The paper investigated the spatial coherence between historical (1600-1899) and instrumental (1900-1950) catalogs, trying to answer three central questions (as out by the author):
i. At which spatial scales does historical seismicity exhibit statistically meaningful agreement with instrumental seismicity?
ii. How does this coherence depend on tectonic setting (e.g. subduction zones, continental collision belts, intraplate regions)?
iii. What are the implications of these results for the use of global historical earthquake catalogs in seismic hazard and tectonic studies?
To do so the author discussed features of the historical and instrumental catalogs and computes spatial probability density fields via isotropic Gaussian kernel smoothing. The author focuses on the event location rather than magnitude and provides insights on the limitations of the use of historical earthquakes catalogs for use in geophysical and hazards research.
The questions posed by the authors are certainly worth of investigation, but I feel that more work is desirable before seeing this work published in NHESS. Below I outline my general comments and the detailed comments are included in pop-up notes (usually associated to highlighted text) in the attached PDF.

I sincerely thank the reviewer for the careful reading of the manuscript and for the constructive and insightful comments provided. I appreciate the reviewer's recognition of the relevance of the scientific questions addressed in this study, namely the scale dependence of historical–instrumental spatial coherence, its variation across tectonic environments, and its implications for the interpretation of historical earthquake catalogues in seismological and hazard-related applications.
The comments have been extremely helpful in improving both the methodological robustness and the presentation of the manuscript. In response to the review, I performed several additional analyses, including sensitivity tests using multiple instrumental reference periods (1900–1950, 1951–1970, and 1971–2025) and multiple instrumental magnitude thresholds ($M \geq 5.0$–6.5). These new results are now incorporated in the revised manuscript and supplementary material. I also revised several sections of the text to clarify methodological choices, strengthen the interpretation of the results, and provide a more balanced discussion of the strengths and limitations of historical earthquake catalogues.
Detailed responses to each comment are provided below.

1. The author needs to justify the use of the early instrumental catalog (1900-1950) instead of the recent one (e.g., 1970-2020): given the aims of his work, one wonders why not using a more reliable and complete instrumental catalog instead of the more uncertain and incomplete one from the first part of the last century;

I thank the reviewer for this important comment. The original use of the 1900–1950 catalogue was motivated by its temporal proximity to the historical period. However, I agree that this choice required explicit justification and that the sensitivity of the results to the instrumental reference period should be tested.
I therefore repeated the analysis using three non-overlapping instrumental reference periods: 1900–1950, 1951–1970, and 1971–2025, all with $M \geq 5.5$. The revised Fig. 6 now shows that absolute coherence values decrease from the early instrumental to the modern instrumental era, reflecting the progressively improved global coverage of instrumental seismicity, especially in remote oceanic and sparsely documented regions. Importantly, however, the scale-dependent behaviour remains stable across all three reference periods: coherence increases monotonically with smoothing scale and displays the same overall trend.
This new analysis clarifies that the absolute value of the coherence metric is benchmark-dependent, whereas the main conclusion of the study—the emergence of historical–instrumental spatial agreement only at regional to continental smoothing scales—is robust with respect to the instrumental reference period.

2. The comparison between historical and instrumental catalogs strikes me as unfair. It is clear that the European historical catalog lists many small earthquakes that would not be included in an instrumental global

catalogs with M>5.5 (addressing point 1 could benefit the aims of the author, as instrumental catalogs for recent decades would include smaller earthquakes);

We thank the reviewer for this important observation. We agree that historical catalogues, particularly in well-documented regions such as Europe, may contain numerous moderate earthquakes that would not be included in a global instrumental catalogue restricted to $M \geq 5.5$. We therefore investigated the sensitivity of our results to the magnitude threshold adopted for the instrumental catalogue.
To address this issue, I performed an additional sensitivity analysis using three non-overlapping instrumental reference periods (1900–1950, 1951–1970, and 1971–2025) and four magnitude thresholds ($M \geq 5.0$, 5.5, 6.0, and 6.5). The results are presented in the new Supplementary Fig. S1.
The analysis shows that absolute coherence values vary with the adopted magnitude threshold, particularly for the early and transitional instrumental periods, where the number of available events is more limited. However, the fundamental behaviour of the coherence metric remains unchanged. In all instrumental eras and for all magnitude thresholds, coherence increases monotonically with smoothing scale and exhibits the same overall scale-dependent pattern.
The modern instrumental catalogue (1971–2025) is especially insensitive to magnitude threshold, despite large differences in the number of earthquakes included at each cutoff. This indicates that the large-scale spatial distribution of global seismicity is relatively stable with respect to magnitude selection when catalogue completeness is high. Although stronger threshold-dependent differences are observed in the earlier instrumental periods, these differences affect primarily the absolute level of coherence rather than its scale dependence.
Because the analysis is based on normalized spatial probability distributions rather than earthquake counts, changes in magnitude threshold modify the number of events included in the comparison but do not alter the principal conclusion of the study: statistically meaningful historical–instrumental agreement emerges only after substantial spatial smoothing and remains observable across a wide range of instrumental magnitude thresholds.
I have added this new sensitivity analysis as Supplementary Fig. S1 and discuss its implications in the revised manuscript.

3. It feels from several sentences in the text that one goal of the author is to "downgrade" (somewhat) the importance of the historical earthquake catalog in the use of hazard and geophysical research. We all know that any earthquake catalog has limitations and that as we go further back in time such limitations are exacerbated by the paucity of data, however I feel that the tone of the text should just reflect that as in its current state it unnecessarily depicts the historical record too negatively.

I thank the reviewer for this important observation. I agree that the original wording of several passages may have conveyed an excessively negative view of historical earthquake catalogues. This was not the intended message of the manuscript.
The objective of the study is not to diminish the scientific value of historical seismicity data, but rather to quantify the spatial scales and conditions under which historical catalogues preserve tectonically meaningful information at the global scale. Historical earthquake catalogues remain indispensable for extending seismic observations beyond the instrumental era, documenting rare destructive earthquakes, and supporting both hazard and seismotectonic investigations.
Following the reviewer's suggestion, I carefully revised the manuscript to adopt a more balanced and nuanced tone throughout. In particular:

- Several statements emphasizing the limitations of historical catalogues were reformulated to acknowledge both their strengths and their constraints.
- Expressions such as “dominated by reporting artefacts” were replaced by more neutral formulations emphasizing the influence of heterogeneous reporting conditions rather than implying a lack of scientific value. The Abstract, Introduction, Results, and Conclusions sections were revised to clarify

that historical catalogues retain a detectable and statistically meaningful large-scale tectonic signal, whose interpretation depends on spatial scale and tectonic context.

- The Conclusions were substantially rewritten to emphasize that the study provides a quantitative framework for identifying the spatial scales at which historical catalogues can be most reliably interpreted, rather than establishing a purely negative assessment of their usefulness.

The revised manuscript now explicitly states that historical earthquake catalogues contain a measurable large-scale imprint of tectonic structure and remain highly valuable for regional studies, long-term seismic investigations, and the documentation of rare events. The main conclusion is therefore not that historical catalogues are unreliable, but that their informational content is strongly scale-dependent and should be interpreted accordingly.

I believe that these revisions address the reviewer's concern and result in a more balanced presentation of the role of historical earthquake catalogues in modern seismological research.