Comparative Analysis of Seismic Clustering: Deterministic Techniques vs. Probabilistic ETAS Model

Giuseppe Falcone [1] Ilaria Spassiani [1] Stefania Gentili [2] Rodolfo Console [3,1] Maura Murru [1] Matteo Taroni [1]

[1]National Institute of Geophysics and Volcanology [2]National Institute of Oceanography and Applied Geophysics - OGS [3]Center of Integrated Geomorphology for the Mediterranean Area

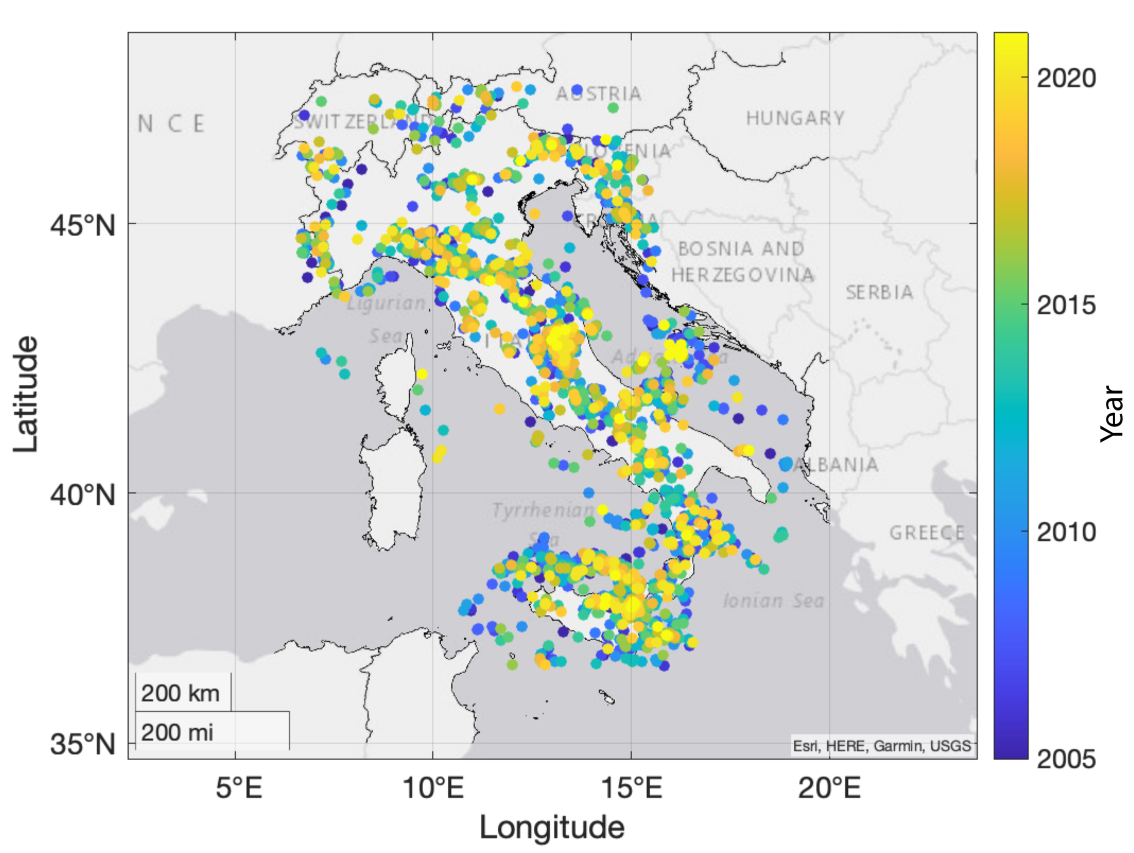
Overview

Short-term seismic clustering, a crucial aspect of seismicity, has been extensively studied in literature. Existing techniques for cluster identification are predominantly deterministic, relying on specific constitutive equations to define spatiotemporal extents. Conversely, probabilistic models, such as the Epidemic Type Aftershock Sequence (ETAS) model, dominate short-term earthquake forecasting. The ETAS model, known for its stochastic nature, has been employed to decluster earthquake catalogs probabilistically. However, the challenge arises when selecting a probability threshold for cluster identification, potentially distorting the model's underlying hypothesis. This study aims to assess the consistency between seismic clusters identified by deterministic window-based techniques specifically, Gardner-Knopoff and Uhrhammer-Lolli-Gasperini and the associated probabilities predicted by the ETAS model for events within these clusters.

The earthquake catalog

Italian seismic catalog, ISIDe (Italian Seismological Instrumental and Parametric Data-BaseE, `http://terremoti.ingv.it/ISIDe`).

- **Temporal interval** 2005 April 18 - 2021 April 30
- **Min and Max magnitude:** ML 0.9 and ML 6.1
- **Completeness magnitude [5]:** $m_0 = 2.9$ (5084 events above completeness)



Cluster identifications methods and ETAS rate

Window-based methods (NESTOREv1.0 [2]). Min mainshock mag $M_m = 4$; the cluster of a mainshock m_i is *the set of earthquakes after m_i within its triggering circular area*. (Foreshocks excluded to avoid possible multiple assignments)

1. Uhrammer-Lolli-Gasperini **(ULG) [3]-[4]:** $d = e^{0.804 \cdot M_m - 1.024}$, $t = 60 + 60(M_m - 4)$
2. Gardner-Knopoff **(GK) [1]:** $d = 10^{0.1238 \cdot M_m + 0.983}$, $t = 10^{A \cdot M_m + B}$, where $(A, B) = (0.032, 2.7389)$ or $(0.5409, -0.547)$ if $M_m \geq 6.5$ or $M_m < 6.5$.

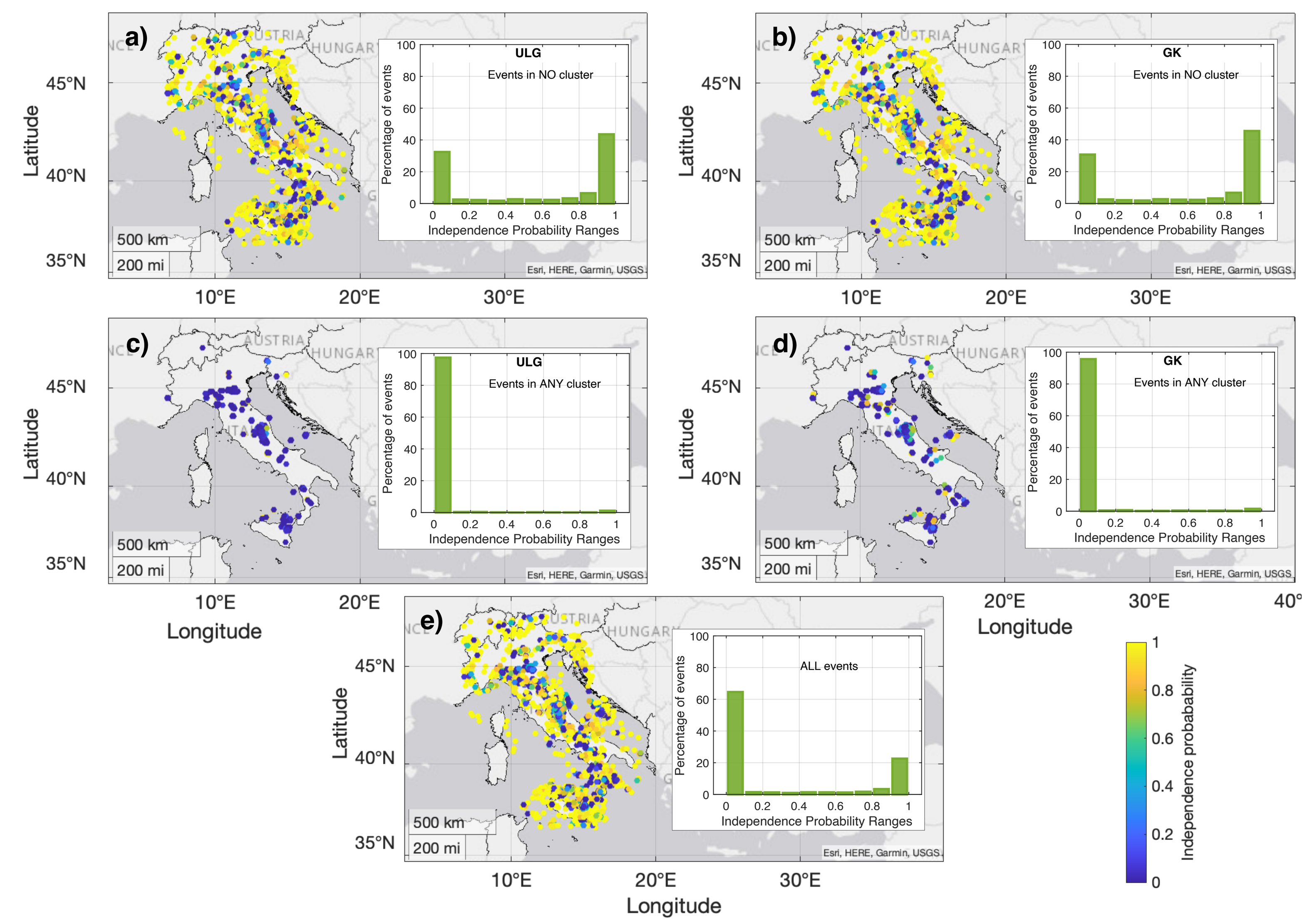
Spatio-temporal ETAS model. The rate $\lambda(t, x, y, m)$ is given by

$$\beta e^{-\beta(m-m_0)} \left\{ f_r \, \lambda_0(x,y) + \sum_{i=1}^{N} H(t-t_i) K \left[\frac{d_i^2}{(x-x_i)^2 + (y-y_i)^2 + d_i^2} \right]^q (t - t_i + c)^{-p} \right\},$$

where f_r is the failure rate, $\beta = b \ln 10$, $d_i = d_0 10^{\alpha(m_i - m_0)/2}$ and $H(\cdot)$ is the step function.

ULG- and GK-clusters; comparison with ETAS independence $\mathbb{P}$

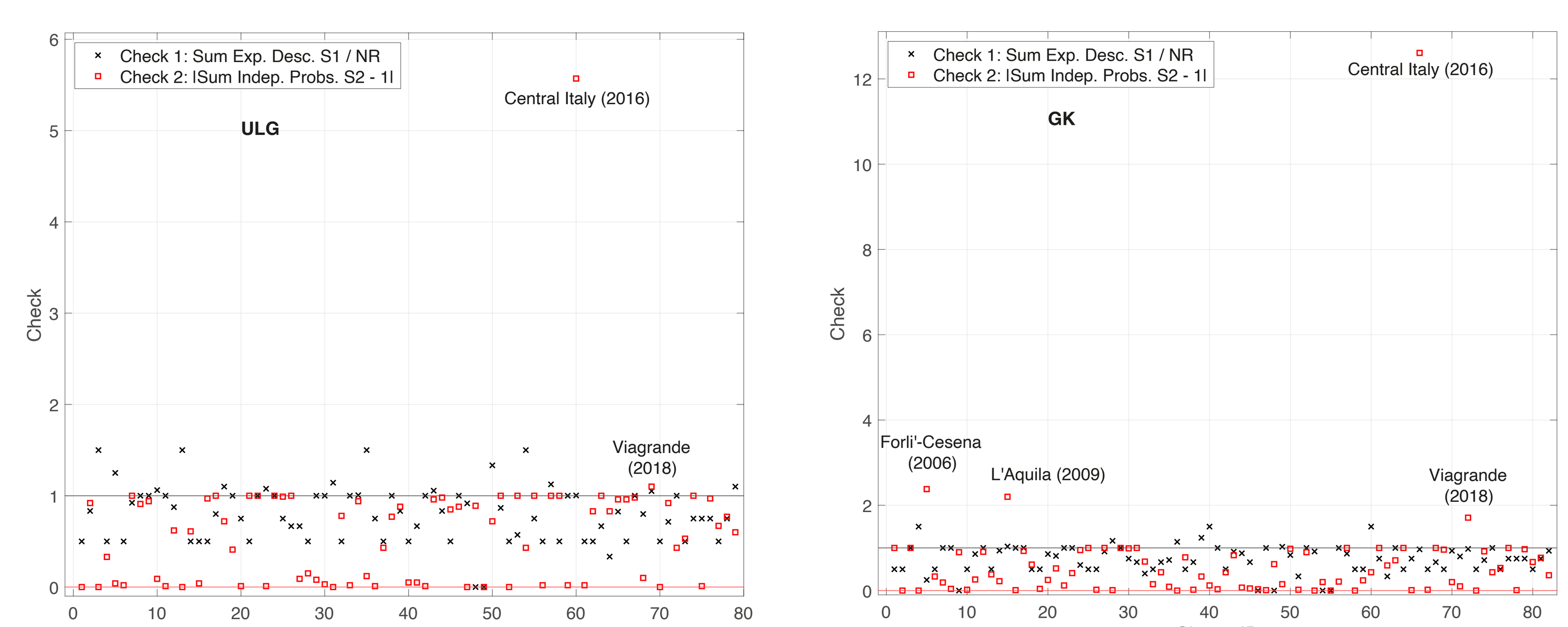
The procedure identified 79 ULG- and 82 GK-clusters. Clusterized events are much less sparse than the others, and the great majority has a very low $\mathbb{P}$ of being independent (< 0.1), with the exception of a few of them (reasonably, the mainshocks).

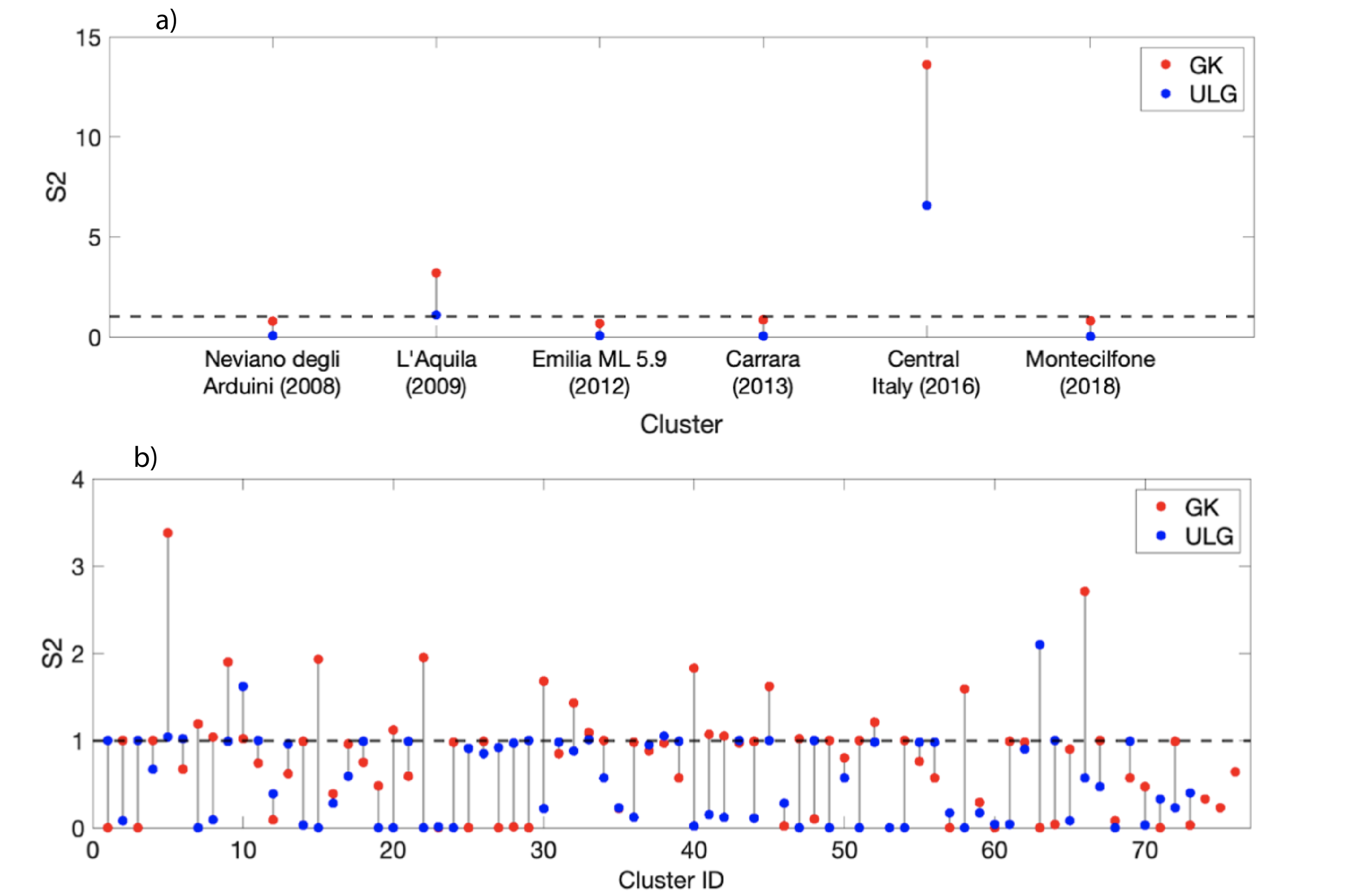


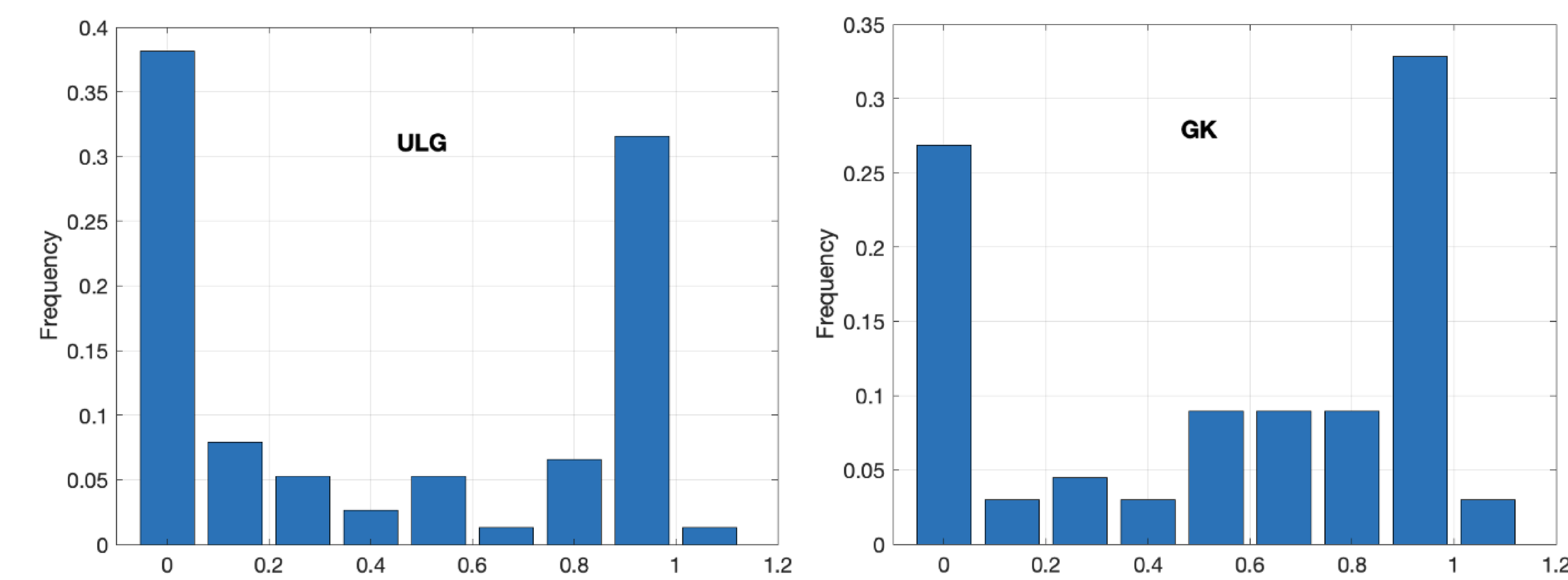
Deterministic VS probabilistic approaches: consistency checks

We assume that the n^{th} cluster (nCL) contains NR events. The nCL cluster is considered consistent with the ETAS model if:

- CHECK 1. *The sum S1 of the expected numbers of events triggered by the NR events in the nCL cluster is close to the number of elements in nCL:* CHECK1 = S1/NR $\sim$ 1. This is because the expected offspring in nCL should reflect its cardinality.
  If S1>NR, we likely have an "over-productive" seismic sequence included in the cluster.
- CHECK 2. *The sum S2 of the independence $\mathbb{P}$ of all the NR events in the nCL cluster is close to 1:* CHECK2 = |S2-1| $\sim$ 0. This is because we expect a "single" cluster to have a "single" independent event.
  If S2 > 1, nCL contains >1 (ETAS-)independent events, and therefore it likely involves strong seismic sequences.







Take home messages

- **No substantial differences for the 2 cluster identification procedures:**
  - GK identifies quite longer and quite wider clusters when considering strong mainshocks;
  - comparable cardinality and mainshocks of the identified clusters.
- **Overall consistency between identified clusters and corresponding ETAS $\mathbb{P}$:**
  - window-based methods need sharp cutoff to include events in a cluster, and require the labeling of events as "mainshock";
  - independence $\mathbb{P}$ tends to increase in the final part of a sequence;
  - swarm-type sequences better captured by a probabilistic approach such ETAS.

Deterministic and probabilistic approaches allow us to consider different perspectives and aspects of seismicity. The probabilistic view is less subjective (no cutoff required). Still, $\mathbb{P}$ is a challenging concept to understand, and carries a certain degree of uncertainty.

No general rule for one approach being preferable to the other, but, be aware of the meaning behind the selected approach and the implications it entails in the results.

References

[1] Gardner and Knopoff. https://doi.org/10.1785/BSSA0640051363. *Bull. Seismol. Soc. Am.*, 64(5):1363–1367, 1974.
[2] Gentili, Brondi, and Di Giovambattista. https://doi.org/10.1785/0220220327. *Seismol. Res. Lett.*, 94(4):2003–2013, 2023.
[3] Lolli and Gasperini. https://doi.org/10.1023/A:1023588007122. *J. Seismol.*, 7:235–257, 2003.
[4] Uhrhammer. Characteristics of northern and central California seismicity. *Earthquake Notes*, 57(1):21, 1986.
[5] Zhuang, Murru, Falcone, and Guo. https://doi.org/10.1093/gji/ggy428. *Geophys. J. Int.*, 216(1):302–318, 2019.