



ST-DBSCAN vs Window-Based Methods: A Comparative Cluster Analysis of the New Zealand Earthquake Catalog

Ester Piegari¹, Stefania Gentili², and Letizia Caravella²

¹University of Naples Federico II, Dipartimento di Scienze della Terra, dell'Ambiente e delle Risorse, Napoli, Italy
(esterpiegari@gmail.com)

²National Institute of Oceanography and Applied Geophysics - OGS, Italy

Seismic catalogs combine background seismicity driven by tectonic loading with clustered earthquakes that reveal stress transfer and fault interactions. Both components are essential for seismic hazard models, which require accurate declustering.

Traditionally, declustering relies on window-based methods widely used in operational seismology for their simplicity and real-time efficiency. However, these methods suffer from rigid geometric constraints, depend on mainshock identification, and are highly sensitive to parameter choices, which may lead to over- or underestimation of earthquake cluster size. Machine learning-based approaches can mitigate these limitations by adapting flexibly to data patterns without rigid geometric or mainshock assumptions. Density-based algorithms such as DBSCAN and OPTICS identify spatial clusters effectively but struggle with spatiotemporal aftershock sequences because they treat time independently from space. ST-DBSCAN addresses this by using separate spatial and temporal radii, enabling flexible space-time clustering critical for aftershock analysis.

In this comparative study, we applied both approaches – ST-DBSCAN and window-based methods – to the New Zealand earthquake catalog to highlight the strengths and limitations of each, analyzing 15 overlapping clusters (>100 events, centroids <10 km apart). We found that ST-DBSCAN better captures fine-scale structures, whereas window-based methods produce more compact large-scale groupings. We analyze in detail the 2010–2013 Canterbury–Christchurch sequence, validating cluster membership against an independent dataset of approximately 150 earthquakes ($M_w > 3.5$), which reveals methodological differences in spatiotemporal resolution.