 4, Lines 234-237 and Fig. 2 and 3 provide information about how the citizen science workshops were performed.

Section 4.3 (Data structuring and validation) is too poor. Please describe each of the key data fields and explain how they have been selected, e.g., did you take inspiration from existing databases? Moreover, please provide a more detailed description of the DB architecture. As a generic comment, Section 4 (Methodology) would benefit from the introduction of a scheme

that depicts the database structure and how the different sources of information have been integrated.

The methodology has been restructured into three sections: *3.1 Data identification*, *3.2 Data collection strategy*, and *3.3 Data processing*. *Data identification* describes the criteria used to define the time span, select the natural hazards, and identify the main variables influencing each hazard. *Data collection strategy* details the sources reviewed (inventoried, written, and oral). *Data processing* explains, for each hazard and source, how the data were incorporated into the final dataset. Tables have been added to summarize the information and improve readability.

The final dataset has been uploaded to an official repository and assigned a DOI. It is publicly available at <https://digital.csic.es/handle/10261/394924> (Lagresa et al., 2025). The *Data availability* section has been updated accordingly, and the link is also provided in the *Results* section (line 297).

The structure of the final dataset is briefly discussed in lines 298-303.

Table S2 (supplement) has been created. It provides a detailed definition of the codes, descriptions, units (where applicable), and examples for each variable collected across the natural hazards included in the dataset.

The discussion in Section 6 is almost uniquely a list of the challenges encountered in the creation and filling of the database. Why are these not instead reported in section 7, currently named Challenges and Future steps? Moreover, the discussion phase does not discuss in detail how the database can support a better understanding of multi-hazard dynamics and how it can be used to support the advancement of comprehensive risk assessments and improved preparedness and response strategies

The discussion section has been fully restructured and rewritten. It is now organized into five subsections, each addressing a key point:

- Section 5.1 *Multi-hazard risk assessment framework* outlines future steps for conducting a multi-hazard risk assessment based on the dataset.
- Section 5.2 *Implications for regional risk management strategies* explains how the study contributes to the development of regional strategies through the GarMultiRisk Project and the resulting Natural Risk Reduction and Management Plan. Even at the dataset stage, relevant insights are provided.
- Section 5.3 *Scalability* discusses how the dataset could be replicated in other regions.
- Section 5.4 identifies the main challenges encountered during data collection and processing.
- Section 5.5 presents the main limitations of the study.

Related to the previous point, the authors mention several times in the manuscript that the database would enable “comprehensive analyses of [hazard] interactions,” but they never explain how these analyses would be performed, and which specific features of the database would facilitate these analyses.

The methodological approach toward a future multi-hazard risk assessment is now explained in Section 5.1. In Section 3.1, we provide details on how the different methodological steps were carried out with the future analysis in mind.