In the manuscript entitled ‘Scale-dependent spatial coherence between historical and instrumental earthquake catalogues at the global scale’ the author attempts to verify to what degree do global earthquake reports accord with the tectonic settings. In my opinion, the main strength of the manuscript is the decision to focus on the spatial distribution rather than on magnitude comparisons, thereby avoiding a well-known weakness of historical catalogues. This is a careful, consistent, and convincing methodological choice. I find the manuscript is well structured and interesting to read.

I thank the reviewer for the positive assessment of the manuscript and for recognizing the rationale behind the methodological framework adopted in this study. I particularly appreciate the acknowledgment of the decision to focus on spatial distributions rather than magnitude-based comparisons, as this choice was specifically motivated by the well-known heterogeneity and incompleteness of historical magnitude information at the global scale. I am also grateful for the reviewer’s comments regarding the overall structure and readability of the manuscript. The detailed suggestions provided throughout the review have helped improve both the clarity of the presentation and the interpretation of the results.

I think that in its present form, the manuscript lacks significant support and justification and could benefit from additional, deeper revision prior to publication. First and foremost, I am not sure the research questions (lines 78-83) are fully addressed. The thing that bothers me most is that the author attempts to resolve these goals solely by quantitative approaches whereas the nature of the pre-instrumental records is qualitative and should be inspected also is association with social/cultural/human context. The author is indeed aware of that (lines 192-193) but seems to ignore these important insights in his interpretation. Simply put, the fact that a given report exists is primarily subjected to human factors rather than tectonic settings. For example, large, populated regions most likely will be reported massively regardless of the type and nature of the nearby tectonic units. Such a discussion, rooted in the historical context of the data, is not present in the manuscript which limits the importance and meaningfulness of results and discredit the ability to accurately address the questions raised by the author.

I thank the reviewer for this thoughtful and important comment. I agree that historical earthquake catalogues differ fundamentally from instrumental catalogues because they are shaped not only by earthquake occurrence itself but also by the historical, social, cultural, and institutional conditions that determine whether an event is documented, transmitted, and preserved in the historical record.
The objective of this study is not to disentangle or model these societal and historical factors. Rather, the aim is to evaluate the extent to which a tectonically meaningful spatial pattern remains detectable in historical earthquake catalogues despite the cumulative effects of these non-geophysical influences. I agree that this distinction was not sufficiently emphasized in the original version of the manuscript.
To address this concern, I have revised the Introduction by adding a new paragraph explicitly stating that historical earthquake catalogues should not be regarded as purely geophysical datasets, but as the result of a complex filtering process involving population distribution, literacy, administrative organization, communication networks, historical documentation practices, and long-term archival preservation. The revised text also clarifies that the coherence metric introduced in this study should not be interpreted as a direct measure of tectonic activity alone, but rather as a measure of the extent to which tectonic structure remains detectable after these historical and societal filters have acted on the earthquake record.
In addition, I revised several parts of the manuscript to adopt a more balanced interpretation of the results and to emphasize that the study seeks to identify the spatial scales at which historical catalogues retain tectonically meaningful information, rather than to evaluate historical catalogues solely in terms of their limitations.
I believe that these revisions clarify both the scope and the limitations of the study and address the reviewer’s concern regarding the historical and societal context of pre-instrumental earthquake records.

The second point is related to the goal of the study: “with the aim of clarifying the extent to which historical seismicity reflects the underlying tectonic organization of earthquake occurrence” (lines 238-240). I find the conclusions of the manuscript regarding this aim is somewhat vague and in sufficient. For instance, to what extent do the reports reflect the underlying tectonic organization? How much? Where exactly? why? Figure 4 (c), 6 and 7 are valuable but the when the inspection is global (worldwide) one would expect a deeper evaluation of the results. While we are on that matter, I would argue that the coherence in Europe region (Figure

4(c)) demonstrates perfectly my previous point that is, the observed coherence is primarily subjected to human factors and NOT because of tectonic organization. I am not saying the author is incorrect – I only think that the conclusions are based on partial examination that lacks the historical context.

I thank the reviewer for this insightful comment. I agree that the original version of the manuscript did not sufficiently articulate the implications of the coherence values for the central objective of the study, namely assessing the extent to which historical seismicity reflects the underlying tectonic organization of earthquake occurrence.
To address this concern, I revised the interpretation of the results in several parts of the manuscript. The revised text now explicitly states that historical earthquake catalogues retain a detectable and statistically meaningful large-scale tectonic imprint, but that this imprint represents only a partial reconstruction of the spatial organization observed in instrumental seismicity. In particular, I now emphasize that even at the largest smoothing scales considered, coherence remains substantially below unity, indicating that historical catalogues capture only part of the global instrumental spatial pattern.
I also expanded the discussion of the spatial meaning of the results. The revised manuscript now clarifies that the strongest historical–instrumental agreement is associated with major tectonically active regions, particularly subduction systems, whereas lower coherence is observed in tectonic environments where historical reporting is less complete or where seismicity is more spatially diffuse. This interpretation is supported by the tectonic-domain decomposition presented in Fig. 7.
I further agree with the reviewer that the observed coherence cannot be attributed exclusively to tectonic controls. In historically well-documented regions such as Europe, favourable reporting conditions likely contribute to the observed agreement between historical and instrumental catalogues in addition to tectonic factors. For this reason, I have explicitly clarified in the revised manuscript that the coherence metric should not be interpreted as a direct measure of tectonic control alone. Rather, it quantifies the extent to which tectonic organization remains detectable within the historical record after the combined effects of historical, societal, and reporting-related filtering processes.
The revised manuscript therefore adopts a more cautious and nuanced interpretation of the results. The principal conclusion is not that coherence provides a direct measure of tectonic organization, but rather that a statistically meaningful tectonic signal remains observable within historical earthquake catalogues despite the strong influence of non-geophysical factors. I believe that these revisions clarify both the meaning and the limitations of the coherence analysis and better address the central questions posed by the study.

The third point is the geographic scales. I had difficulties in understand the scale definitions. The smoothing scale in km is fine and understandable, but I didn't find any definitions for the geographic grid (line 115) as well as its characteristics (e.g., size, area etc) and why they were chosen (is it 200 km, line 159?). Additionally, throughout the manuscript the geographic definitions could be sharped. For instance, what is 'fine spatial resolution (line 247) or regional scale (line 114)

I thank the reviewer for this useful observation. I agree that the original manuscript did not provide sufficiently explicit definitions of the spatial scales and geographic discretization used throughout the analysis.
To address this issue, I revised the Methods section to clarify the spatial framework adopted in the study. In particular, I now explicitly state that all analyses are performed on a regular global grid with a spatial resolution of 0.5° × 0.5°, corresponding to a nominal cell size of approximately 55 km at the equator. The revised text also clarifies that the grid resolution is substantially finer than the smoothing scales considered and therefore does not control the observed scale-dependent behaviour.
I have additionally clarified the distinction between grid resolution and smoothing scale. The Gaussian smoothing parameter σ is now explicitly defined as the effective spatial averaging scale of the analysis and is systematically varied over six values (50, 75, 100, 150, 200, and 300 km). These scales therefore characterize the spatial resolution at which coherence is evaluated and should not be confused with the dimensions of the underlying geographic grid.
Following the reviewer's suggestion, I also introduced explicit definitions for qualitative expressions such as "fine spatial resolution" and "regional to continental scales". In the revised manuscript, smoothing scales of 50–100 km are defined as fine-scale analyses, whereas smoothing scales of 150–300 km correspond to regional and continental-scale spatial averaging. These definitions are now used consistently throughout the manuscript.

I believe that these revisions substantially improve the clarity, reproducibility, and interpretation of the spatial framework adopted in the study.

Forth, the breakdown by tectonic domains is interesting, but the classification into only three domains and the use of a fixed distance of 200 km from plate boundaries seem simplistic (line 159). It would be worth justifying this choice or testing the sensitivity to alternative definitions of domains and distances of influence.

I thank the reviewer for this valuable suggestion. I agree that the choice of a fixed 200 km boundary influence distance requires justification and sensitivity testing.
To address this concern, I performed an additional sensitivity analysis in which the tectonic-domain decomposition was repeated using alternative boundary influence distances of 100 km and 300 km. The results are presented in the new Supplementary Fig. S2.
The analysis shows that although the absolute coherence contribution assigned to each tectonic domain varies as expected when the spatial extent of the domain masks changes, the qualitative behaviour of the results remains remarkably stable. In particular, the relative ordering of tectonic-domain contributions and their scale dependence are preserved across all tested distances.
These results indicate that the principal conclusions of the study do not depend on the specific 200 km threshold adopted in the main analysis. I have therefore retained the 200 km value as a representative intermediate choice while explicitly documenting the robustness of the results through the new sensitivity analysis. Relevant text has been added to both the Methods and Results sections, and the new Supplementary Fig. S2 has been included in the revised manuscript.

Fifth, in some sections and paragraphs, I find the phrasing to be vague and cumbersome (see or instance lines 188-194). Perhaps another round of edit will sharpen the content and meaning of the text.

Thank you for this observation. I agree that several passages in the original manuscript could be expressed more clearly and concisely. I therefore performed an additional editorial revision of the manuscript, with particular attention to the example indicated by the reviewer (former lines 188–194), simplifying sentence structure, reducing redundancy, and improving clarity of expression. These revisions do not affect the results or interpretations, but improve the readability and precision of the manuscript throughout.

Few more comments:

- Figure 1: Enlarge sub-figure (a) for it is the most important representation. If that map will be structured correctly, then the others might be omitted. Also emphasize the epicentres' location – one can hardly detect some of them. What does the author mean by 'All maps are shown using equal-area cartographic projections with coastlines for geographical reference' (line 397-398)? Perhaps it would be better to describe the actual Coordinate Reference System and the datum used.

  Thank you for this suggestion. Figure 1 has been redesigned to improve readability and visual emphasis on the global distribution of historical earthquake epicentres. The size and visibility of epicentral symbols have been increased, particularly in panel (a), which now provides a clearer representation of the global dataset coverage. I retained the additional panels because they document the different spatial characteristics of the three historical catalogues considered in the study and therefore provide useful context for the interpretation of the results.
  I also revised the figure caption to replace the generic description of the map projection with an explicit statement of the geographic reference system and cartographic projection used. The revised caption now specifies that earthquake locations are referenced to the WGS84 geographic coordinate system and displayed using a Robinson projection.

- Lines 418-419: "The resulting maps primarily reflect the evolving spatial heterogeneity of historical documentation and catalog compilation, rather than physical variations in seismicity". See my first comment

Thank you for this observation. I agree that the original wording was overly categorical and could be interpreted as attributing the observed spatial patterns almost exclusively to reporting-related factors. In the revised manuscript, I replaced this statement with a more balanced formulation acknowledging that historical earthquake distributions reflect the combined influence of documentation and reporting processes, catalogue compilation practices, and genuine tectonic controls. The revised text also emphasizes that the relative importance of these factors varies with spatial scale, which is one of the central conclusions of the study.

- The use of Gaussian smoothing is good. I think ANNI (average nearest neighbor index) might add further statistical support/verification

Thank you for this valuable suggestion. I agree that nearest-neighbour statistics such as the Average Nearest Neighbour Index (ANNI) can provide useful information regarding the clustering characteristics of earthquake distributions. However, the primary objective of this study is not to quantify clustering within individual catalogues, but rather to assess the spatial agreement between historical and instrumental seismicity through a scale-dependent comparison of spatial probability density fields. Because ANNI characterizes the internal spatial structure of a single point pattern and does not directly measure the similarity between two spatial distributions, I considered it complementary rather than central to the objectives of this work. For this reason, I have not introduced a separate ANNI analysis.

- Lines 198-202 contain an internal contradiction which undermine the core claim of the manuscript. In the 1st sentence, the author suggests dominance of major tectonic structure while in the second he underscores the influence of reporting bias: "These features broadly correspond to major tectonic structures, such as subduction zones and active plate boundaries, suggesting that historical earthquake locations contain a non-random tectonic signal when considered at sufficiently large spatial scales. However, even at these scales, the density fields remain far from globally uniform, underscoring the persistent influence of reporting bias"

Thank you for pointing out this ambiguity. I agree that the original wording could be interpreted as presenting tectonic controls and reporting bias as mutually exclusive explanations, which was not my intention. The central conclusion of the study is that both factors contribute to the observed spatial patterns, with their relative importance varying across spatial scales. I have therefore revised this passage to clarify that the large-scale correspondence with major tectonic structures indicates the presence of a detectable tectonic signal, while substantial reporting-related heterogeneity remains superimposed on that signal. The revised text now emphasizes the combined influence of tectonic and historical-documentary factors rather than treating them as competing interpretations.

- Figure 3 might benefit from adding the polylines of the main tectonic units to emphasize the reporting densities superimposed on the tectonic settings.

Thank you for this suggestion. I agree that a direct comparison between reporting-density patterns and tectonic structures improves the interpretation of the results. I have therefore revised Figure 3 by superimposing the major plate boundaries from the PB2002 global plate-boundary model (Bird, 2003) on all density maps. This addition facilitates the visual assessment of the relationship between historical reporting density and large-scale tectonic structures, while preserving the original purpose of the figure. The figure caption has been updated accordingly.

- Figure 4: isn't the results too obvious? The instrumental records clearly reflect the setting of the tectonic units since they are complete (or close to complete) while the historical records are partial and depend on human and historical context. What do we learn that we didn't know from these results?

Thank you for this observation. I agree that the qualitative contrast between historical and instrumental seismicity may appear intuitive. However, the purpose of Figure 4 is not to demonstrate that instrumental catalogues provide a more complete representation of global seismicity. Rather, it serves

as a visual reference for the central question of the study: whether historical earthquake catalogues retain a measurable tectonic signal despite strong reporting heterogeneity.
The key contribution of the manuscript is not the visual comparison itself, but the subsequent quantitative assessment of how much spatial agreement exists, how this agreement varies with spatial scale, and how it differs among tectonic settings (Figs. 6 and 7). Figure 4 therefore provides the observational context required to interpret the coherence analyses rather than constituting the primary result of the study.
To further clarify this point, I have revised the text surrounding Figure 4 and strengthened the connection between the qualitative comparison and the subsequent quantitative analyses. In addition, Figure 4 has been updated by superimposing the major plate boundaries from the PB2002 model (Bird, 2003), allowing a direct visual comparison between historical reporting density, instrumental seismicity, their spatial overlap, and the large-scale tectonic framework. This modification emphasizes that the objective of the figure is not simply to illustrate differences in catalogue completeness, but to provide the visual context for the quantitative coherence analyses that follow.

- Lines 207-208: 'The global spatial coherence, quantified as the integrated overlap between the two density fields, remains substantially below unity, indicating that only a fraction of the instrumental spatial pattern is captured by historical reporting.' What does the author mean by 'below unity'?

  Thank you for this observation. I agree that the expression "below unity" was unnecessarily terse and could be unclear to readers unfamiliar with the definition of the coherence metric. In the revised manuscript, I replaced this wording with a more explicit statement indicating that the observed coherence values are approximately 0.13–0.19, which is well below the theoretical maximum value of 1 corresponding to perfect spatial agreement between the two distributions. This clarification makes the interpretation of the coherence metric more transparent.

- Lines 214-215: I partially agree with the statement. Yes, there are inconsistencies in historical magnitude estimations but that does not mean the data is 'unsuitable for direct quantitative comparison' but rather requires calibration and constraints within reasonable boundaries.

  Thank you for this observation. I agree that the original wording was overly strong. My intention was not to suggest that historical magnitude estimates cannot be used quantitatively, but rather that their use in a global multi-catalogue comparison would require additional calibration and homogenization procedures that fall outside the scope of the present study.
  To clarify this point, I revised the text to acknowledge that historical magnitudes can provide valuable quantitative information when appropriate calibration and harmonization approaches are applied. The revised manuscript now emphasizes that the decision to focus on spatial information was motivated by the substantial heterogeneity and uncertainty of magnitude estimates across the different historical catalogues considered, rather than by an intrinsic inability of historical magnitude data to support quantitative analyses.

- Figure 5 should appear in the methodology section where the author explains why he does not rely on magnitudes.

  Thank you for this suggestion. I agree that Figure 5 is closely related to the rationale for focusing on spatial information rather than magnitude-based metrics. However, I chose to retain the figure within the Results section because it presents empirical characteristics of the datasets, including magnitude availability and inter-catalogue consistency, rather than methodological assumptions.

- Lines 216-222: why this comparison was chosen and what do we learn from the results?

  Thank you for this comment. I agree that the original text focused primarily on describing the observed coherence curves and did not sufficiently explain the rationale for the comparison or its broader implications.
  The purpose of the analysis shown in Figure 6 is not simply to quantify the overlap between historical and instrumental catalogues, but to determine whether a stable and reproducible relationship exists

between spatial scale and historical–instrumental agreement. This relationship provides a quantitative measure of the spatial scales at which historical earthquake records begin to recover information that is consistent with the large-scale organization of instrumental seismicity.
To make this point clearer, I expanded the interpretation of Figure 6 in the revised manuscript and complemented the analysis with additional sensitivity tests using different instrumental reference periods (1900–1950, 1951–1970, and 1971–2025) and different magnitude thresholds (Supplementary Fig. S1). These additional analyses show that, although absolute coherence values vary, the overall coherence–scale relationship remains remarkably stable. This result demonstrates that the emergence of historical–instrumental agreement at regional to continental scales is a robust property of the historical record rather than an artefact of a particular instrumental catalogue or magnitude threshold.
The revised text now explicitly discusses what is learned from this comparison: namely, that historical catalogues retain a measurable tectonic signal, but that this signal becomes detectable only after sufficient spatial aggregation and remains intrinsically limited even at the largest scales considered.

- Figures 6 and 7 could merge into one figure. I would also add visual representation on a map to the 3 inspected tectonic units

  Thank you for this suggestion. I considered merging Figures 6 and 7, but decided to retain them as separate figures because they address different aspects of the analysis. Figure 6 focuses on the scale dependence and robustness of historical–instrumental coherence, whereas Figure 7 examines how this coherence is partitioned among different tectonic environments. Combining both analyses into a single figure would reduce readability and obscure their distinct objectives.
  I agree, however, that a visual representation of the tectonic-domain classification is useful. To address this point, I added a new supplementary figure (Fig. S3) showing the global classification of grid cells into subduction, collision, and intraplate domains based on the PB2002 model and the adopted 200-km boundary-influence distance. The Methods section now explicitly refers to this figure when describing the tectonic-domain framework.

- The last paragraph of the results section contain inaccuracy: one can detect tectonic units from historical data also for local/regional scale (e.g., the DST or the NAF regions)

  Thank you for this observation. I agree that the original wording was too strong and could be interpreted as implying that historical earthquake catalogues do not preserve tectonic information at local or regional scales. This was not my intention.
  The objective of the present study is to assess historical–instrumental coherence at the global scale, where reporting heterogeneity exerts a strong influence on the observed spatial patterns, particularly at small smoothing scales. However, I agree that some historically well-documented regions can preserve recognizable tectonic signatures even at local and regional scales, including major fault systems and plate-boundary structures.
  To address this point, I revised the final paragraph of the Results section. The revised text now clarifies that the increasing emergence of tectonic patterns with spatial scale refers to the global-scale behaviour quantified in this study and does not exclude the possibility that local or regional tectonic structures may already be identifiable in regions characterized by sufficiently rich historical documentation.

- The author is the sole contributor then why does he use the first-person plural(we) (e.g., line 241)

  In the revised manuscript, I reviewed the text and replaced first-person plural expressions with either singular or impersonal formulations, as appropriate.

- Line 249-256: the author describes coherence of the different tectonic units, but it does not necessarily arise from the results. Please further explain.

  Thank you for this comment. I agree that the original text was primarily descriptive and did not sufficiently explain the interpretation of the observed differences among tectonic domains.
  To address this point, I expanded the discussion of Figure 7 in the Results section. The revised text now explains why subduction zones provide the largest contribution to the total coherence, relating

this result to their high long-term seismic activity and the presence of several historically well-documented convergent margins. The text also clarifies that collision belts contribute less because they occupy a comparatively smaller fraction of the classified tectonic domain and are represented by fewer globally distributed seismic provinces, whereas intraplate regions contribute through their large spatial extent despite generally lower seismic activity. In addition, the revised manuscript now explicitly links the interpretation of Fig. 7 to the tectonic-domain classification shown in Supplementary Fig. S3.
These additions make the reasoning connecting the observed coherence patterns to the tectonic-domain framework more explicit.

- Repeating sentences and meanings: e.g., lines 256-259 were already stated in line 213

Thank you for pointing this out. I agree that the original Results section contained some redundancy in the interpretation of the coherence analysis. In particular, similar conclusions regarding the persistence of a large-scale tectonic signal in historical catalogues were stated in consecutive paragraphs.
To improve clarity and avoid repetition, I revised the relevant section by removing redundant statements and retaining only the information necessary to support the interpretation of the coherence results. The revised text is more concise while preserving the scientific meaning of the discussion.

- Lines 261-262: I am not sure I agree with this statement. The author should further explain and justify. The same goes for lines 267-269

To address this concern, I revised the concluding discussion and adopted a more cautious interpretation of the results. In particular, I removed wording that could be interpreted as implying a formal limit on the tectonic information content of historical catalogues and instead emphasized that the proposed framework provides a quantitative assessment of the spatial scales at which agreement between historical and instrumental seismicity becomes more evident.
The revised Conclusions now focus on the scale-dependent nature of the observed coherence, the influence of tectonic setting, and the practical implications for the interpretation of historical earthquake catalogues, while avoiding broader claims that are not directly supported by the analysis.

- Data and python code: should be but are not accessible

The Python scripts used in this study will be deposited in a public repository upon acceptance of the manuscript. A permanent repository link will be included in the final published version of the article.