https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

Scale-dependent spatial coherence between historical and instrumental earthquake catalogues at the global scale

Antonino D'Alessandro
Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Nazionale Terremoti, Rome, Italy
antonino.dalessandro@ingv.it

Abstract

Historical earthquake catalogues extend seismic observations back by several centuries and are widely used in seismic hazard and tectonic studies, yet their global-scale informational content remains difficult to quantify due to strong spatial, temporal, and magnitude-dependent reporting biases. In this study, we present a quantitative, spatially explicit assessment of the consistency between global historical (1600–1899) and early instrumental (1900–1950, $M \geq 5.5$) earthquake catalogues. Rather than relying on magnitude-based comparisons, we represent earthquake occurrence as spatial probability density fields obtained through Gaussian smoothing and define a scale-dependent spatial coherence metric based on the overlap between historical and instrumental distributions. This approach allows us to isolate large-scale tectonic signal from localized reporting artefacts and to systematically explore the role of spatial scale. Our results show that spatial coherence between historical and instrumental seismicity is low at small scales and increases monotonically with smoothing length, reaching moderate values only at regional to continental scales. Even at the largest scales considered, coherence remains well below unity, indicating that only a limited fraction of the instrumental spatial pattern is recoverable from historical data. Decomposition by tectonic domain reveals that subduction zones dominate the historical–instrumental agreement, while continental collision belts and intraplate regions contribute substantially less. These findings demonstrate that global historical earthquake catalogues contain a detectable but intrinsically limited imprint of tectonic structure. Meaningful use of historical seismicity at the global scale therefore requires explicit consideration of spatial scale and tectonic

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

context. The framework proposed here provides a transparent and reproducible basis for evaluating the reliability of historical earthquake data in seismic hazard and global seismotectonic applications.

1. Introduction

Earthquake catalogs constitute the primary empirical basis for seismic hazard assessment, tectonic studies, and long-term analyses of earthquake occurrence. While instrumental seismic networks provide spatially and temporally homogeneous observations only since the early–mid twentieth century, historical earthquake catalogs extend the temporal window of observation back by several centuries, offering the potential to investigate rare large events and long-term seismic patterns (e.g. Guidoboni and Stucchi, 2011; Albini et al., 2014).

Despite their importance, historical earthquake catalogs are affected by profound and well-known limitations. The reporting of pre-instrumental earthquakes depends on population density, literacy, administrative organization, and the survival of historical sources, resulting in highly heterogeneous spatial and temporal coverage (Ambraseys, 1989; Guidoboni et al., 2007). In addition, magnitude information is often incomplete, uncertain, or retrospectively inferred from macroseismic intensity data, introducing systematic biases that complicate quantitative analyses (Bakun and Wentworth, 1997; Stucchi et al., 2013). These limitations are particularly critical when historical data are used in a probabilistic seismic hazard context, where assumptions on completeness and representativeness directly affect hazard estimates (Albini et al., 2015).

At the global scale, several compilations of historical seismicity are currently available, including the Global Historical Earthquake Catalogue (GHEC; Albini et al., 2014), the European Pre-Instrumental Catalogue (EPICA; Rovida et al., 2022), and the NOAA Significant Earthquake Database (NOAA, 2023). These datasets differ substantially in scope, selection criteria, and methodological philosophy, but they share common limitations related to reporting completeness and spatial representativeness. Recent studies have emphasized that catalog completeness should not be regarded as a binary property, but rather as a scale-, magnitude-, and region-dependent quantity, whose uncertainty must be explicitly considered (Woessner et al., 2015; Gómez Capera et al., 2020). As a result, the degree to which historical catalogs capture the underlying tectonic signal of global seismicity remains an open and debated question.

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

Previous investigations have primarily addressed this issue at regional scales, often through qualitative comparisons or by focusing on specific well-documented areas such as Europe, the Mediterranean, or Japan (Ambraseys and Jackson, 1998; Guidoboni and Ebel, 2009). At the global scale, however, a systematic and quantitative assessment of the spatial consistency between historical and instrumental seismicity is still lacking. In particular, little attention has been paid to the spatial scales at which historical earthquake distributions may transition from being dominated by reporting artifacts to reflecting genuine tectonic structure.

A key challenge arises from the intrinsic mismatch between historical and instrumental datasets. Instrumental catalogs, even in their early phases, provide more uniform detection thresholds and spatial coverage, especially for moderate-to-large earthquakes (e.g. Gutenberg and Richter, 1954; Utsu, 2002). For this reason, early instrumental seismicity can be used as an approximate reference for the large-scale spatial organization of earthquake occurrence. Spatially explicit and probabilistic comparison frameworks, already widely adopted in earthquake forecast testing and model evaluation (Schorlemmer et al., 2010; Helmstetter et al., 2007), offer a natural basis for such analyses.

In this study, we introduce a quantitative, spatially explicit framework to assess the coherence between historical and instrumental earthquake catalogs at the global scale. Rather than relying on magnitude-based comparisons, which are strongly affected by historical incompleteness (Mignan and Woessner, 2012), we focus on the spatial distributions of earthquake occurrence. Historical (1600–1899) and early instrumental (1900–1950, $M \geq 5.5$) seismicity are represented as spatial probability density fields obtained through Gaussian smoothing, allowing us to define a scale-dependent measure of spatial coherence.

By systematically varying the smoothing scale and decomposing the analysis by tectonic domain, we aim to answer three central questions:

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?

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

By explicitly quantifying the limits and conditions under which historical earthquake data reflect tectonic reality, this work provides a robust methodological basis for their appropriate and informed use in global-scale analyses.

2. Data

This study relies exclusively on openly accessible earthquake catalogues in order to ensure transparency, reproducibility, and long-term traceability of the results. Two distinct classes of data are considered: historical earthquake catalogues documenting pre-instrumental seismicity, and an early instrumental catalogue used as a spatial reference for global earthquake occurrence.

Historical seismicity is represented by three major compilations covering the period 1600–1899: the Global Historical Earthquake Catalogue (GHEC; Albini et al., 2014), the European Pre-Instrumental Earthquake Catalogue (EPICA; Rovida et al., 2022), and the NOAA Significant Earthquake Database (NOAA, 2023). GHEC provides a globally harmonized dataset derived from critically reviewed regional and national catalogues, with explicit documentation of sources and selection criteria. EPICA represents the most up-to-date and systematically revised compilation of pre-instrumental earthquakes for Europe and the Mediterranean region. The NOAA database complements these resources by including globally distributed significant historical earthquakes, albeit with heterogeneous reporting standards and variable parameter completeness.

All historical events reported between 1600 and 1899 were initially retained. Quality control procedures were applied to exclude events lacking geographic coordinates or a defined year of occurrence. No attempt was made to homogenize historical magnitude estimates or to convert macroseismic intensities into magnitudes, as such procedures introduce additional epistemic uncertainty and are known to be strongly dependent on region- and period-specific assumptions (Bakun and Wentworth, 1997; Stucchi et al., 2013). Consequently, historical data are treated exclusively in terms of spatial occurrence, without using magnitude as a quantitative variable.

To provide a reference representation of the large-scale spatial organization of global seismicity, an early instrumental earthquake catalogue covering the period 1900–1950 was constructed from the United States Geological Survey (USGS) database. A conservative magnitude threshold of $M \geq 5.5$ was adopted to ensure near-global detectability even during the early instrumental era (Gutenberg and Richter, 1954; Utsu, 2002).

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

Although this catalogue does not achieve the completeness of modern global datasets, it offers substantially more homogeneous spatial coverage than historical records and is therefore used here as a proxy for the underlying tectonic control on earthquake occurrence at regional to continental scales.

All catalogues were projected onto a common global geographic grid and processed using identical spatial discretization. Potential duplicate events across catalogues were identified based on spatial and temporal proximity and flagged for diagnostic purposes, but not forcibly merged, in order to preserve the original information content of each dataset. Subsequent analyses are based on spatial probability density fields derived from event locations, rather than on individual earthquake parameters. This strategy minimizes sensitivity to catalogue-specific uncertainties in magnitude, depth, and origin time, and ensures that the results reflect robust spatial patterns rather than artefacts of catalogue construction.

3. Methods

The methodological framework adopted in this study is designed to quantify the spatial consistency between historical and instrumental earthquake catalogues in a scale-dependent and tectonically informed manner. The analysis is entirely based on earthquake locations and avoids the direct use of magnitude information, which is known to be highly heterogeneous and incomplete in historical datasets.

Both historical (1600–1899) and instrumental (1900–1950) earthquake catalogues are first represented as discrete spatial point processes on a common global geographic grid. To enable a meaningful comparison between datasets characterized by different levels of spatial completeness and reporting noise, earthquake locations are transformed into continuous spatial probability density fields through isotropic Gaussian kernel smoothing. For a given catalogue k(historical or instrumental), the smoothed spatial density field $p_k^{\sigma}(\mathbf{x})$ is defined as

$$p_k^{\sigma}(\mathbf{x}) = \frac{1}{N_k} \sum_{i=1}^{N_k} \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\| \mathbf{x} - \mathbf{x}_i \|^2}{2\sigma^2}\right), \qquad (1)$$

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

where $\mathbf{x}_i$ denotes the epicentral location of the i-th earthquake in catalogue k, N_k is the total number of events in that catalogue, and $\sigma$is the Gaussian smoothing length scale. This normalization ensures that each density field integrates to unity over the global domain and can therefore be interpreted as a spatial probability density function.

The smoothing scale $\sigma$is treated as a free parameter and systematically varied in the range 50–300 km. This interval spans scales from sub-regional, where historical reporting noise is expected to dominate, to continental scales, where the imprint of large-scale tectonic structure should emerge. All analyses are performed on identical grids and with identical smoothing parameters for both historical and instrumental datasets, ensuring methodological consistency.

To quantify the spatial agreement between the two density fields, we define a scale-dependent spatial coherence metric $C(\sigma)$as the total overlap between the historical and instrumental probability density functions,

$$C(\sigma) = \int_{\Omega} \min\left(p^{\sigma}_{\mathrm{hist}}(\mathbf{x}), p^{\sigma}_{\mathrm{inst}}(\mathbf{x})\right) \mathrm{d}\mathbf{x}, \qquad (2)$$

where Ωdenotes the global spatial domain. By construction, $0 \leq C(\sigma) \leq 1$, with $C(\sigma) = 1$ corresponding to identical spatial distributions and $C(\sigma) = 0$ indicating no spatial overlap. This metric emphasizes shared large-scale spatial structure while being insensitive to localized discrepancies, making it particularly suitable for comparing datasets with heterogeneous completeness.

To investigate the role of tectonic setting, the global coherence is further decomposed into contributions from distinct tectonic domains. Tectonic regions are defined using the PB2002 global plate boundary model (Bird, 2003), which provides a geometrically consistent representation of plate boundaries and their kinematic classification. Each grid cell is assigned to one of three broad tectonic domains: subduction zones, continental collision belts, or intraplate regions. Grid cells located within 200 km of a PB2002 boundary segment are classified according to the dominant boundary type, while remaining cells are assigned to the intraplate domain.

For each tectonic domain d, a domain-specific coherence contribution $C_d(\sigma)$ is computed as

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

$$C_d(\sigma) = \int_{\Omega_d} \min\left(p^{\sigma}_{\text{hist}}(\mathbf{x}), p^{\sigma}_{\text{inst}}(\mathbf{x})\right) \mathrm{d}\mathbf{x}, \qquad (3)$$

where $\Omega_d \subset \Omega$ denotes the spatial domain associated with tectonic setting d. By construction, the sum of the domain-specific contributions exactly reproduces the global coherence,

$$C(\sigma) = \sum_d C_d\,(\sigma).$$

This decomposition allows the relative contribution of different tectonic environments to the overall historical–instrumental agreement to be quantified as a function of spatial scale.

No temporal smoothing is applied beyond the selection of the analysis periods, and no declustering is performed. This choice reflects the focus of the study on long-term spatial patterns rather than on short-term seismic interactions. Sensitivity to catalogue size and smoothing scale is addressed explicitly through the systematic exploration of σ, while uncertainties related to magnitude, depth, and focal mechanism are intentionally excluded from the analysis to avoid introducing poorly constrained assumptions.

4. Results

The spatial distribution of historical earthquake reporting exhibits pronounced heterogeneity at the global scale. Figure 1 illustrates the geographic coverage of historical seismicity between 1600 and 1899 as derived from the combined GHEC, EPICA, and NOAA datasets. Reported events are strongly concentrated in Europe and the Mediterranean region, Japan, parts of East Asia, and selected areas of the Americas, whereas vast regions, including most of Africa, oceanic domains, and large intraplate areas, display sparse or absent reporting. This pattern reflects the combined influence of tectonic activity and historically contingent reporting conditions rather than a homogeneous representation of global seismicity.

The temporal evolution of historical reporting further emphasizes its non-stationary character. Figure 2 shows the smoothed temporal evolution of the number of reported historical earthquakes using a wide moving window. The gradual increase in reporting rate over time, particularly from the eighteenth century onward,

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

does not exhibit clear episodic behaviour that could be robustly attributed to changes in global seismicity rates. Instead, the observed trends are consistent with progressive improvements in documentation, communication, and archival preservation. This temporal variability highlights the difficulty of interpreting raw historical event counts in terms of physical seismic processes.

Spatial probability density maps of historical seismicity, constructed through Gaussian smoothing, provide further insight into the scale-dependent nature of historical information (Fig. 3). At relatively small smoothing scales, the density maps are dominated by localized clusters associated with well-documented regions, while at larger scales persistent large-scale features emerge. 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.

A direct comparison between historical and instrumental spatial density fields is presented in Figure 4. The instrumental reference catalogue (1900–1950, $M \geq 5.5$) displays a spatial pattern closely aligned with global plate boundaries, whereas the historical density field shows only partial correspondence. The resulting overlap between the two distributions is spatially limited and concentrated in a small number of regions. 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.

The limitations associated with historical magnitude information are illustrated in Figure 5. The fraction of historical events with reported magnitudes varies strongly through time and across catalogues, with a marked increase only in the late nineteenth century. Complementary cumulative magnitude distributions reveal substantial inconsistencies between catalogues, reflecting differences in magnitude assignment practices and data availability. These results confirm that historical magnitude data are unsuitable for direct quantitative comparison with instrumental catalogues at the global scale and justify the exclusive focus on spatial information in the present analysis.

The dependence of historical–instrumental spatial coherence on the smoothing scale is quantified in Figure 6. The coherence $C(\sigma)$ increases monotonically with increasing σ, from values of approximately 0.13 at 50 km to about 0.19 at 300 km. This behaviour indicates that small-scale discrepancies, dominated by reporting noise

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

and localized incompleteness, progressively diminish as spatial averaging increases. No clear saturation is observed at the smallest scales, whereas a gradual approach to an asymptotic value occurs at larger scales, suggesting that only a limited fraction of the instrumental spatial structure is recoverable from historical data even after substantial smoothing.

The role of tectonic setting in controlling spatial coherence is further examined in Figure 7, where the global coherence is decomposed into contributions from subduction zones, continental collision belts, and intraplate regions. Subduction zones provide the dominant contribution to the total coherence across all smoothing scales, followed by intraplate regions, while collision zones contribute comparatively less. The relative contribution of each tectonic domain remains stable with increasing σ, indicating that the scale dependence of coherence is largely driven by the emergence of large-scale tectonic patterns rather than by changes in the relative importance of different tectonic environments.

Overall, the results demonstrate that historical earthquake catalogues contain a detectable but limited imprint of global tectonic structure. This imprint becomes apparent only when spatial information is analysed at sufficiently large scales and remains strongly modulated by tectonic setting. At smaller scales, historical seismicity is dominated by reporting artefacts, whereas at larger scales a partial convergence toward the instrumental spatial pattern is observed.

6. Conclusions

This study presents a quantitative, spatially explicit assessment of the consistency between global historical and early instrumental earthquake catalogues, with the aim of clarifying the extent to which historical seismicity reflects the underlying tectonic organization of earthquake occurrence. By adopting a scale-dependent approach based on spatial probability density fields and a robust coherence metric, we explicitly separate the influence of reporting bias from large-scale tectonic signal.

The results demonstrate that global historical earthquake catalogues contain a detectable but limited imprint of tectonic structure. Spatial coherence between historical (1600–1899) and instrumental (1900–1950, $M \geq 5.5$) seismicity is systematically low at small spatial scales and increases monotonically with increasing smoothing length, reaching moderate values only at regional to continental scales. This behaviour indicates

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

that historical earthquake locations are dominated by reporting artefacts at fine spatial resolution, while meaningful tectonic information emerges only after substantial spatial averaging.

The decomposition of spatial coherence by tectonic domain highlights a strong control exerted by tectonic setting. Subduction zones provide the dominant contribution to historical–instrumental agreement at all scales, reflecting both their high seismic productivity and their long-standing documentation in historical sources. Continental collision belts and intraplate regions exhibit markedly lower coherence, underscoring the combined effects of distributed deformation, lower seismicity rates, and heterogeneous historical reporting. These findings demonstrate that the informational content of historical catalogues is not uniform across tectonic environments and must be interpreted accordingly.

The pronounced inconsistencies in historical magnitude reporting further confirm that global historical earthquake catalogues are unsuitable for direct magnitude-based quantitative analyses. Consequently, spatial representations that do not rely on magnitude information provide a more robust basis for extracting physically meaningful information from historical data at the global scale.

From a broader perspective, the results place an objective upper bound on the amount of tectonic information that can be recovered from global historical earthquake catalogues. While these datasets remain invaluable for regional studies and for documenting rare destructive events, their use in global-scale seismic hazard or tectonic analyses requires explicit consideration of spatial scale, tectonic context, and reporting bias. Approaches that neglect these factors risk overinterpreting patterns that are primarily controlled by non-geophysical processes.

The framework developed here offers a transparent and reproducible methodology for evaluating the spatial reliability of historical seismicity and can be readily extended to other catalogues, time periods, or alternative spatial representations. More generally, this study contributes to a clearer and more quantitative understanding of the role and limitations of historical earthquake data in modern seismological research.

Code availability

All data processing, analysis, and figure generation were performed using custom Python scripts developed specifically for this study. The code implements the full workflow described in the Methods section, including

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

data download, quality control, spatial smoothing, computation of scale-dependent spatial coherence, and tectonic-domain decomposition. The scripts are available from the authors upon reasonable request.

Data availability

This study is based exclusively on openly accessible earthquake catalogues. Historical seismicity data were obtained from the Global Historical Earthquake Catalogue (GHEC), the European Pre-Instrumental Earthquake Catalogue (EPICA), and the NOAA Significant Earthquake Database. Early instrumental earthquake data (1900–1950) were retrieved from the United States Geological Survey (USGS) earthquake catalogue. Tectonic domain definitions are based on the PB2002 global plate boundary model. All datasets used in this study are publicly available from the original data providers through their respective online repositories, as cited in the References section. No proprietary or restricted-access data were used.

Author contribution

Antonino D'Alessandro conceptualized the study and designed the methodological framework. He compiled and curated the historical and instrumental earthquake datasets, developed the analysis code, and performed all data processing, statistical analyses, and figure generation. Antonino D'Alessandro interpreted the results and wrote the manuscript (Conceptualization, Methodology, Software, Data curation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing)

Competing interests

The author declare that he has no conflict of interest.

Disclaimer

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

The views and interpretations expressed in this manuscript are those of the author and do not necessarily reflect the official policies or positions of the institutions with which the author is affiliated. The use of publicly available datasets does not imply endorsement by the data providers, who bear no responsibility for the analyses, interpretations, or conclusions presented in this study.

Acknowledgements

The author acknowledges the institutions and individuals who contributed, directly or indirectly, to this research. The study benefited from openly accessible earthquake catalogues and related resources provided by international and national data centers, including the Global Historical Earthquake Catalogue (GHEC), the European Pre-Instrumental Earthquake Catalogue (EPICA), the NOAA Significant Earthquake Database, and the United States Geological Survey (USGS). The author also acknowledges the availability of the PB2002 global plate boundary model, which was essential for the tectonic-domain analyses.

This work made use of research infrastructure and computational resources available at the author's host institution, which supported data processing, analysis, and visualization. The author is grateful to colleagues for informal discussions that helped clarify methodological choices and interpret the results. Any remaining errors or omissions are the sole responsibility of the author.

Financial support

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

7. References

Albini, P., Musson, R. M. W., Rovida, A., Locati, M., Gomez Capera, A. A., and Stucchi, M.: The Global Historical Earthquake Catalogue (GHEC), *Ann. Geophys.*, **57**, S0435, https://doi.org/10.4401/ag-6309, 2014.

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

Albini, P., Rovida, A., Locati, M., Gomez Capera, A. A., and Stucchi, M.: The challenge of assessing seismic hazard in historical time, *Ann. Geophys.*, **58**, S0101, https://doi.org/10.4401/ag-6623, 2015.

Ambraseys, N. N.: *Earthquakes in the Mediterranean and Middle East*, Cambridge University Press, Cambridge, UK, 1989.

Ambraseys, N. N. and Jackson, J. A.: Faulting associated with historical and recent earthquakes in the Eastern Mediterranean region, *Geophys. J. Int.*, **133**, 390–406, https://doi.org/10.1046/j.1365-246X.1998.00508.x, 1998.

Bakun, W. H. and Wentworth, C. M.: Estimating earthquake location and magnitude from seismic intensity data, *Bull. Seismol. Soc. Am.*, **87**, 1502–1521, 1997.

Bird, P.: An updated digital model of plate boundaries, *Geochem. Geophys. Geosyst.*, **4**(3), 1027, https://doi.org/10.1029/2001GC000252, 2003.

Gómez Capera, A. A., Albini, P., Stucchi, M., and Rovida, A.: Completeness of historical earthquake catalogues and its impact on seismic hazard, *Nat. Hazards Earth Syst. Sci.*, **20**, 1905–1923, https://doi.org/10.5194/nhess-20-1905-2020, 2020.

Guidoboni, E. and Ebel, J.: *Earthquakes and Tsunamis in the Past: A Guide to Techniques in Historical Seismology*, Cambridge University Press, Cambridge, UK, 2009.

Guidoboni, E. and Stucchi, M.: Seismic history and seismic hazard, in: *Seismic Hazard Assessment*, Springer, Dordrecht, 1–24, https://doi.org/10.1007/978-94-007-0418-4_1, 2011.

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

Guidoboni, E., Ferrari, G., Gasperini, P., Mariotti, D., Comastri, A., Tarabusi, G., and Valensise, G.: CFTI4Med, Catalogue of Strong Earthquakes in Italy (461 B.C.–1997), INGV, https://doi.org/10.6092/INGV.IT-CFTI4MED, 2007.

Gutenberg, B. and Richter, C. F.: *Seismicity of the Earth and Associated Phenomena*, 2nd edn., Princeton University Press, Princeton, USA, 1954.

Helmstetter, A., Kagan, Y. Y., and Jackson, D. D.: Comparison of short-term and long-term earthquake forecast models for southern California, *Bull. Seismol. Soc. Am.*, **97**, 405–425, https://doi.org/10.1785/0120060090, 2007.

Mignan, A. and Woessner, J.: Estimating the magnitude of completeness for earthquake catalogs, *J. Geophys. Res.*, **117**, B08302, https://doi.org/10.1029/2011JB008888, 2012.

NOAA: Significant Earthquake Database, National Centers for Environmental Information (NCEI), https://www.ngdc.noaa.gov/hazel/view/hazards/earthquake/search, 2023.

Rovida, A., Locati, M., Camassi, R., Lolli, B., and Gasperini, P.: The European Pre-Instrumental Earthquake Catalogue (EPICA), *Earth Syst. Sci. Data*, **14**, 5369–5389, https://doi.org/10.5194/essd-14-5369-2022, 2022.

Schorlemmer, D., Zechar, J. D., Werner, M. J., Field, E. H., Jackson, D. D., and Jordan, T. H.: First results of the Regional Earthquake Likelihood Models experiment, *Seismol. Res. Lett.*, **81**, 338–343, https://doi.org/10.1785/gssrl.81.3.338, 2010.

Stucchi, M., Albini, P., Rovida, A., Locati, M., Gomez Capera, A. A., and Viganò, D.: Assessing the completeness of macroseismic intensity data, *Bull. Seismol. Soc. Am.*, **103**, 243–256, https://doi.org/10.1785/0120120022, 2013.

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

384 Utsu, T.: Statistical features of seismicity, in: *Int. Handb. Earthquake Eng. Seismol.*, Part A, Academic Press,
385 719–732, https://doi.org/10.1016/S0074-6142(02)80210-5, 2002.
386
387 Woessner, J., Laurentiu, D., Giardini, D., Crowley, H., Cotton, F., Grünthal, G., et al.: The 2014 European
388 Seismic Hazard Model: key components and results, *Bull. Earthquake Eng.*, **13**, 3553–3596,
389 https://doi.org/10.1007/s10518-015-9795-1, 2015.

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

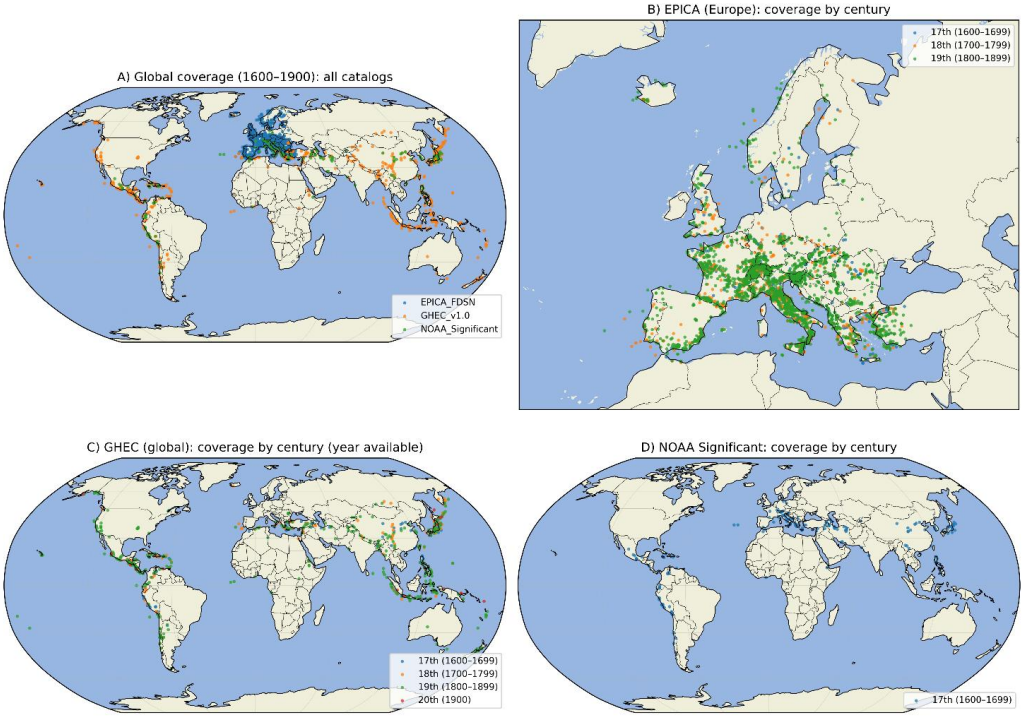


Figure 1. *Spatial coverage of historical earthquake catalogs considered in this study for the period 1600–1900. (A) Global distribution of events from the three datasets used: the Global Historical Earthquake Catalogue (GHEC v1.0), the European Pre-Instrumental Earthquake Catalogue (EPICA), and the NOAA Significant Earthquake Database. (B) Spatial distribution of EPICA events over Europe, coloured by century (17th, 18th, and 19th centuries). (C) Global distribution of GHEC events coloured by century, restricted to earthquakes for which a reliable occurrence year could be extracted from the original metadata. (D) Global distribution of NOAA Significant Earthquake Database events, coloured by century. All maps are shown using equal-area cartographic projections with coastlines for geographical reference. Differences in spatial coverage between panels primarily reflect intrinsic heterogeneities in historical documentation, catalog compilation criteria, and temporal completeness rather than true variations in seismicity.*

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

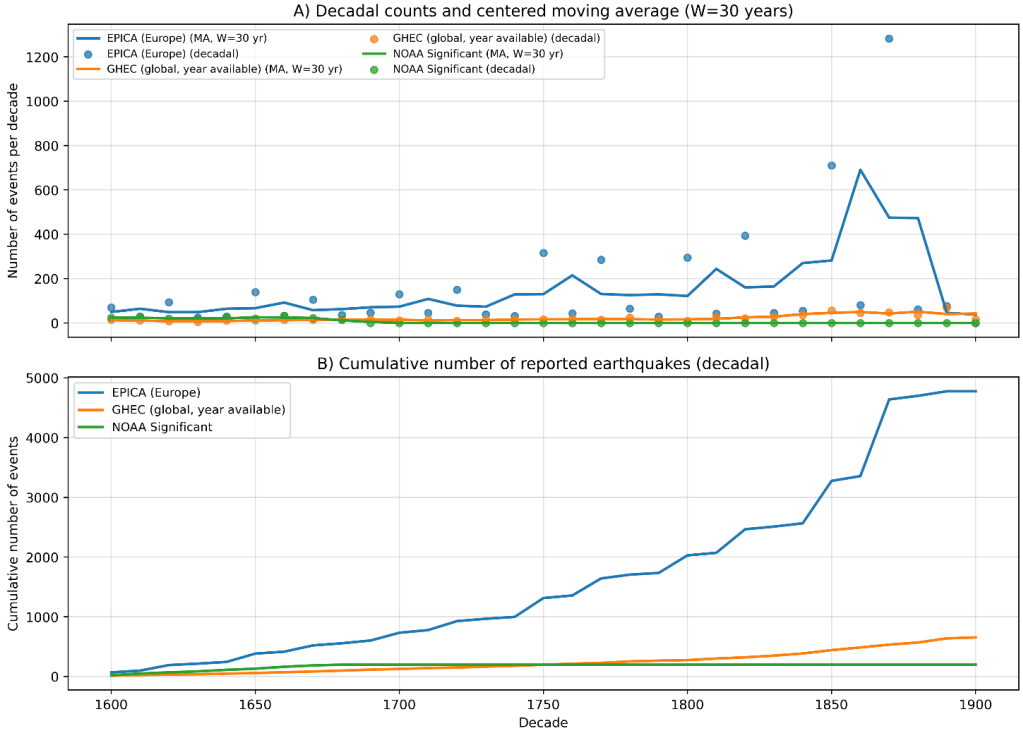
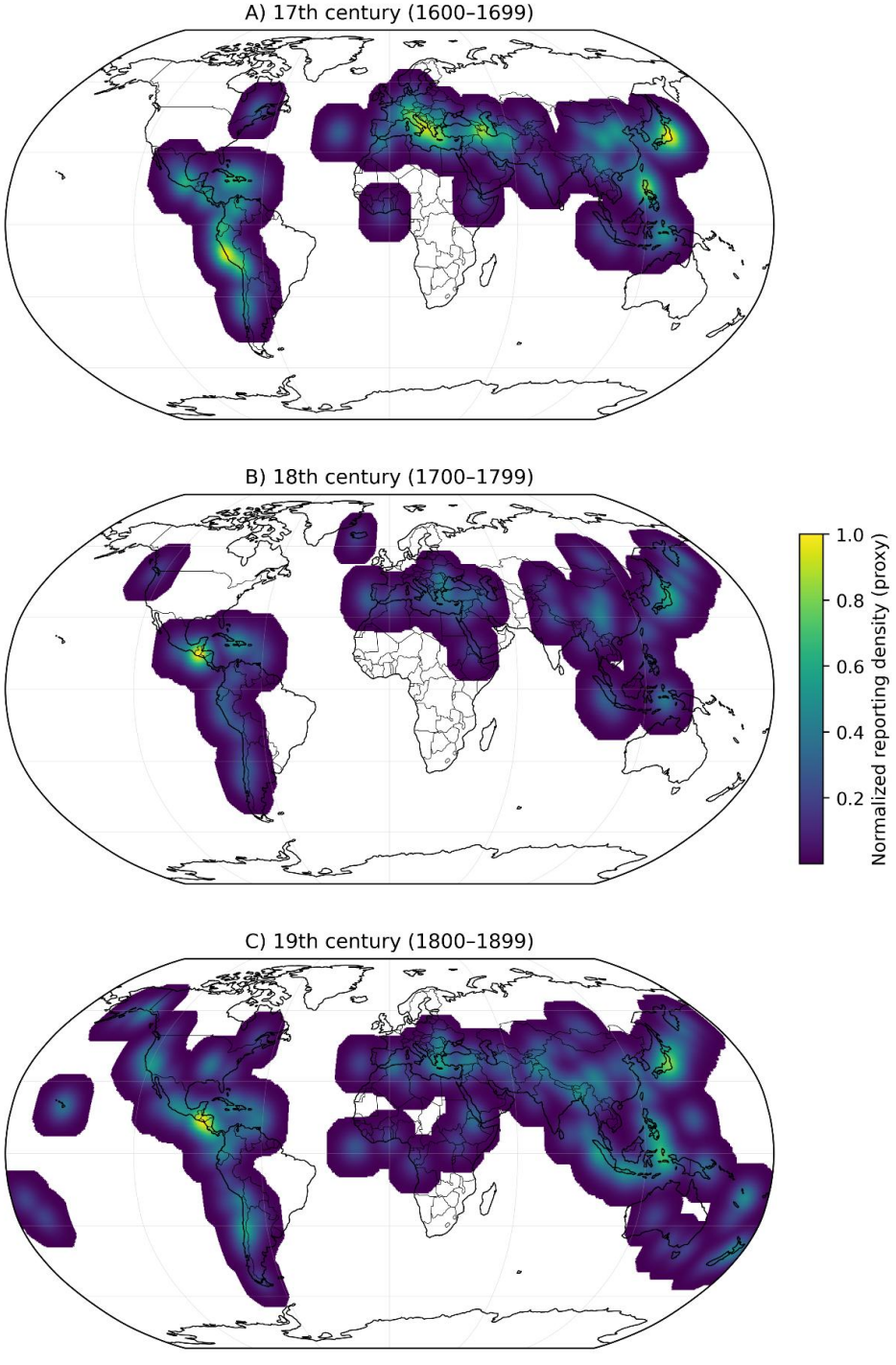


Figure 2. *Temporal evolution of reported historical earthquakes in the study period 1600–1900. (A) Decadal counts* $N(t)$ *for each catalog (symbols) together with a centred moving-average estimate* $\bar{N}_W(t)$ *(solid lines; window length* $W = 30$ *years) to stabilise temporal variability associated with historical dating uncertainty and heterogeneous reporting practices. (B) Cumulative number of reported earthquakes computed from the decadal counts, highlighting differences in long-term catalog growth. EPICA represents European pre-instrumental reporting, whereas GHEC and NOAA provide global compilations with different inclusion criteria; consequently, the curves primarily reflect temporal completeness and documentary/compilation effects rather than true changes in seismicity rates.*

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026



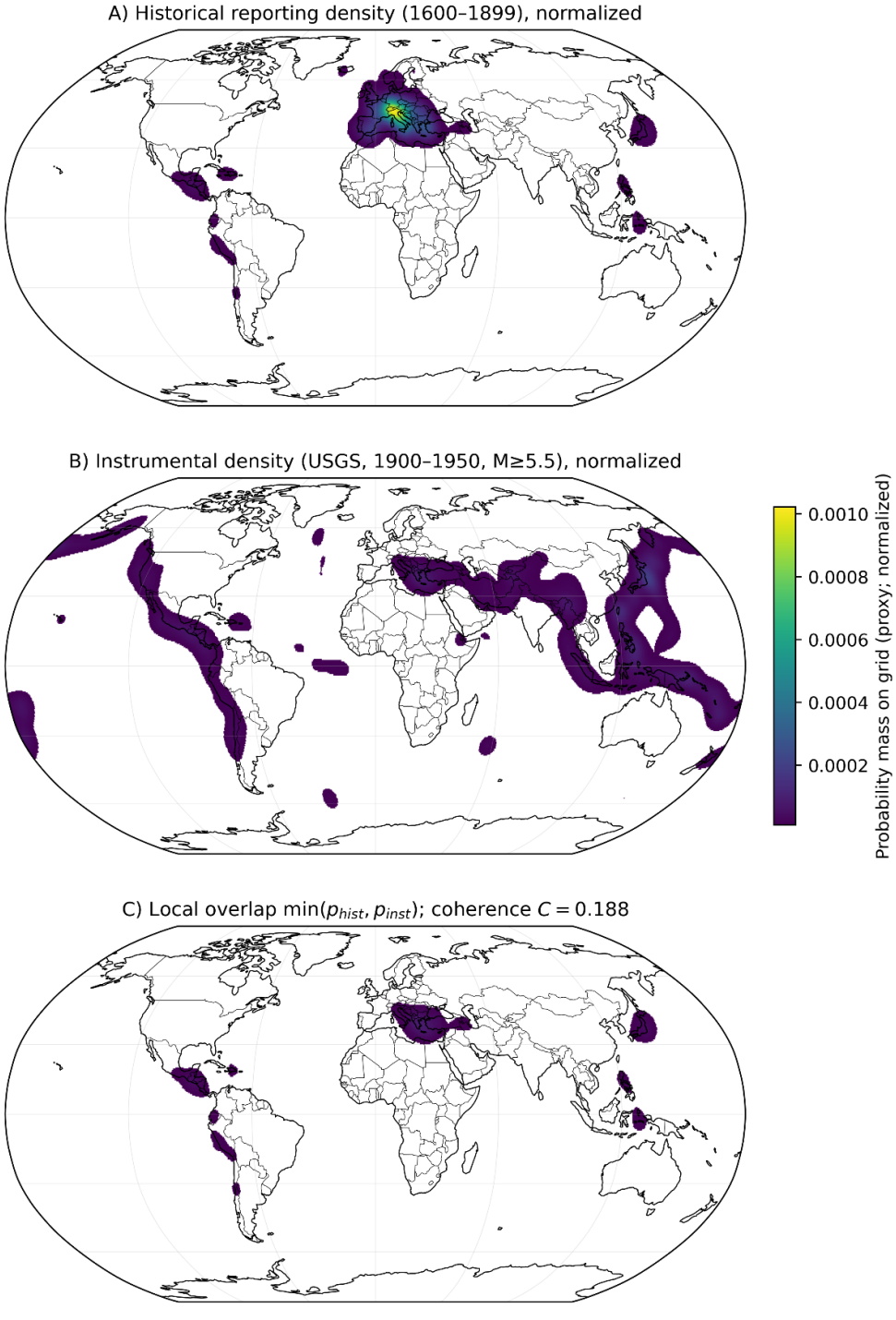
410

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

Figure 3. *Spatial patterns of historical earthquake reporting across centuries (1600–1899) derived from the global catalogs used in this study. Panels show the kernel-smoothed spatial density of reported earthquakes (coverage proxy) for (A) 1600–1699, (B) 1700–1799, and (C) 1800–1899, mapped on a global Robinson projection with coastlines and political borders for reference. For each century, the density field is computed from event locations on a regular geographic grid, smoothed to highlight coherent reporting clusters, and then normalized to the maximum value within that century to enable qualitative comparison of spatial coverage patterns. White areas indicate grid cells with negligible or null reporting density. The resulting maps primarily reflect the evolving spatial heterogeneity of historical documentation and catalog compilation, rather than physical variations in seismicity.*

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

420



421

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

Figure 4. *Comparison between historical and early instrumental earthquake reporting at the global scale. (A) Kernel-smoothed spatial density of reported historical earthquakes for the period 1600–1899, normalized to a probability mass function on a regular geographic grid. (B) Corresponding normalized spatial density for early instrumental seismicity derived from the USGS catalog for 1900–1950, considering events with magnitude* $M \geq 5.5$. *(C) Local spatial overlap between historical and instrumental reporting, computed as* $min\,(p_{hist}, p_{inst})$ *at each grid cell. The scalar coherence index* $C = \sum_i min\,(p_{hist,i}, p_{inst,i})$ *quantifies the global spatial consistency between the two distributions. All maps are shown in Robinson projection with coastlines and political borders for reference; white areas indicate negligible probability mass. The limited overlap highlights the strong spatial bias of historical earthquake documentation relative to the early instrumental record, reflecting differences in population density, archival preservation, and catalog compilation rather than physical variations in seismicity.*

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

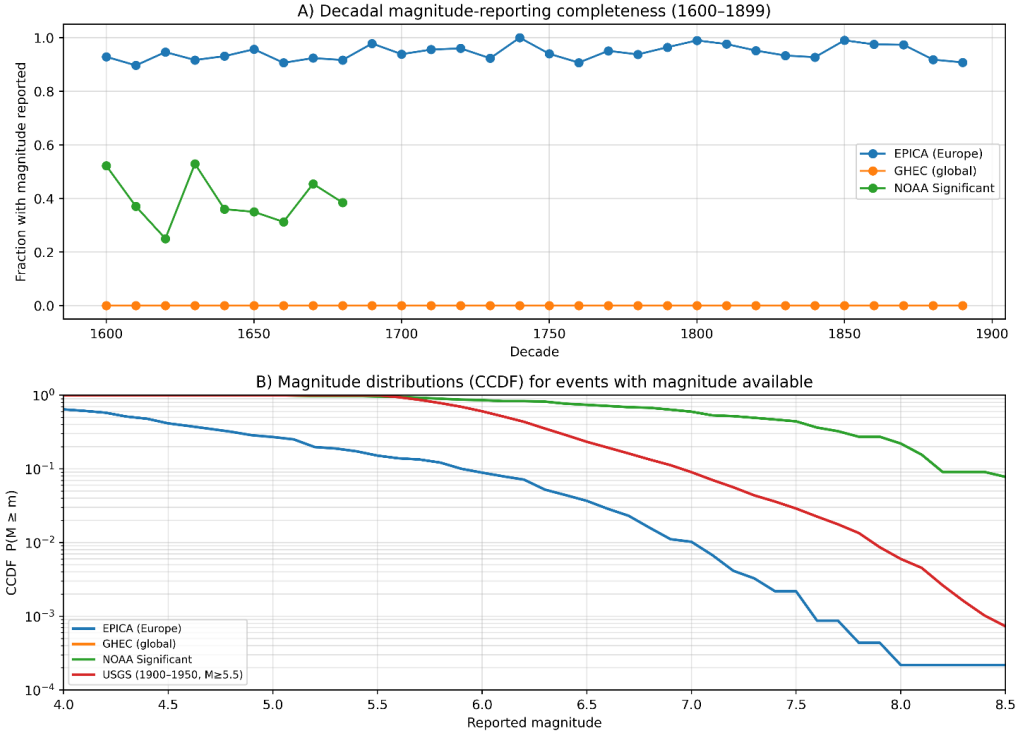


Figure 5. *Magnitude reporting bias in historical earthquake catalogs. (A) Decadal fraction of earthquakes with a reported magnitude for the main historical catalogs considered in this study (EPICA, GHEC, NOAA Significant) over the period 1600–1899. This metric quantifies the temporal completeness of magnitude parametrization and highlights substantial differences in reporting practices across datasets. (B) Complementary cumulative distribution functions (CCDFs) of reported magnitudes for events with available magnitude estimates. Historical catalogs are compared with an early instrumental reference dataset (USGS, 1900–1950, M ≥ 5.5). Deviations from linear Gutenberg–Richter scaling and systematic offsets between catalogs reflect reporting thresholds, selective documentation, and heterogeneous magnitude definitions, emphasizing that historical magnitude distributions primarily represent reporting bias rather than underlying seismicity.*

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

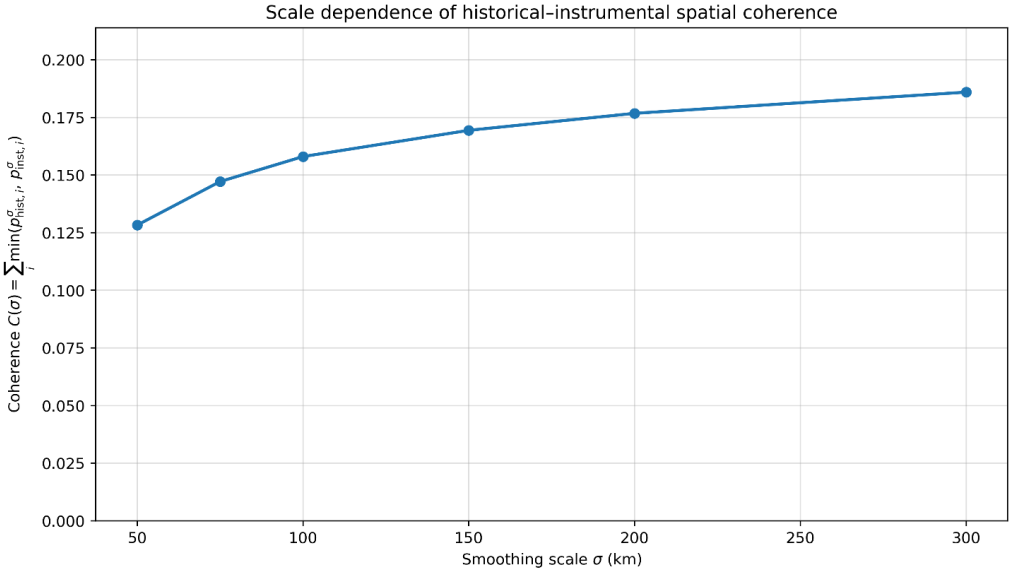


Figure 6. *Scale dependence of historical–instrumental spatial coherence. Scale dependence of the spatial coherence $C(\sigma)$ between historical (1600–1899) and instrumental (1900–1950, $M \geq 5.5$) earthquake reporting densities as a function of the Gaussian smoothing scale σ. For each value of σ, both catalogs are converted into normalized spatial probability density functions on a common global grid using isotropic Gaussian kernel smoothing. Coherence is quantified as*

$$C(\sigma) = \int min\,(p^{\sigma}_{hist}(\boldsymbol{x}), p^{\sigma}_{inst}(\boldsymbol{x}))\, d\boldsymbol{x},$$

which measures the total spatial overlap between the two density fields. The coherence increases monotonically with smoothing scale, indicating that agreement between historical and instrumental spatial patterns emerges progressively as small-scale heterogeneities and reporting noise are averaged out. The absence of saturation at small σhighlights the strong influence of local historical reporting biases, whereas the gradual approach to an asymptotic value at larger scales reflects the recovery of a common large-scale tectonic signal shared by both datasets.

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

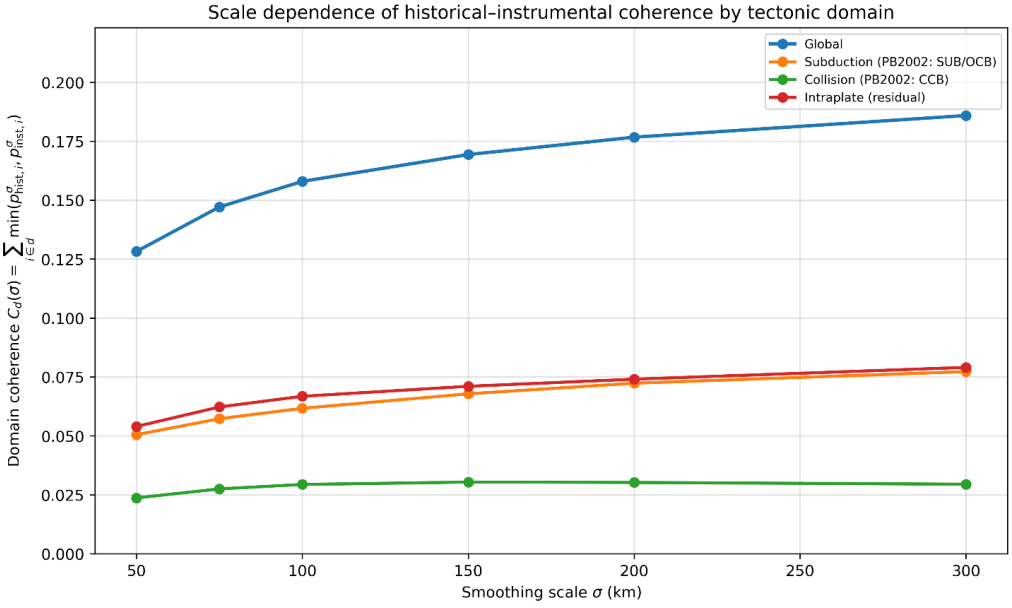


Figure 7. *Scale dependence of historical–instrumental spatial coherence by tectonic domain. Scale dependence of the spatial coherence $C_d(\sigma)$ between historical (1600–1899) and instrumental (1900–1950, $M \geq 5.5$) earthquake reporting densities, evaluated separately for major tectonic domains. For each smoothing scale σ, coherence is defined as*

$$C_d(\sigma) = \sum_{i \in d} min\ (p^{\sigma}_{hist,i}, p^{\sigma}_{inst,i}),$$

where $p^{\sigma}_{hist,i}$ and $p^{\sigma}_{inst,i}$ are the normalized, Gaussian-smoothed reporting densities in grid cell i, and the sum is restricted to cells belonging to domain d. Tectonic domains are defined using the PB2002 global plate-boundary model (Bird, 2003): subduction zones (SUB/OCB), continental collision zones (CCB), and intraplate regions (residual). Grid cells are assigned to a boundary-related domain within a 200 km influence distance from the nearest PB2002 boundary segment; remaining cells are classified as intraplate. The global coherence (blue curve) increases smoothly with σ, reflecting enhanced overlap at larger spatial scales. Subduction and intraplate domains provide the dominant contributions to the total coherence across all scales, whereas collision zones exhibit systematically lower and weakly scale-dependent coherence. The sum of the domain-

https://doi.org/10.5194/egusphere-2026-521
Preprint. Discussion started: 24 March 2026

specific coherences exactly reproduces the global coherence at each scale, confirming internal consistency of the decomposition.