

GEM Global Homogenised Earthquake Catalogue

v2026

Kirsty Bayliss

GEM Hazard Team

June 23 – 25, 2026
University of Zagreb, Croatia

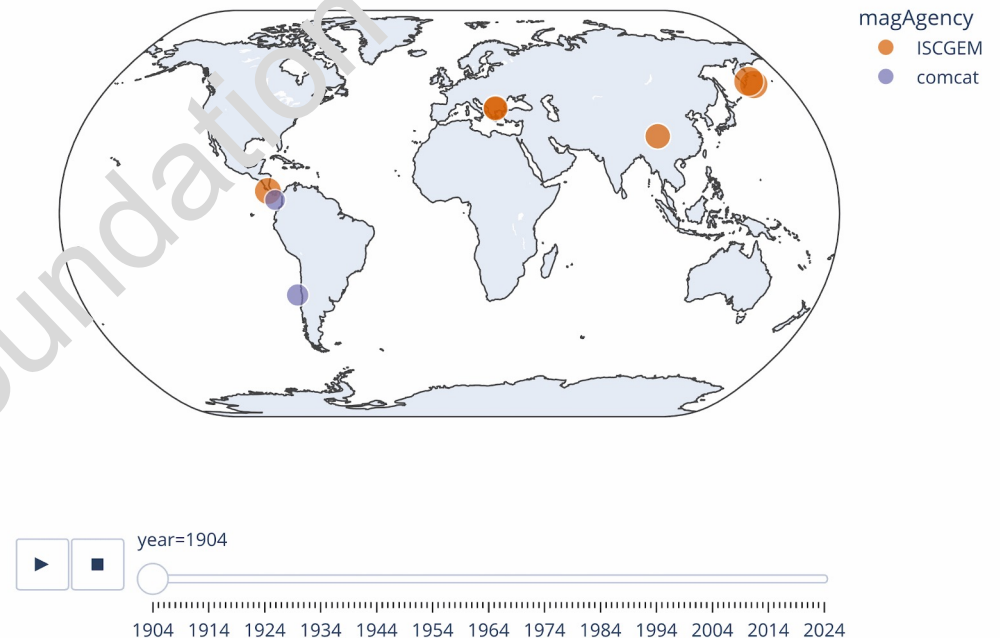
GEM CONFERENCE
FROM FAULTS TO FUTURE SCENARIOS

Motivation

- Catalogues are a fundamental dataset for PSHA
- Need reliable catalogues capturing rates of earthquakes
- Requires combining multiple catalogues into single 'homogenised' catalogue with single magnitude type (M_w)
- Global 'backbone' for developing regional catalogues
- ISC-GEM for onshore, but we wanted to extend to smaller magnitudes and enhance offshore coverage

Merging and homogenisation process

- Merge three global catalogues: ISC-GEM, Global centroid moment tensor (GCMT), comcat (from USGS, M4+)
- GCMT provides newer events (not yet reviewed by ISC)
- comcat provides smaller and recent events with varying magnitude types
- Use tools in the OQ-MBTK to add unique events
- prioritise ISC-GEM > GCMT > comcat



Event locations and duplications

gcmt locations

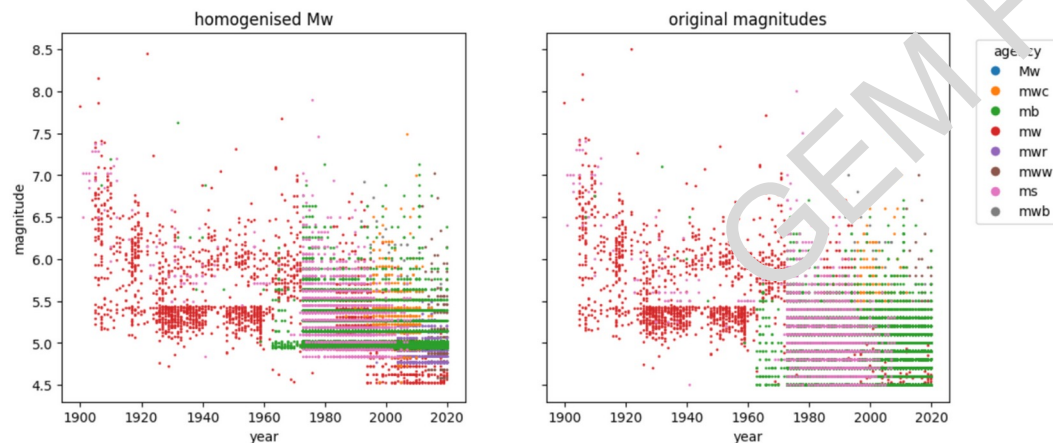
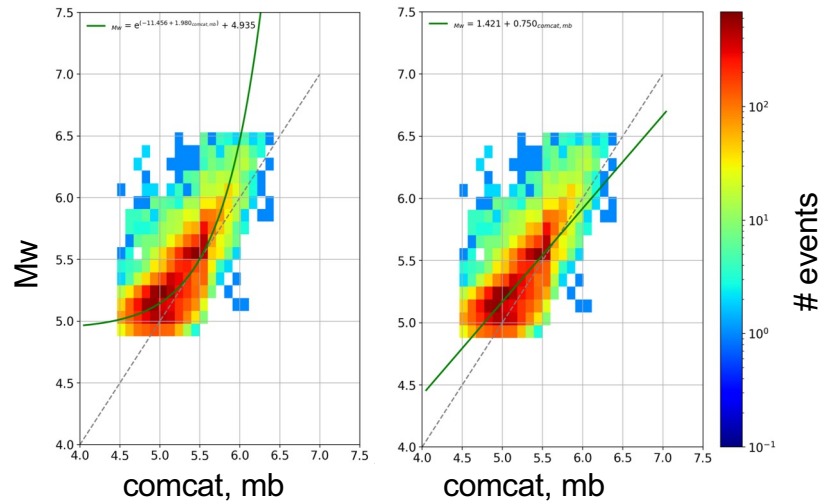
- GCMT inversion results in time and locations reflecting the 'centre' of rupture, compared to locations reported in traditional catalogues
- for large earthquakes this can be up to ~100km (and several minutes!) difference

Location differences between catalogues, and between catalogue revisions

- for events <1960 sometimes > 100km between catalogues
- Improved locations in revised catalogues can lead to duplications - some of these will be removed in declustering



Merging and homogenisation process



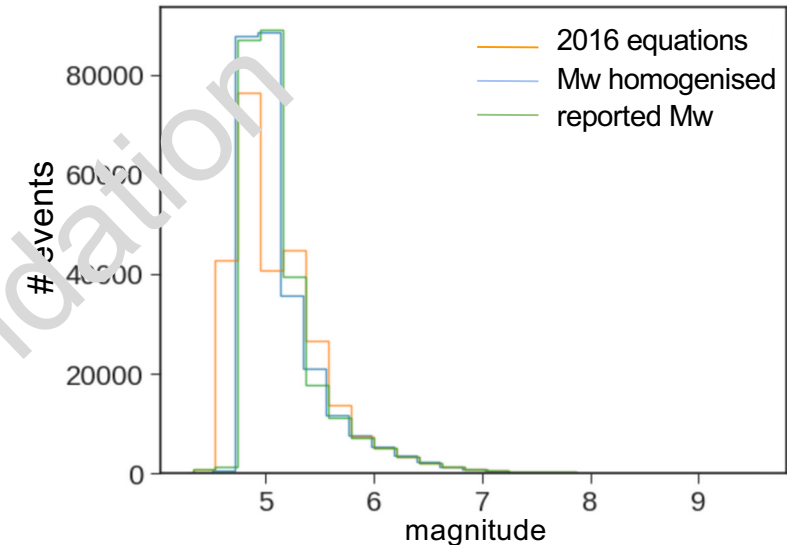
comcat events require homogenisation for mb, ms and various Mw magnitudes - need to determine suitable conversions

mb magnitude conversions are challenging

Even conversions that seem sensible can lead to issues in the final catalogue!

Magnitude homogenisation and uncertainties

- 2016 homogenisation equations for mb results in spikes in higher magnitude bins, reduction in ~M5 events (orange)
- used a back-projection method to test impact of homogenisation choices
- Global catalogue between 5.5 -7 is ~stable GR, so how do different magnitude choices affect the b-value?
 - Assumes GR scaling at $M < 5.5$, does not consider variable completeness



- Propagated uncertainty from original catalogue + uncertainty from homogenisation included

$$\text{Comcat mw: } 0.111 + 0.9806 * m$$

$$\text{Comcat mww: } 0.194 + 0.975 * m$$

$$\text{Comcat mwb: } 0.373 + 0.935 * m$$

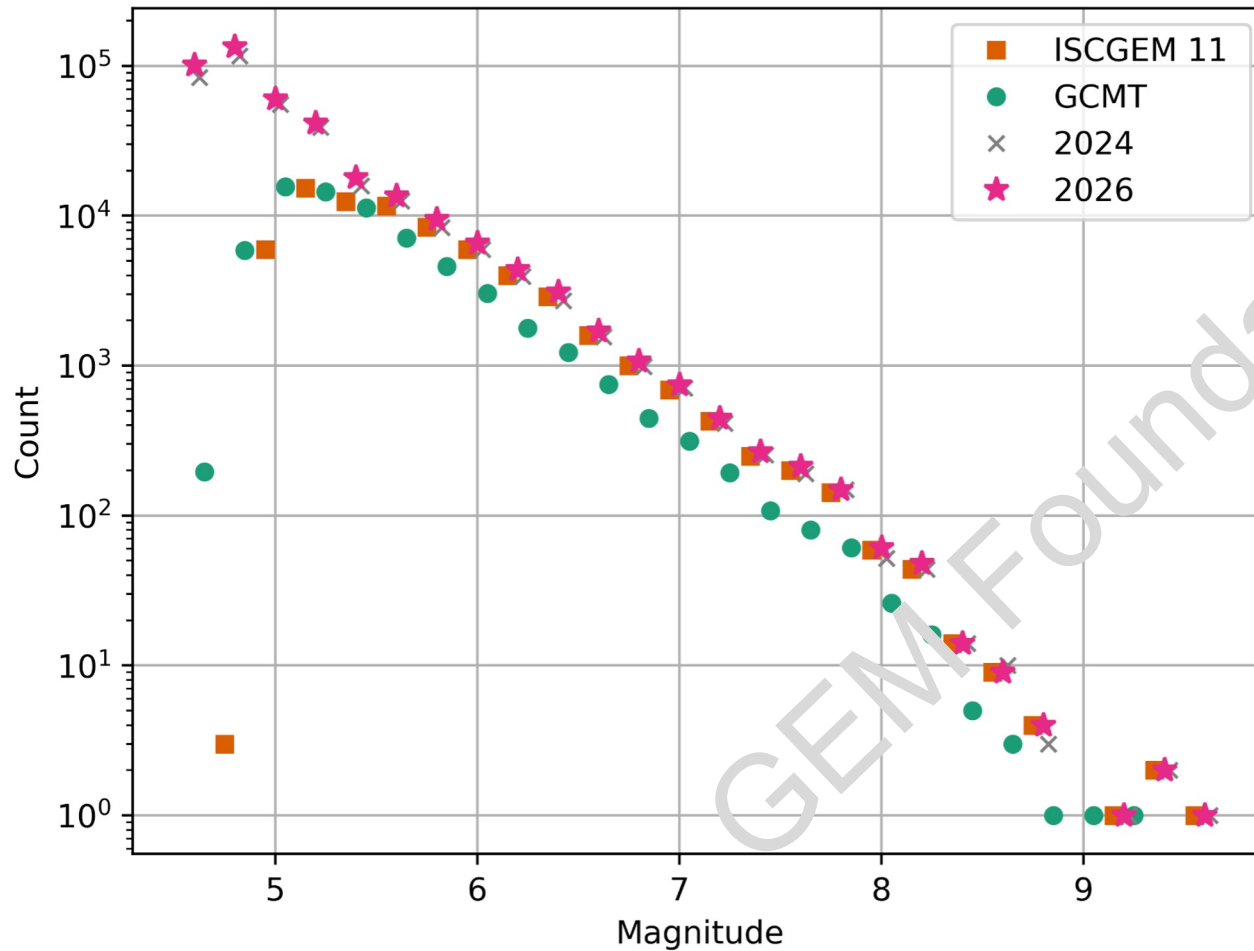
$$\text{Comcat mwr: } 1.504 + 0.725 * m$$

$$\text{Comcat mwc: } 0.099 + 0.985 * m$$

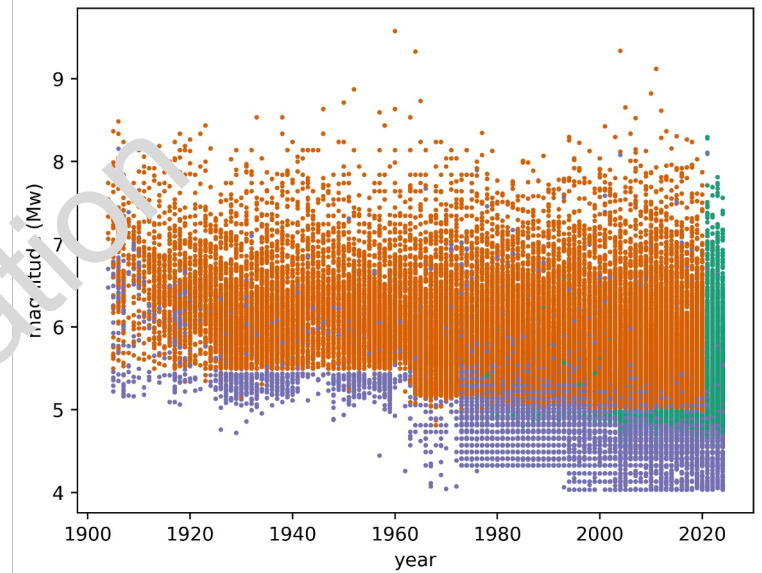
$$\text{Comcat ms: } 0.902 + 0.874 * m$$

$$\text{Comcat mb: } 1.091 + 0.809 * m \text{ where } m < 6.5$$

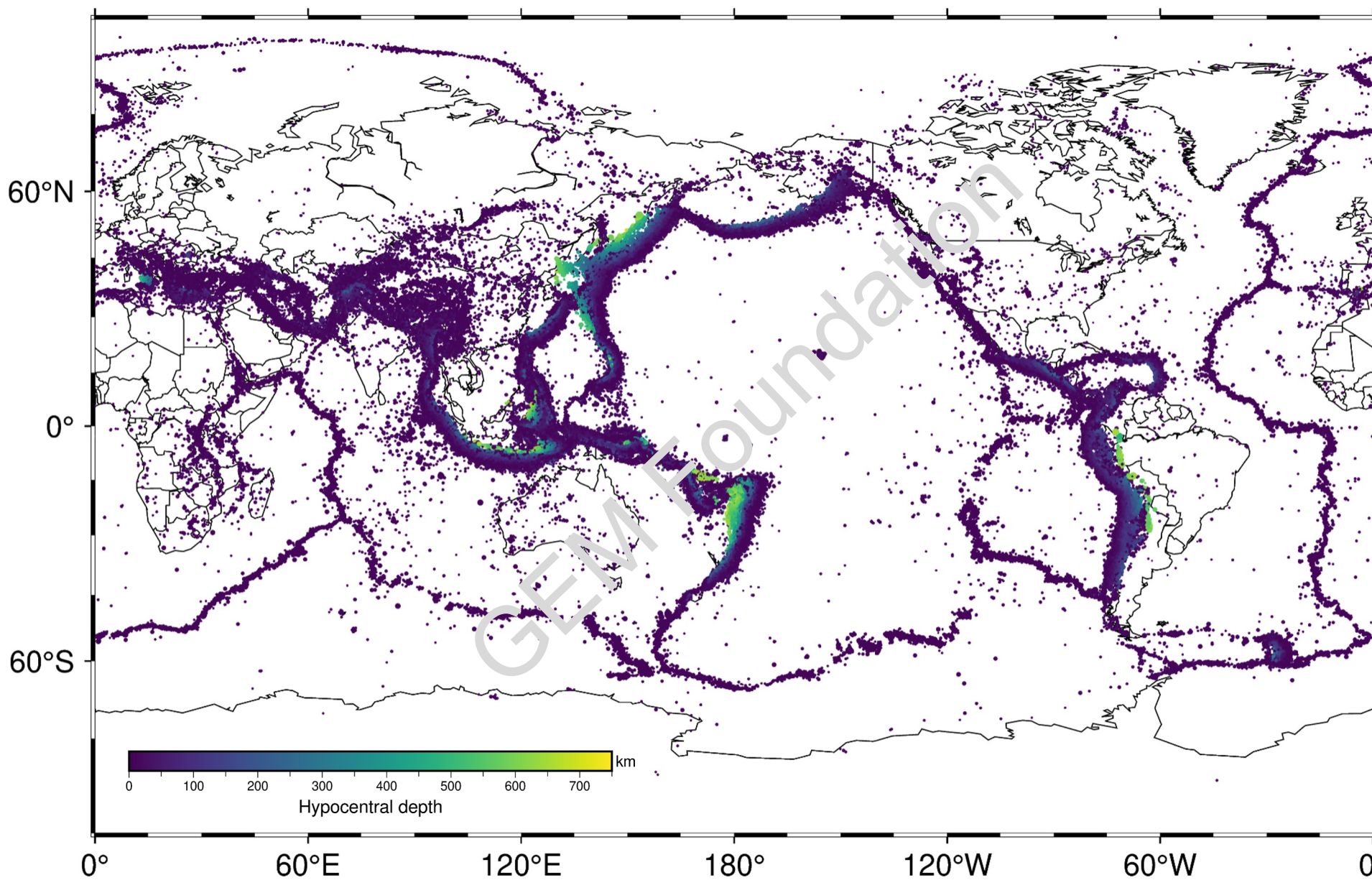
Global magnitude-frequency distribution



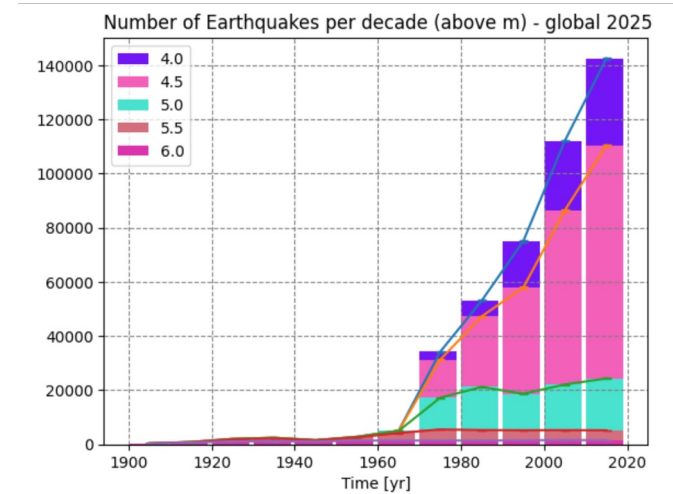
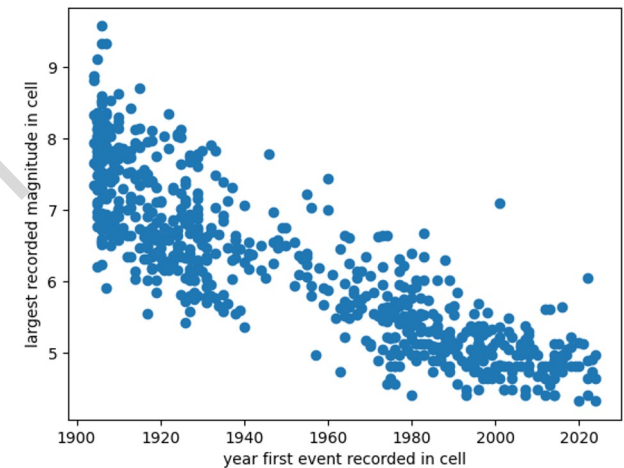
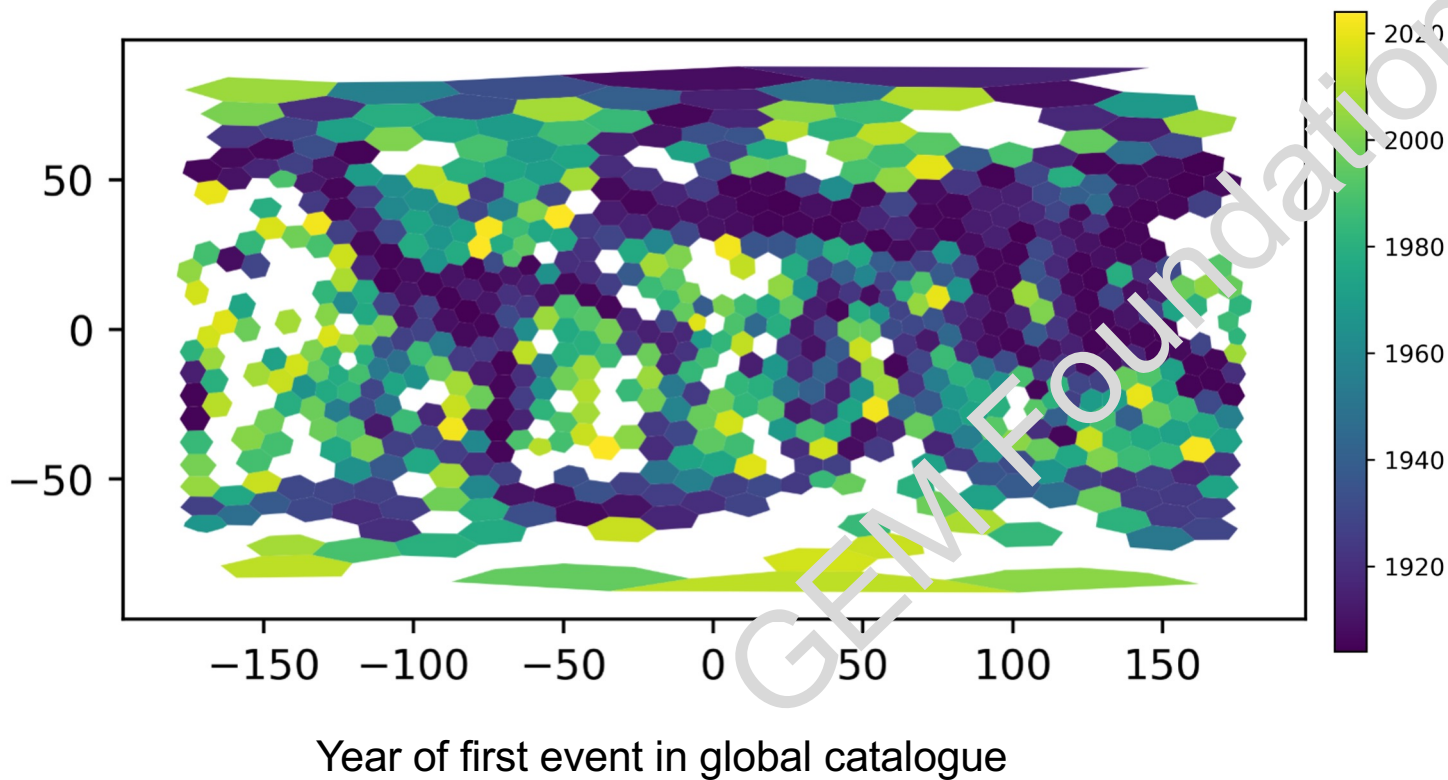
*Note slight jitter to bins



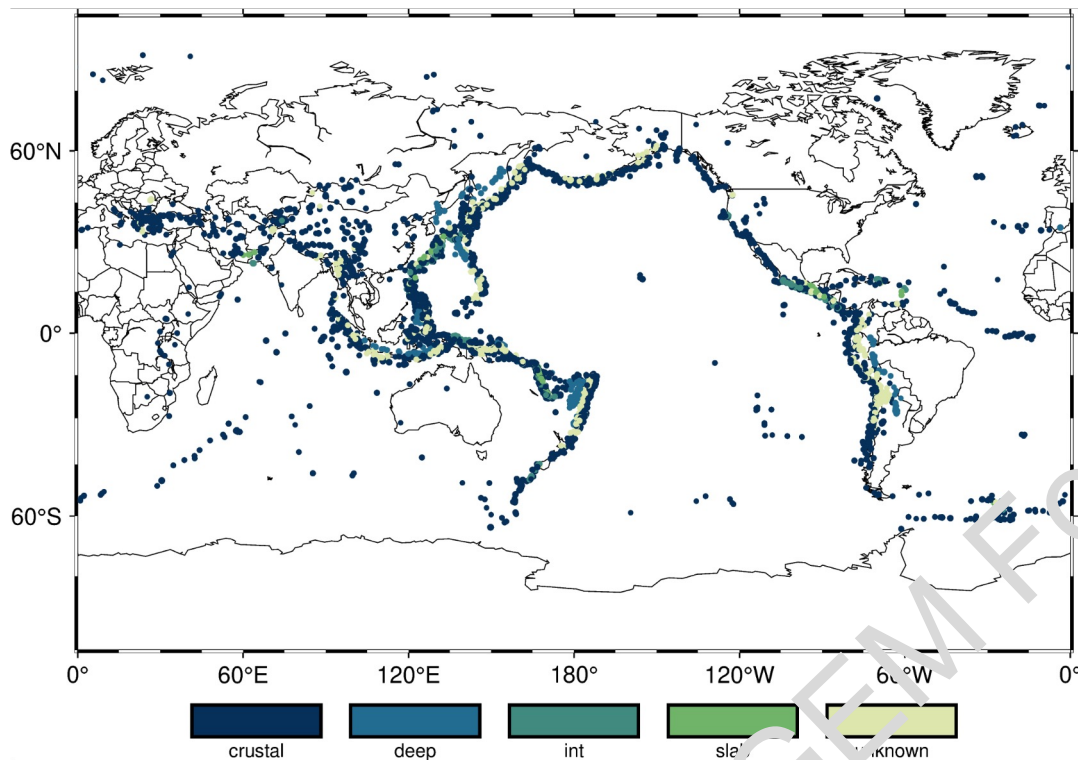
~495 000 events from 1904-2024



Towards global completeness



Planned improvements: Classification

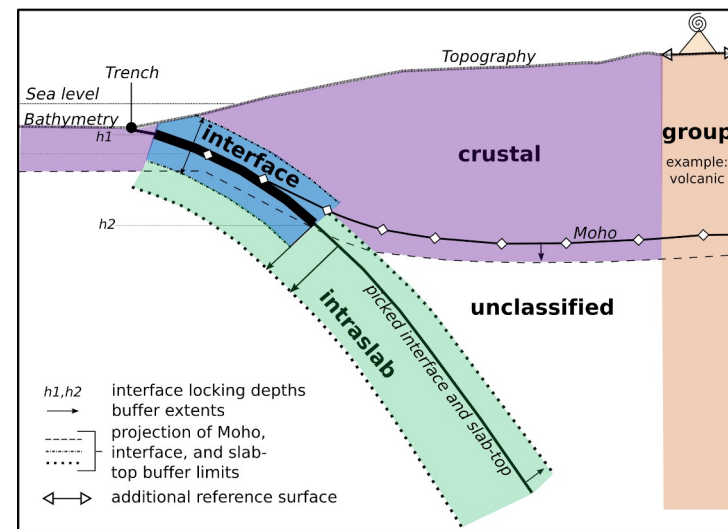


Preliminary global classification for $M_w 5.5$ events, from GEESE

Next step for the global catalogue: global classification by tectonic region

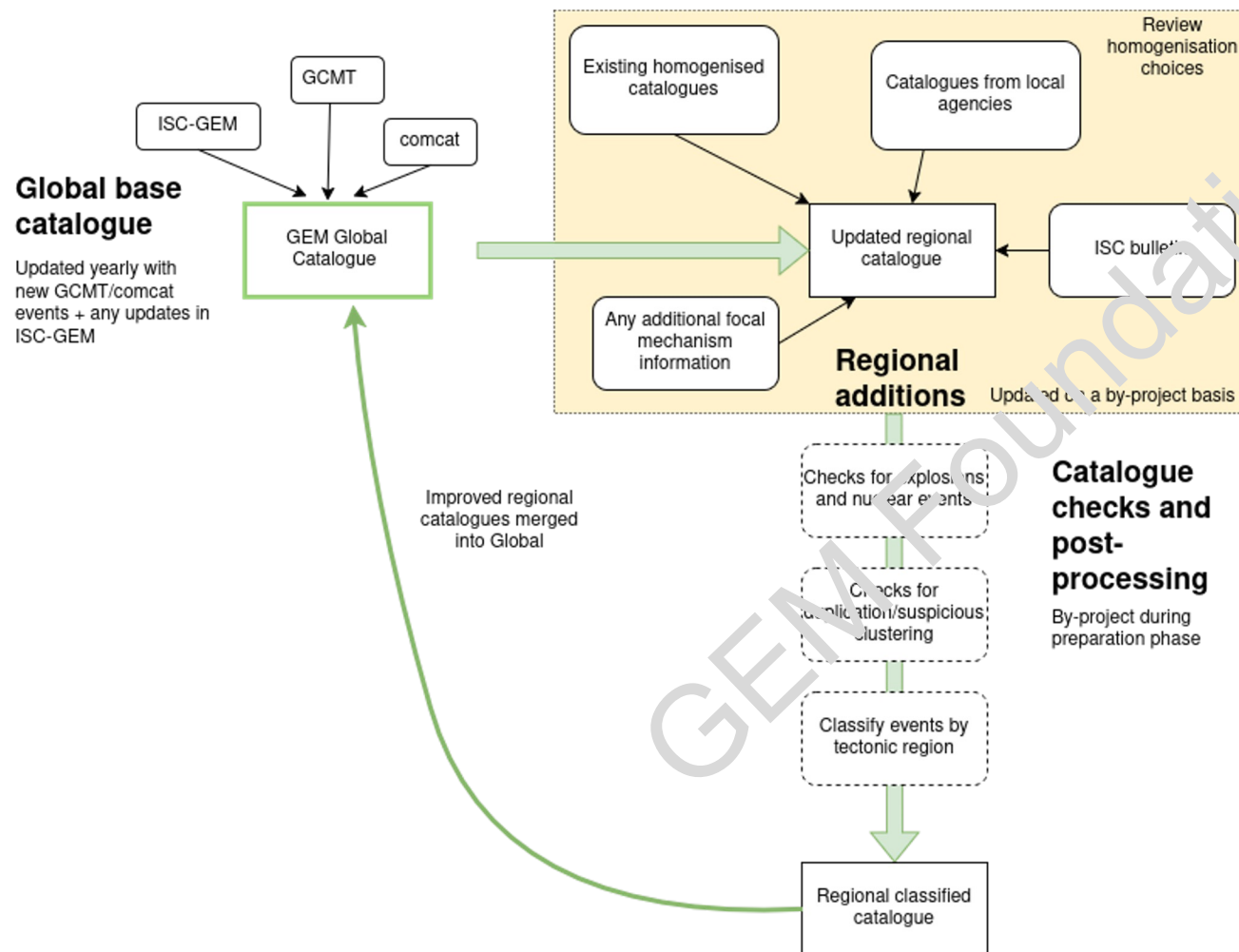
Subduction models require classification of events into crustal, interface and slab components.

Aim to compile 3D subduction geometries to allow global event classification

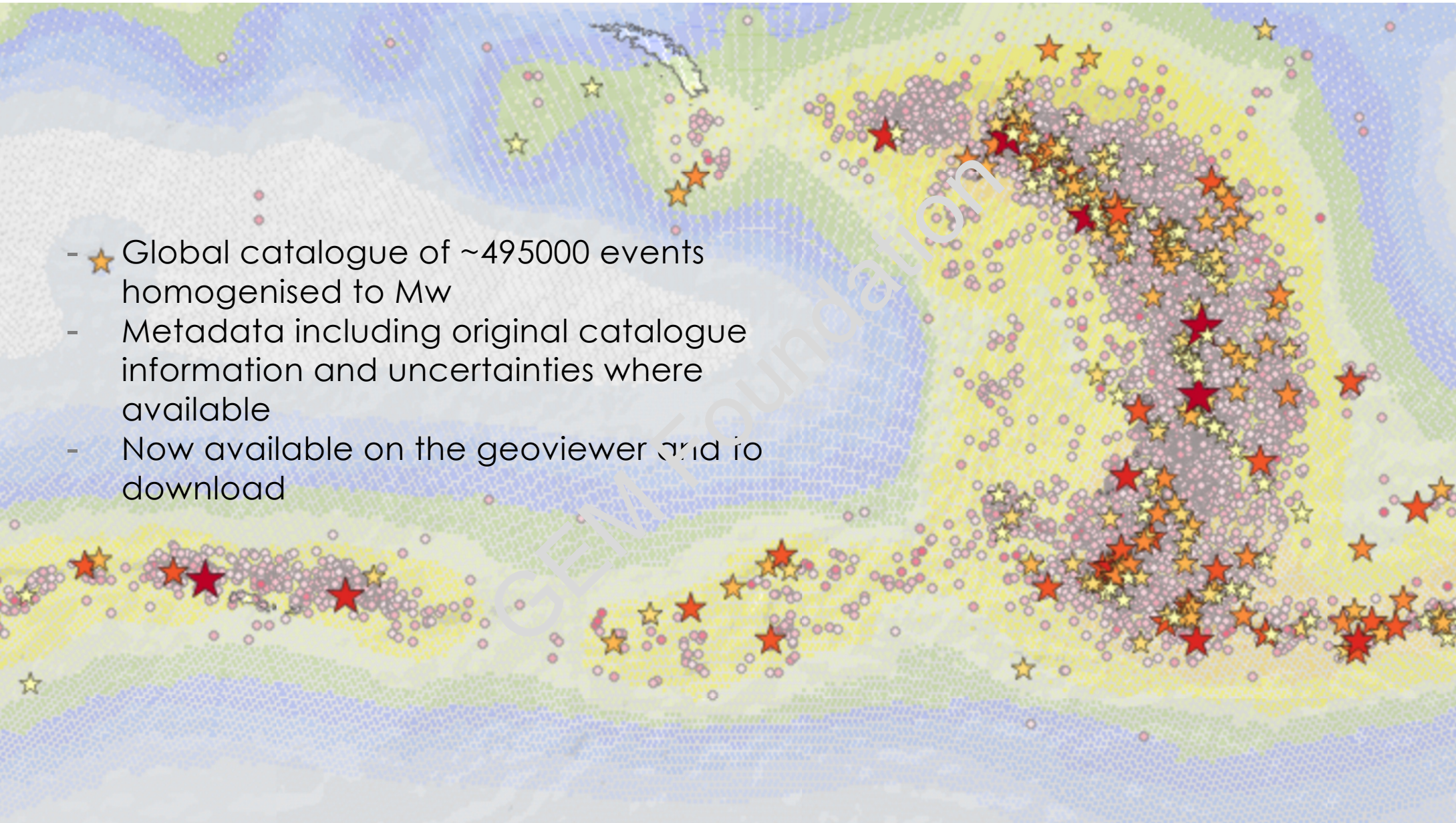


From Pagani et al., 2020

Future plans



- Global catalogue to be updated regularly (~once/year)
- Higher resolution local catalogues to be merged into global as far as possible
- Improved tools for merging homogenised catalogues, keeping track of sources and uncertainties as far as possible
- Possible further improvements, including more focal mechanism data and more tools for checking for spurious events/duplications

- 
- A world map with a color-coded background representing different geological or tectonic regions in shades of blue, green, yellow, and grey. The map is densely populated with symbols representing seismic events. These symbols include small pink circles, larger yellow stars, and prominent red stars. The highest concentration of events is in the western Pacific Ocean, particularly around Japan and the Philippines, where red stars are most frequent. Other significant clusters are visible along the Mid-Atlantic Ridge and the western coast of the Americas. A faint, diagonal watermark reading 'GEM Foundation' is overlaid across the center of the map.
- ★ Global catalogue of ~495000 events homogenised to Mw
 - Metadata including original catalogue information and uncertainties where available
 - Now available on the geoviewer and to download

Thank you!

Please attribute to the GEM Foundation with a link to:
<https://www.globalquakemodel.org>



Except where otherwise noted, this work is licensed under a
Creative Commons Attribution-NonCommercial-NoDerivatives 4.0
International License (CC BY-NC-ND 4.0)
<https://creativecommons.org/licenses/by-nc-nd/4.0/>



June 23 – 25, 2026

University of Zagreb, Croatia

GEM CONFERENCE

FROM FAULTS TO FUTURE SCENARIOS



GEM Global Block Model

Richard Styron

GEM Hazard Team

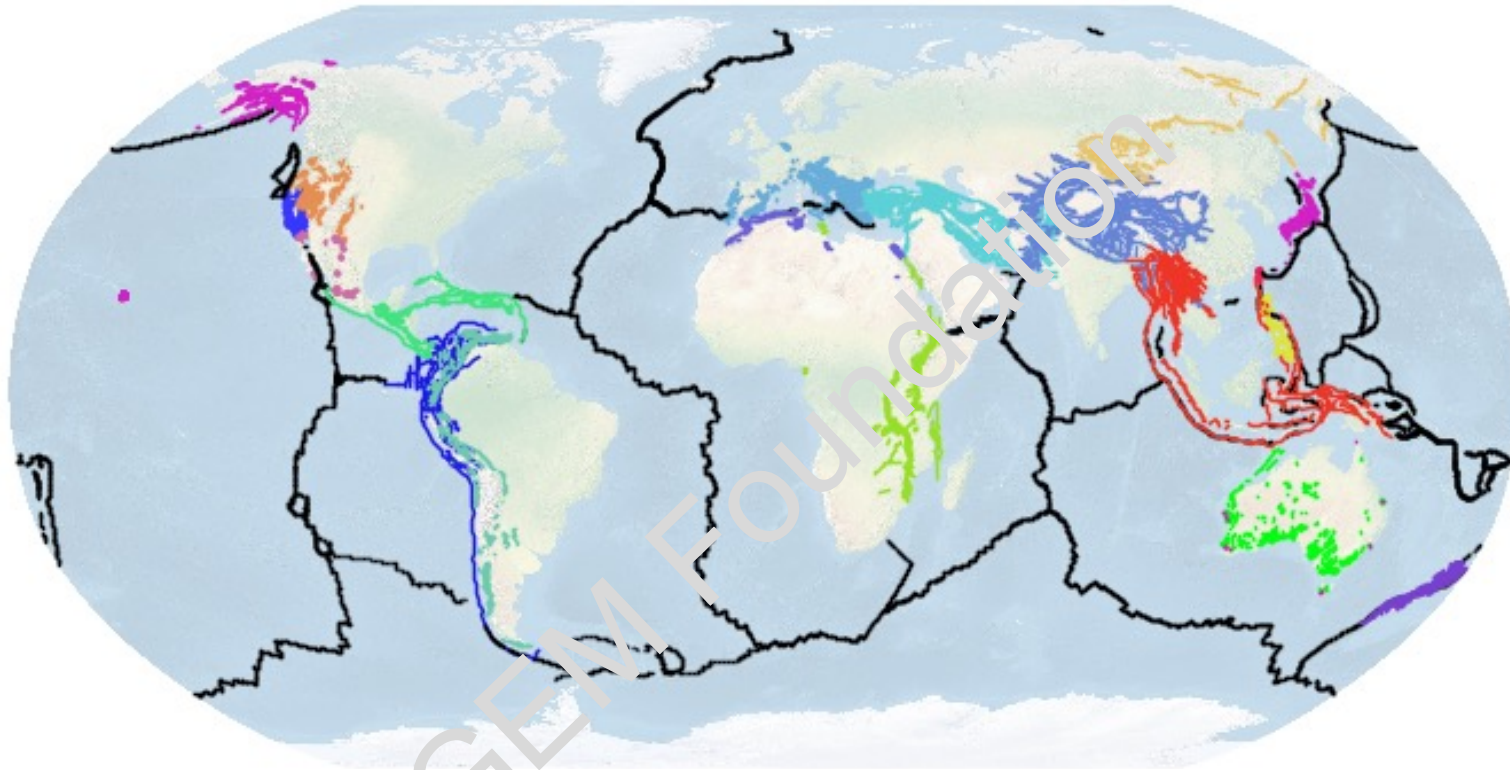
June 23 – 25, 2026
University of Zagreb, Croatia

GEM CONFERENCE
FROM FAULTS TO FUTURE SCENARIOS

Use of faults in PSH(R)A

- Most hazardous earthquakes occur on mapped (or mappable) faults
- Fault sources in PSHA provide much more accurate rupture location, geometry, and kinematics compared to distributed seismic sources
- Fault slip rates provide budget for earthquake occurrence
- Fault catalogs have different limitations than earthquake catalogs

GEM Global Active Faults Database (GEM GAF-DB), 2020 release

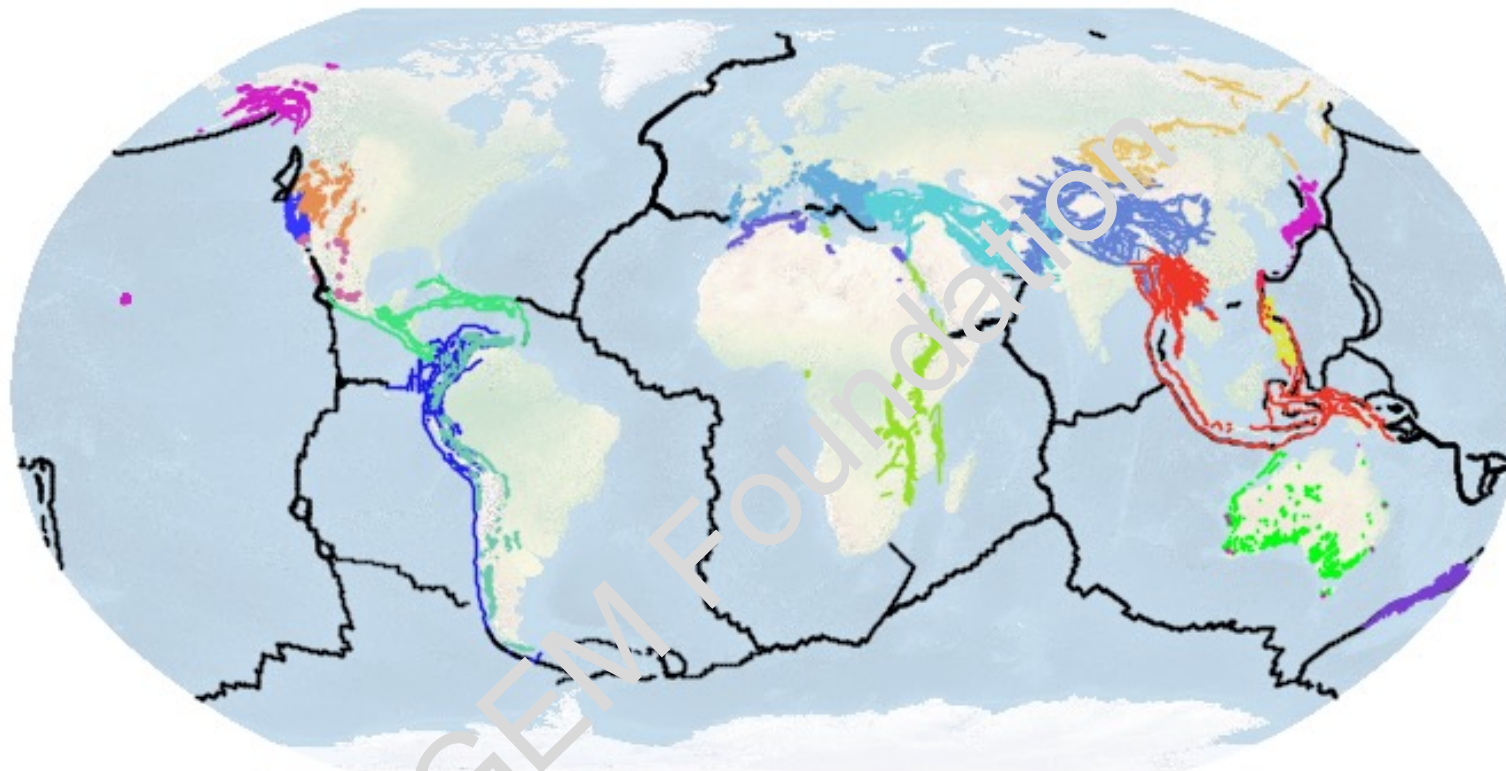


- First known global compilation
- Mosaic of ~20 regional or thematic databases
 - Some for PSHA, some not

Styron and Pagani, 2020



GEM Global Active Faults Database (GEM GAF-DB), 2020 release



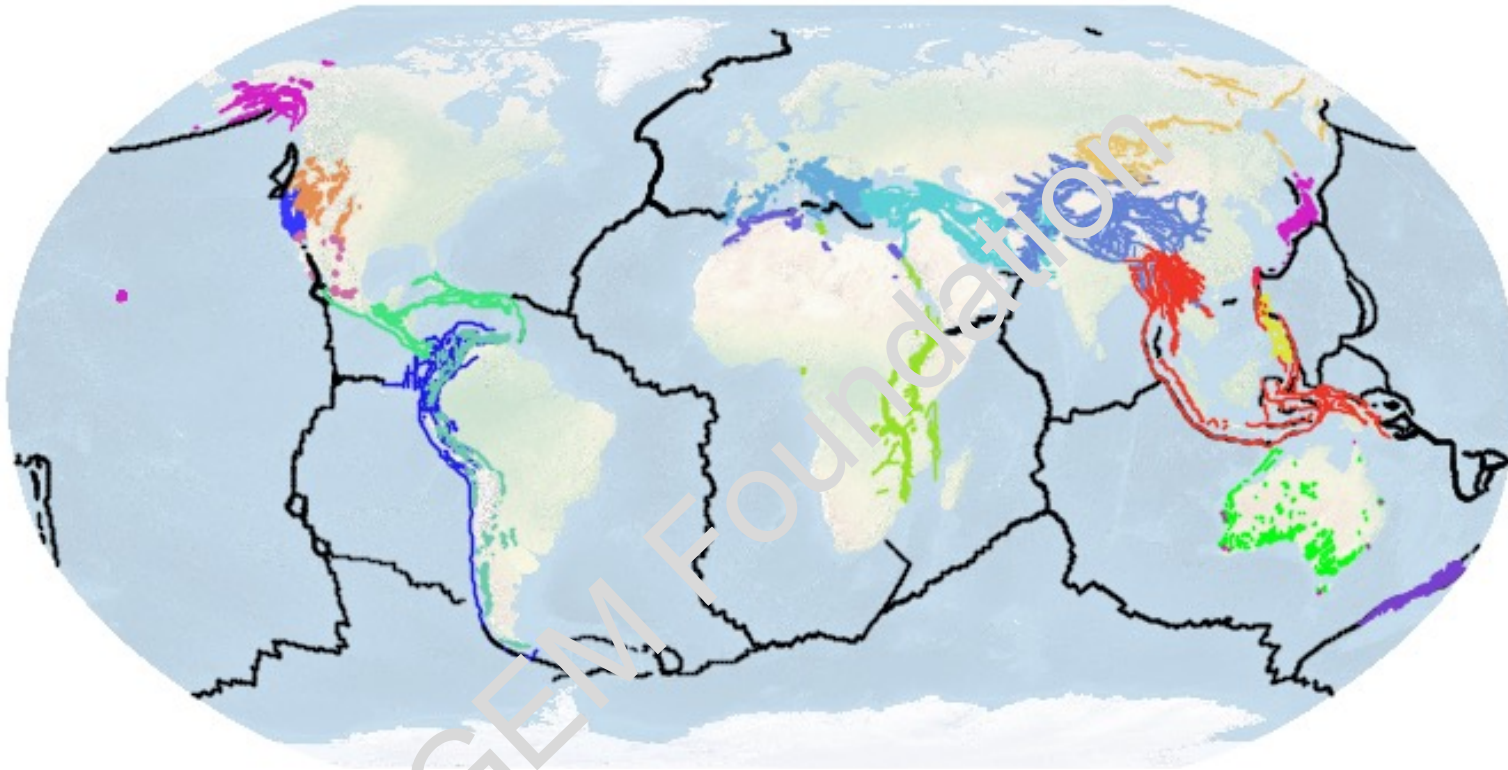
GEM GAF-DB Fault Sources

- | | | |
|---------------------------------|--------------------------|-----------------------|
| — Active Tectonics of the Andes | — GEM N. Africa | — SHARE |
| — Bird Plate Boundaries | — GEM N.E. Asia | — Shyu Taiwan |
| — EMME | — HimaTibetMap | — USGS Hazfaults 2014 |
| — EOS SE Asia | — Litchfield NZ 2013 | — Villegas Mexico |
| — GEM Faulted Earth | — Macgregor AfricaFaults | — AUS_FSD |
| — GEM Carib Central Am | — PHIVOLCS | — UCERF3 |
| | — SARA | |

Styron and Pagani, 2020



GEM Global Active Faults Database (GEM GAF-DB), 2020 release



- No slip rates
- Fault geometries not always best for PSHA



Improving fault data to advance PSHRA

Accuracy

- Improve map resolution of traces
- Traces represent seismic sources, not surface deformation
- More accurate slip rate estimates (*no more guessimates*)

Precision

- Lower slip rate uncertainties
- Correlated slip rate uncertainties
- Fewer branches/samples to explore parameter space

Clarity

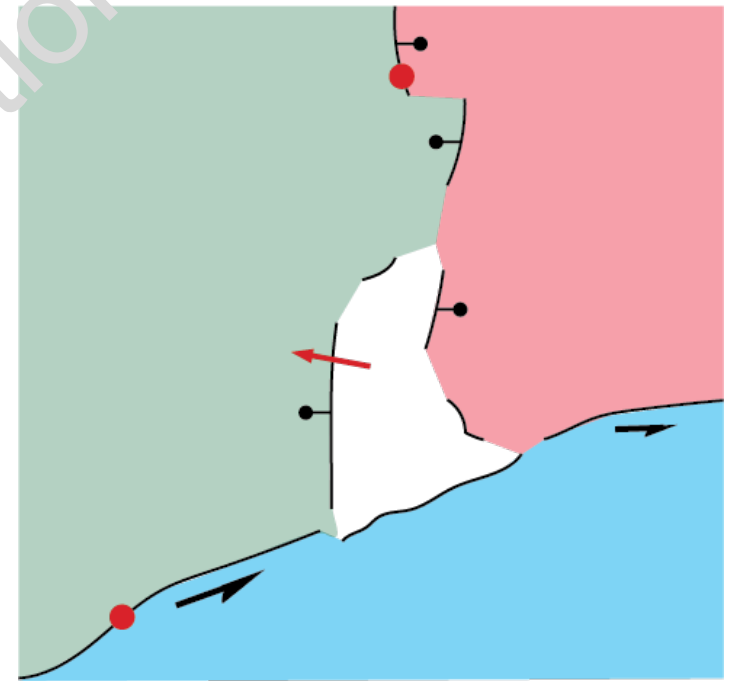
- Clearer links between model parameters and data
- Simpler models with fewer logic tree branches

How? Remap everything and link all available data and theory (systematize it)!



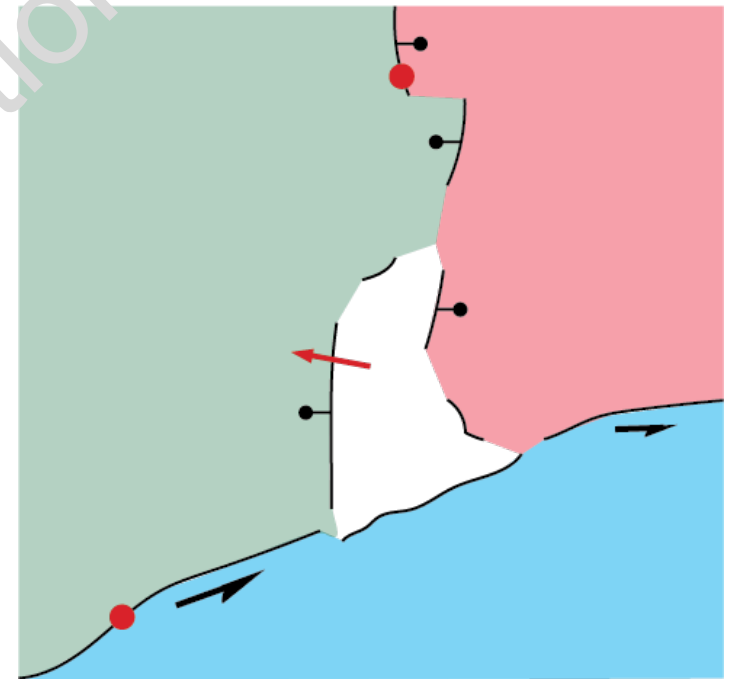
Block models: systematization of Earth deformation data

- Based on plate tectonic theory and math
- Blocks are spatial polygons bounded by faults (and off-fault boundaries)
- Block motion completely described by 3 numbers (3-vector, Euler pole)
- Any combination of geodetic data (block motion compared to reference frame) and geologic slip rate data (relative motion of adjacent blocks) can solve system



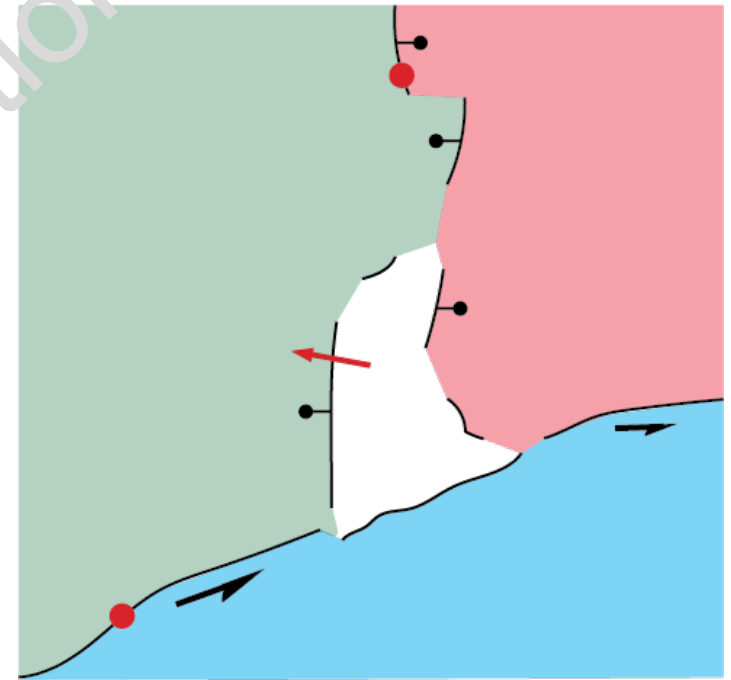
Block models and fault slip rates

- All fault slip rates determined by relative block motions
- *Precision* of fault slip rate estimates greatly increased by integration of many data + constraints
 - Accuracy probably is as well
- Slip rate variance and covariance from data variance and block geometry



Block model data and inversion

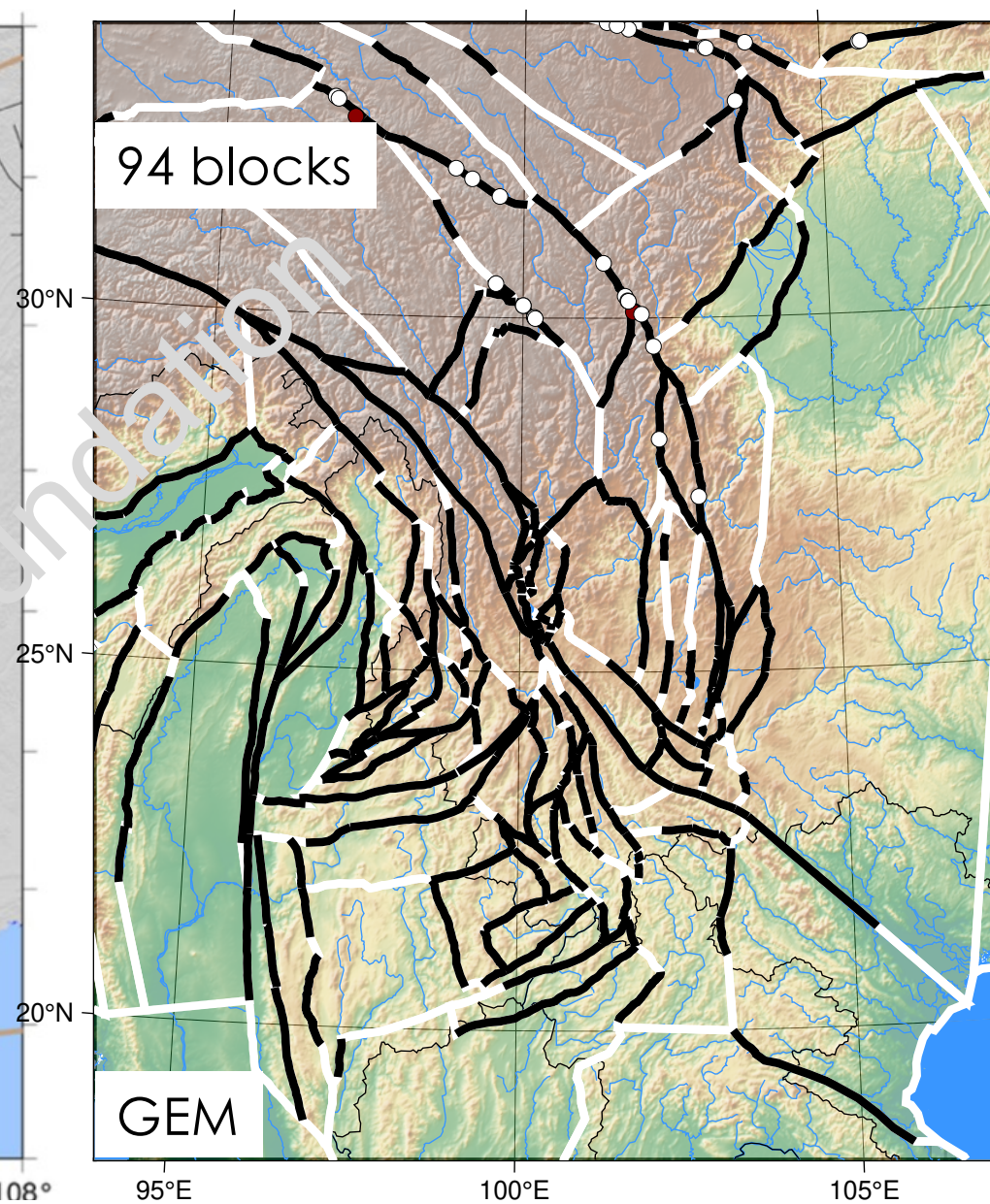
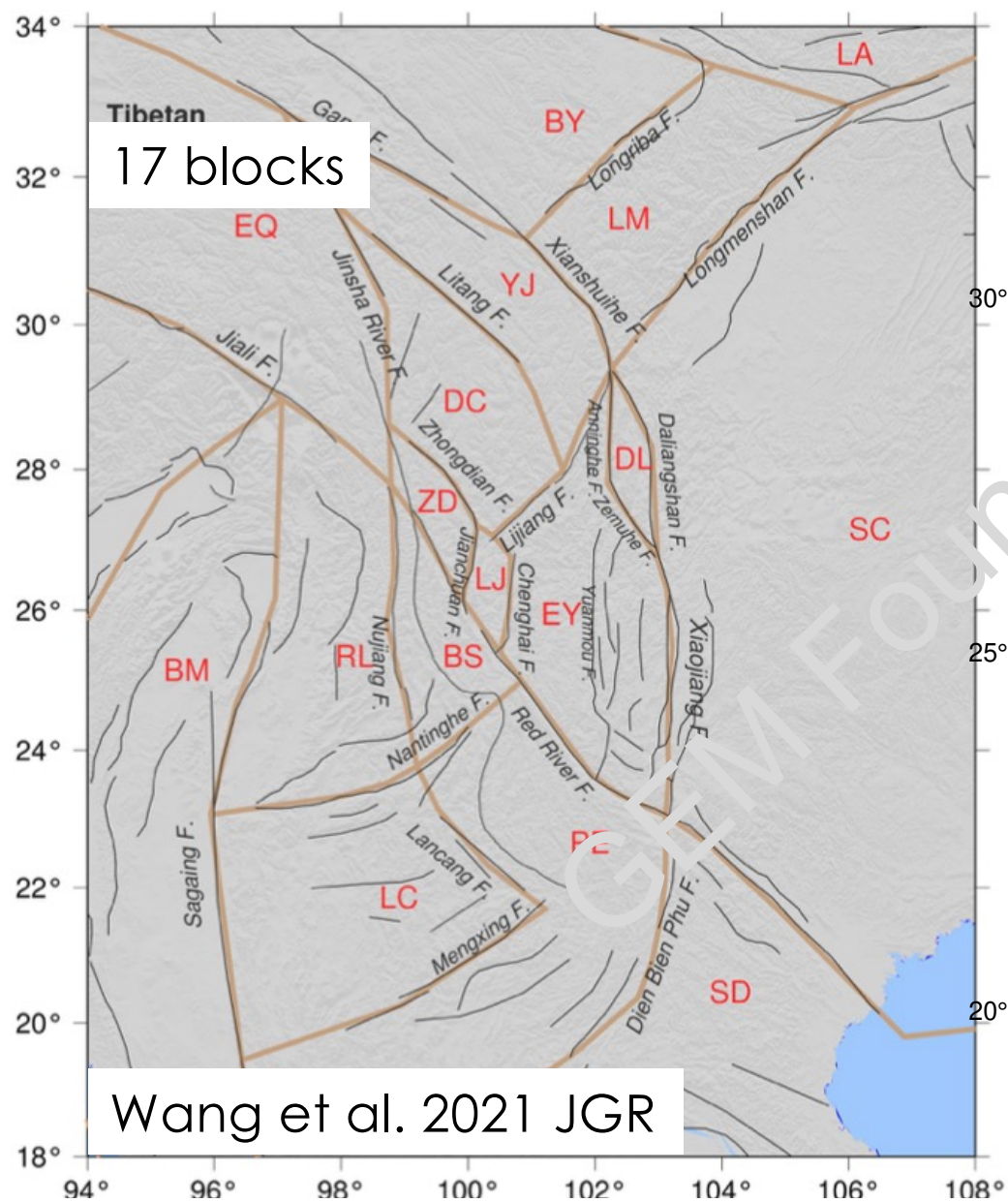
- ~50,000 GNSS + InSAR velocities
- 100s geologic slip rates
- Weighted least squares inversion (finds global optimum)
- Regularization and/or priors for subduction zone locking, blocks with no data
- Uncertainties (+covariance) from Monte Carlo perturbations of data

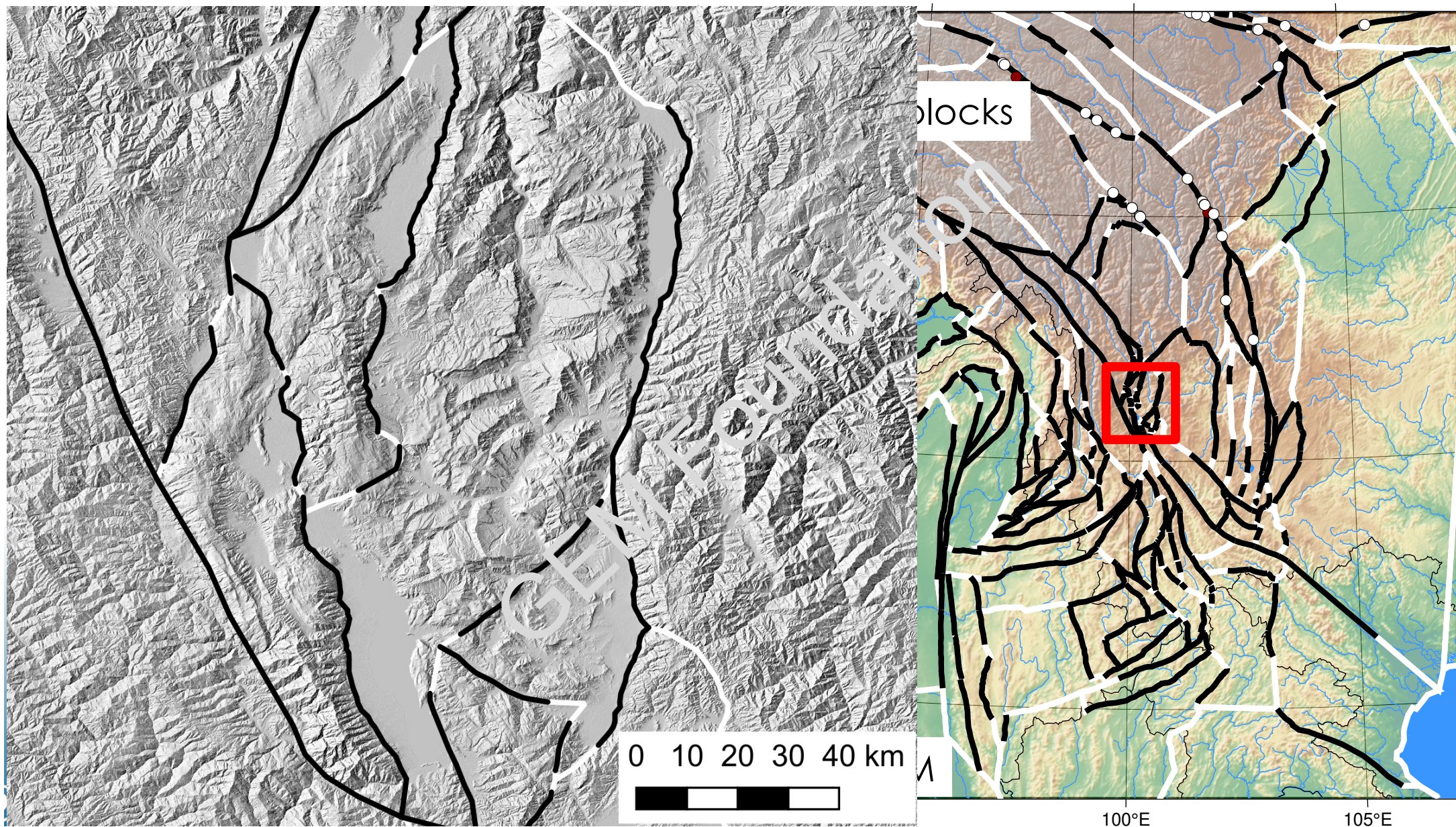


Fault and block mapping

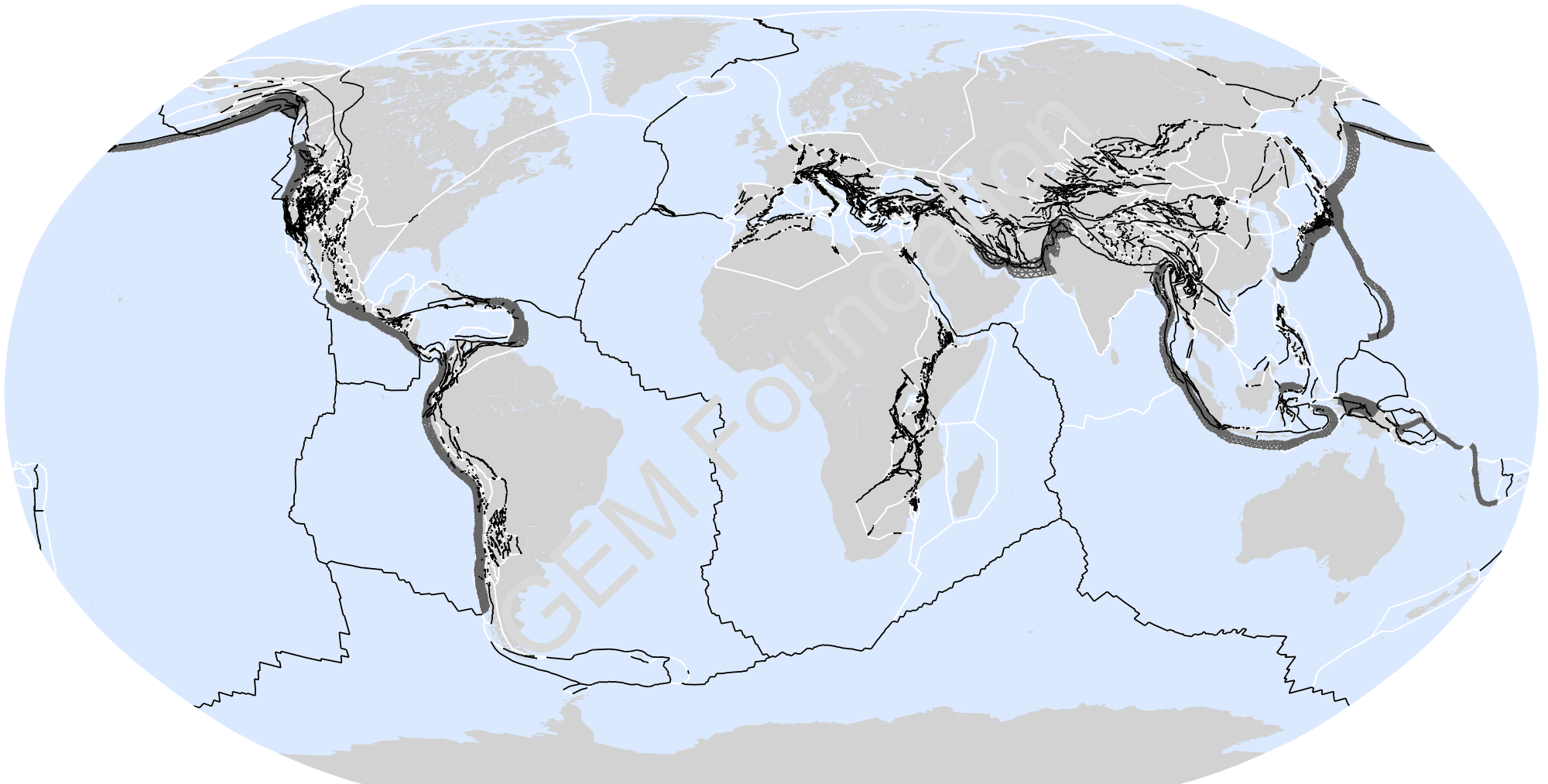
- Almost all faults remapped
 - Mapping from tectonic geomorphology, seismic and geodetic data, previous mapping
 - Some faults (mid-ocean ridges, challenging areas such as Japan) based more directly on existing fault data
- Faults and blocks mapped together
 - Necessary for high-quality models
 - As many faults as deformation data allow
- Fault mapping target 1:100,000 map resolution
 - Higher where datasets allow
 - Lower offshore
 - Each fault should be independent seismic source





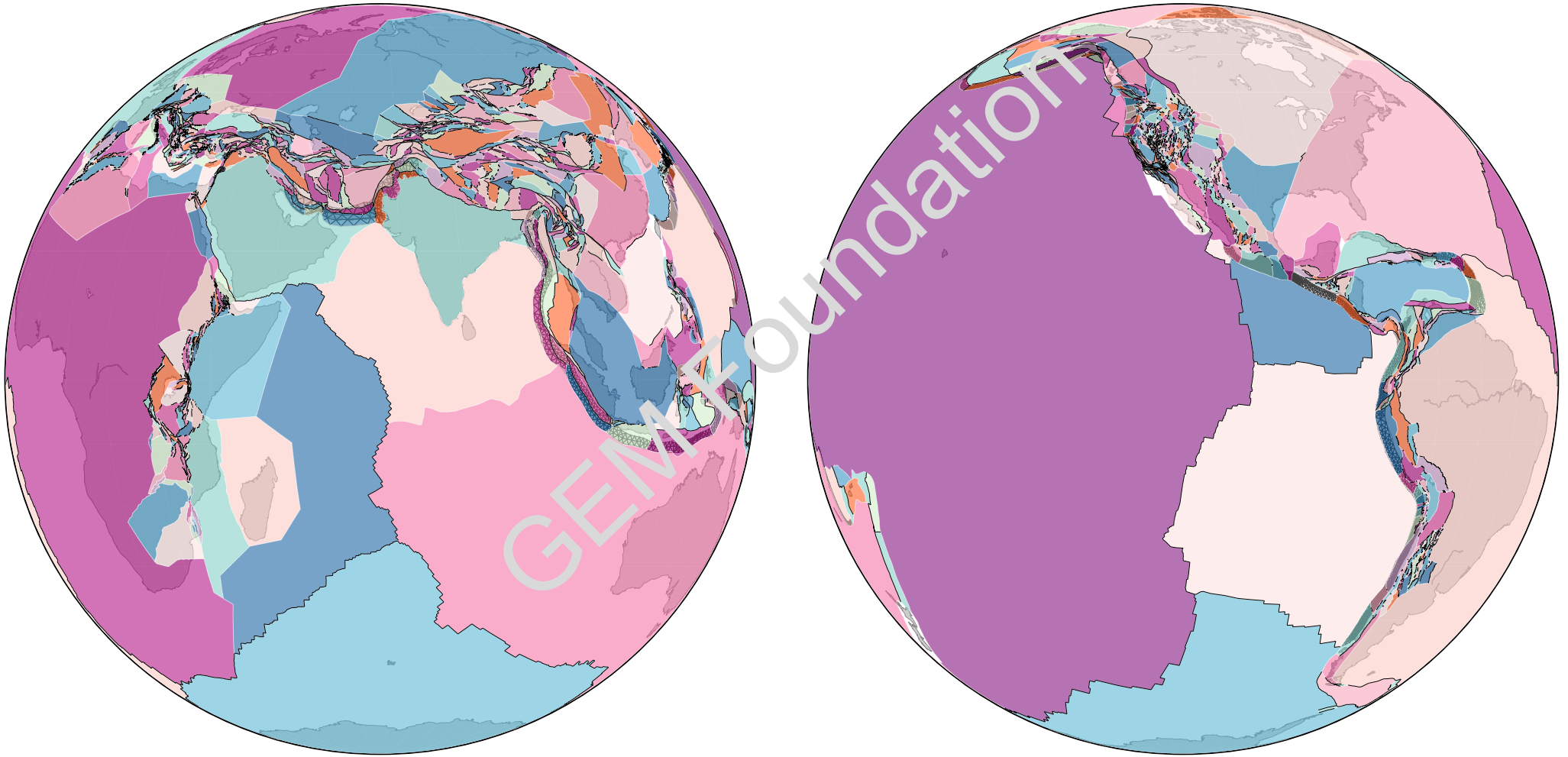


Total model geometry



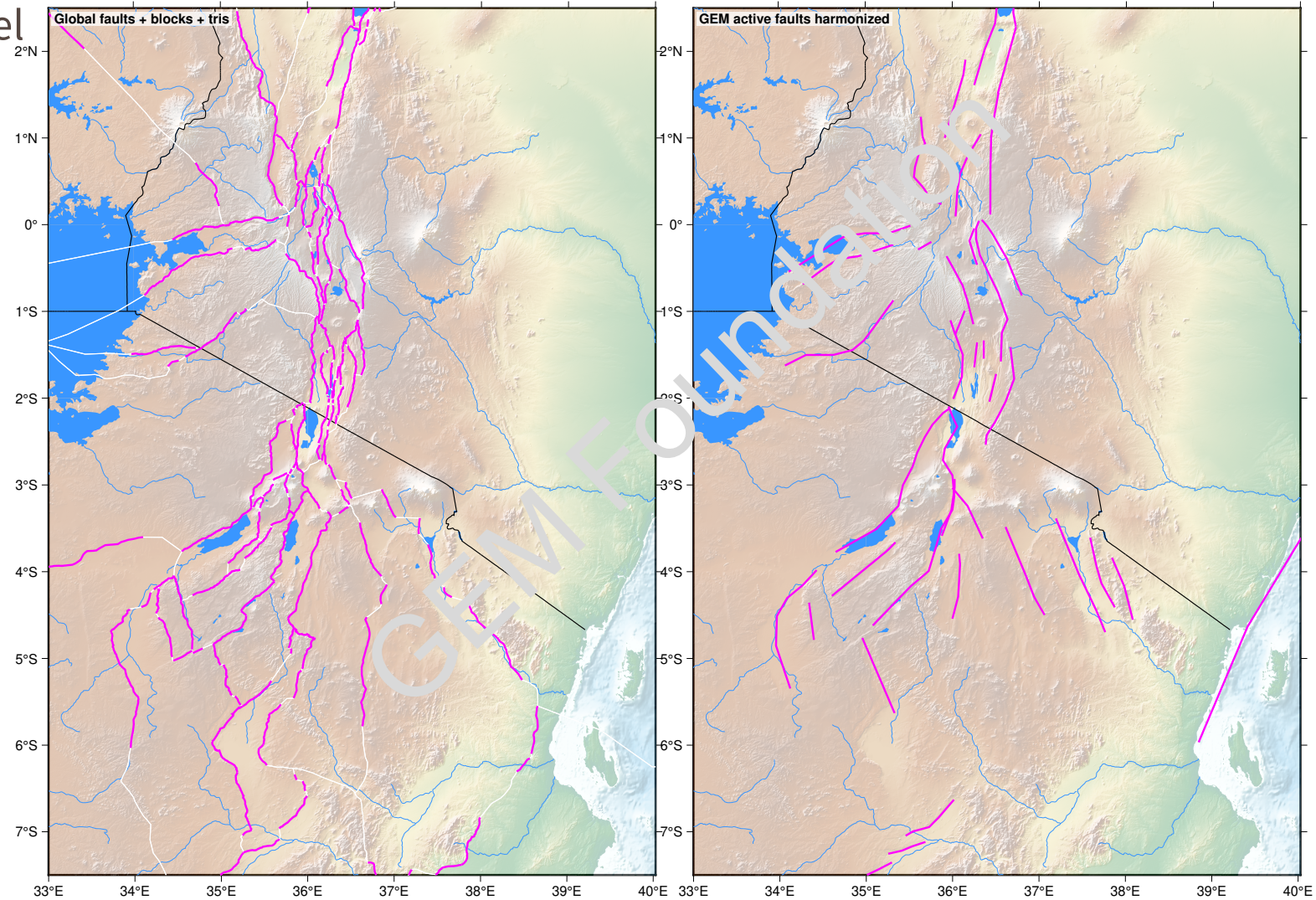
7,500 faults, 2000 blocks

More model geometry



Old vs. New: Kenya and Tanzania

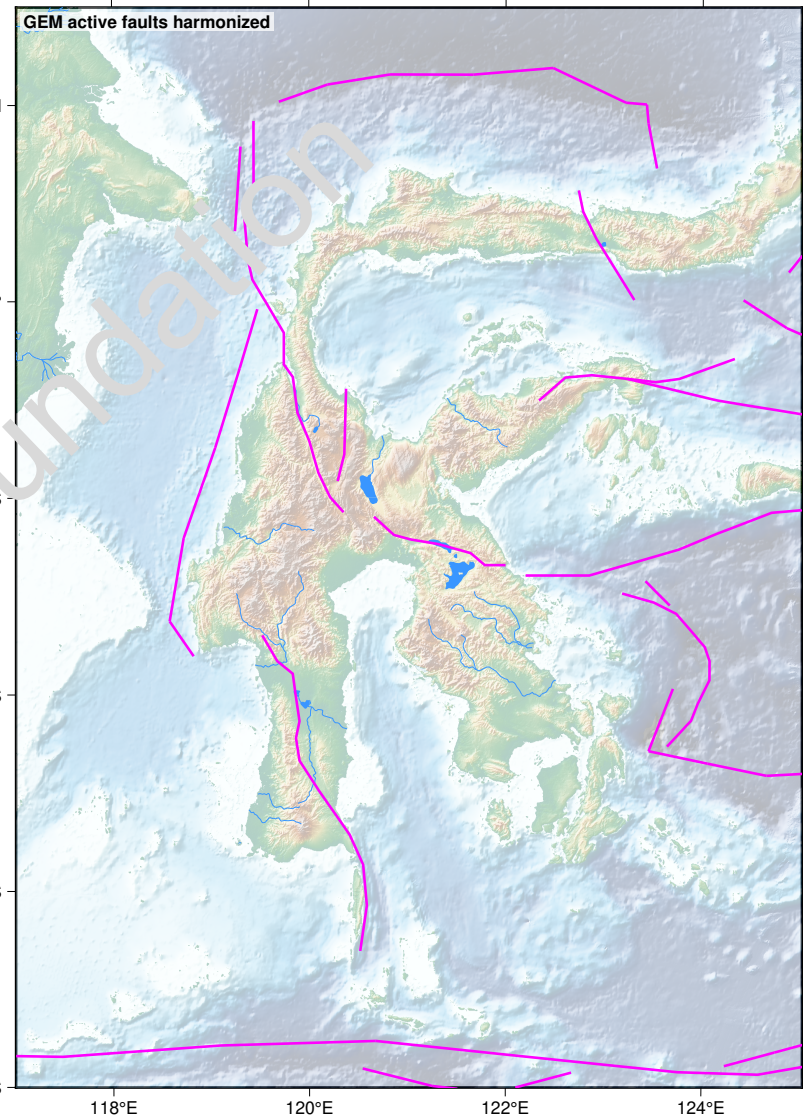
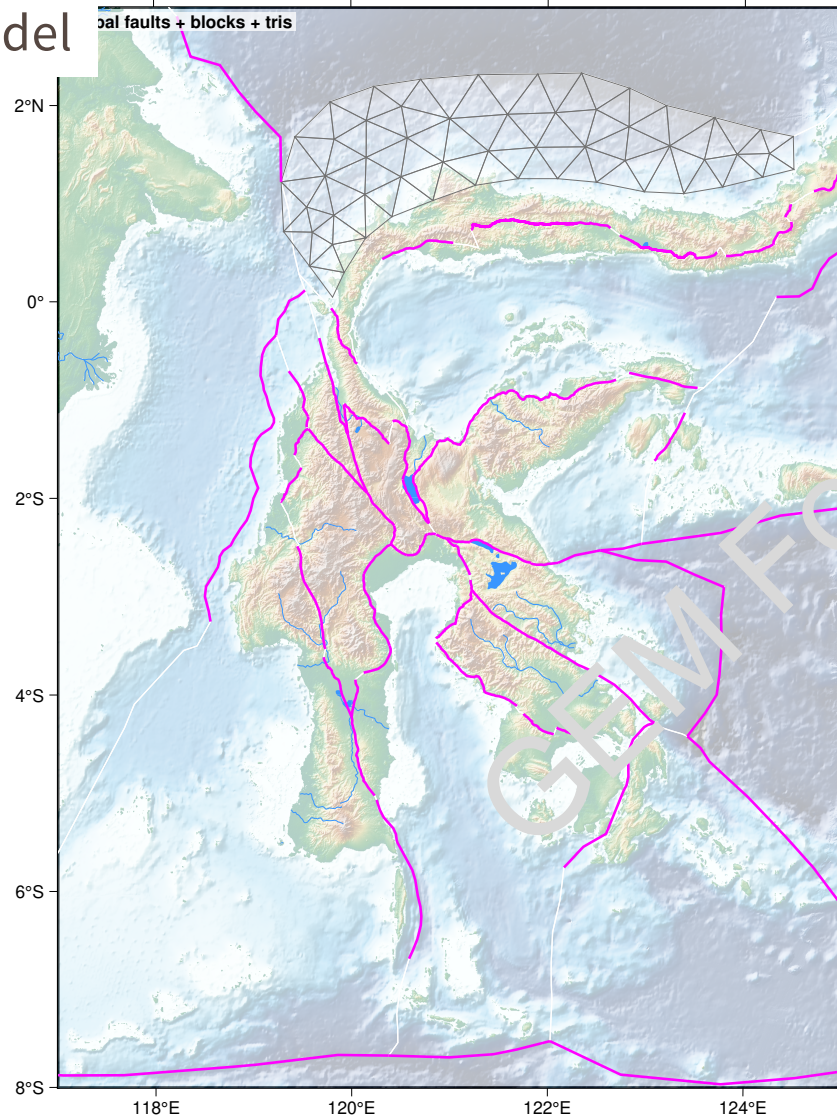
Block model



GAF-DB

Old vs. New: Sulawesi

Block model



GAF-DB

Subduction

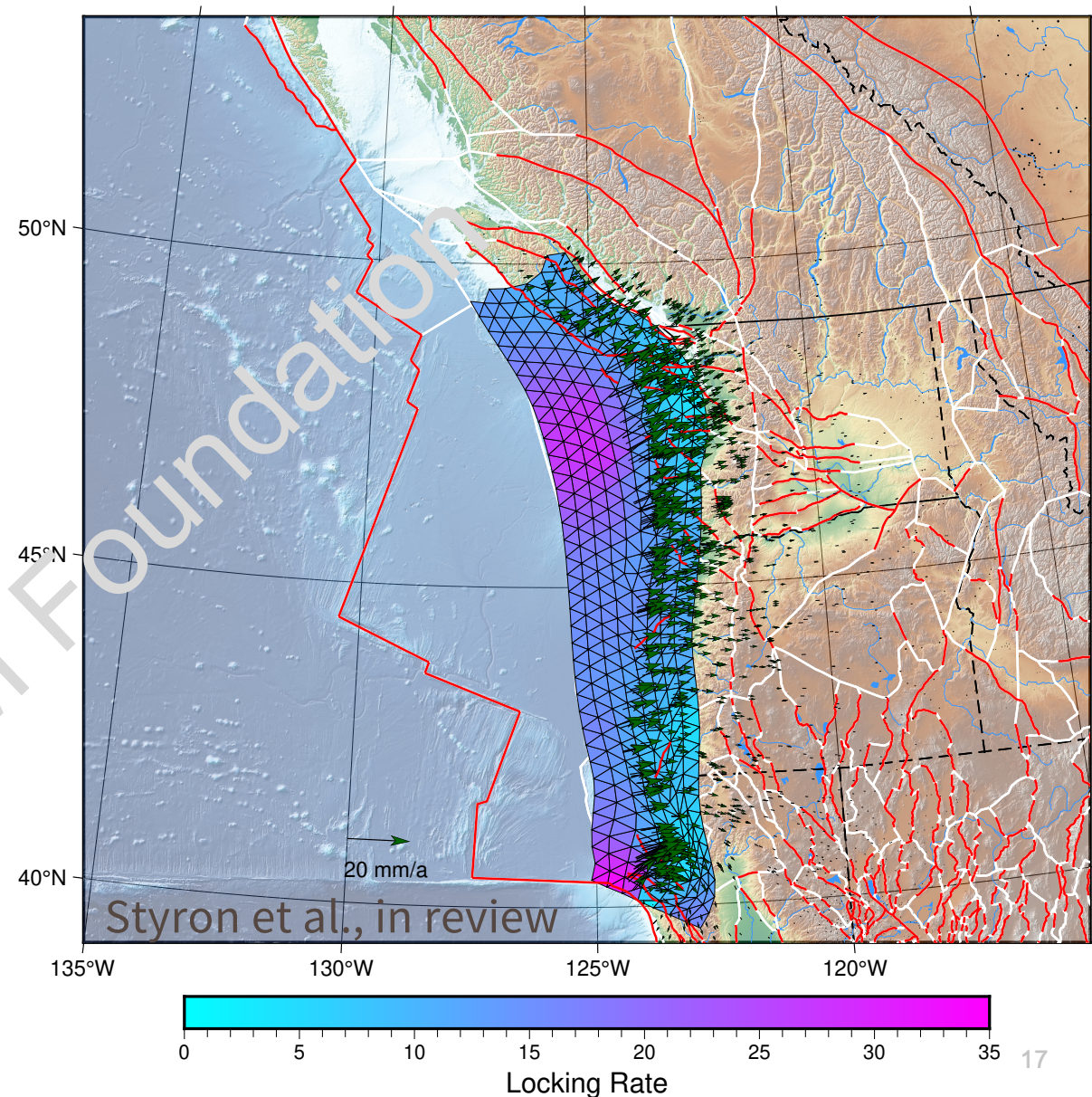
Spatially-variable coupling rate
calculated simultaneously with
block motions

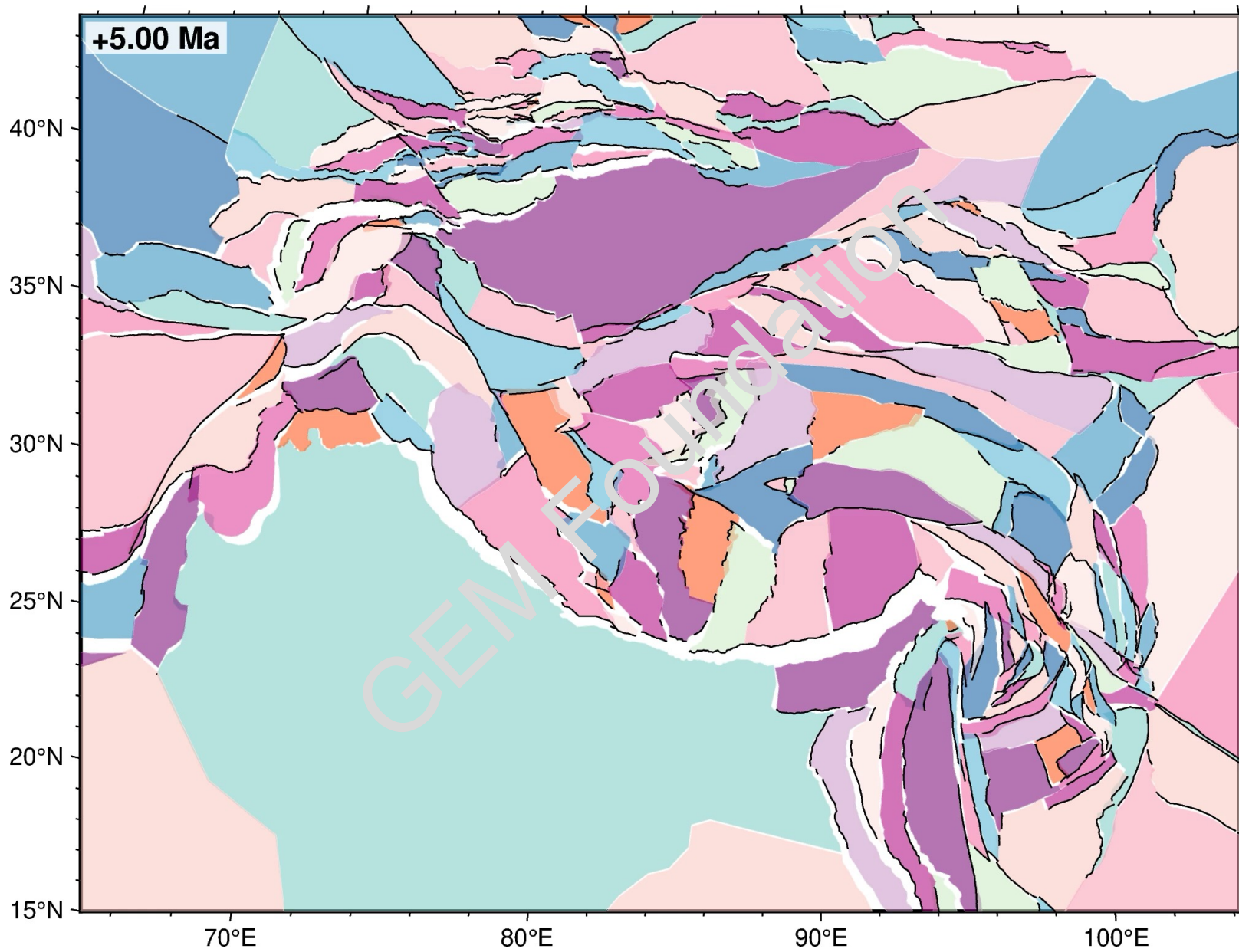
Mathematically underdetermined;
options for regularization and priors

Geometries for nearly all
subduction zones

- Most based on Slab 2.0
- Only upper ~60 km

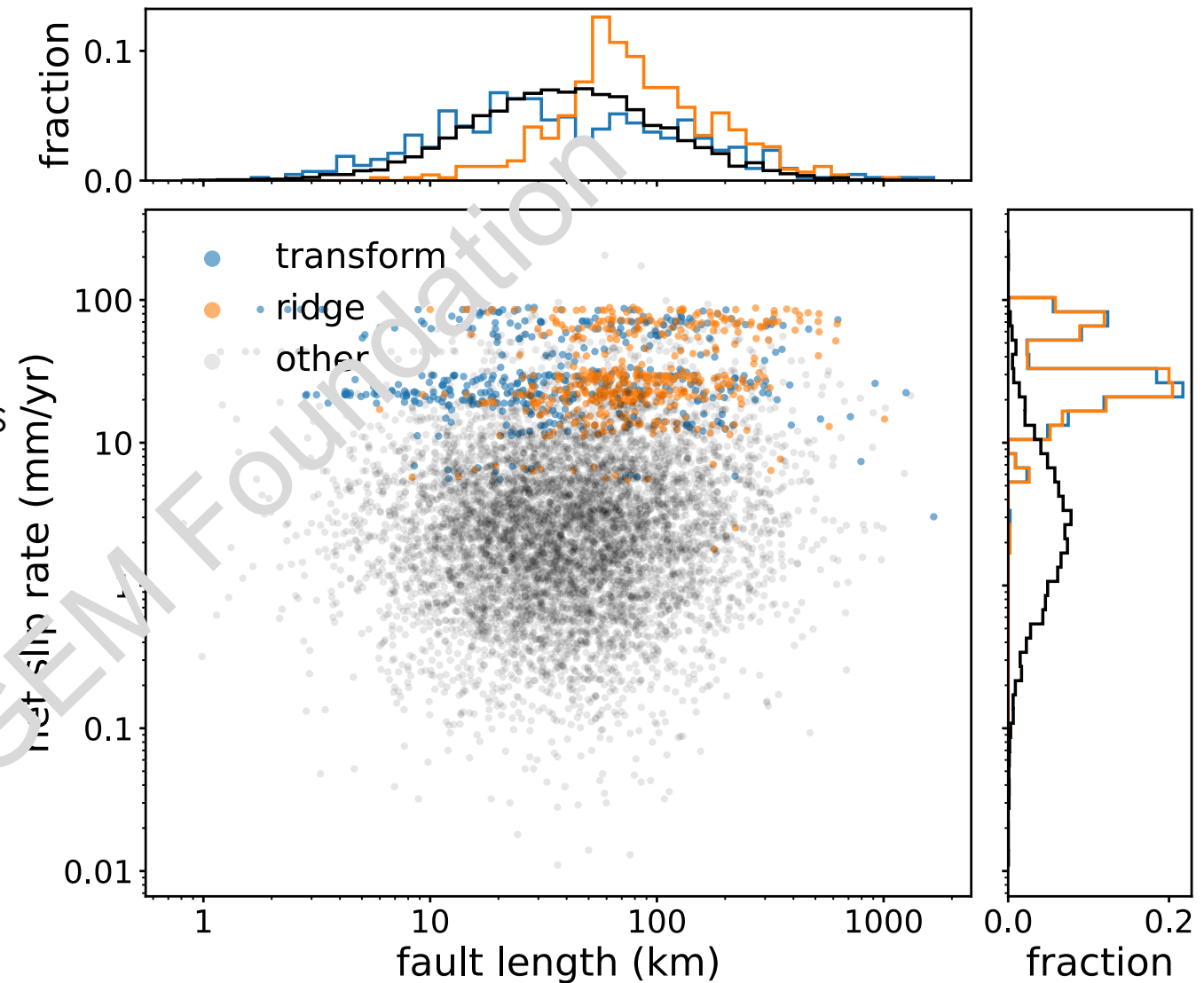
Results can be used for PSHA (with
Fermi)



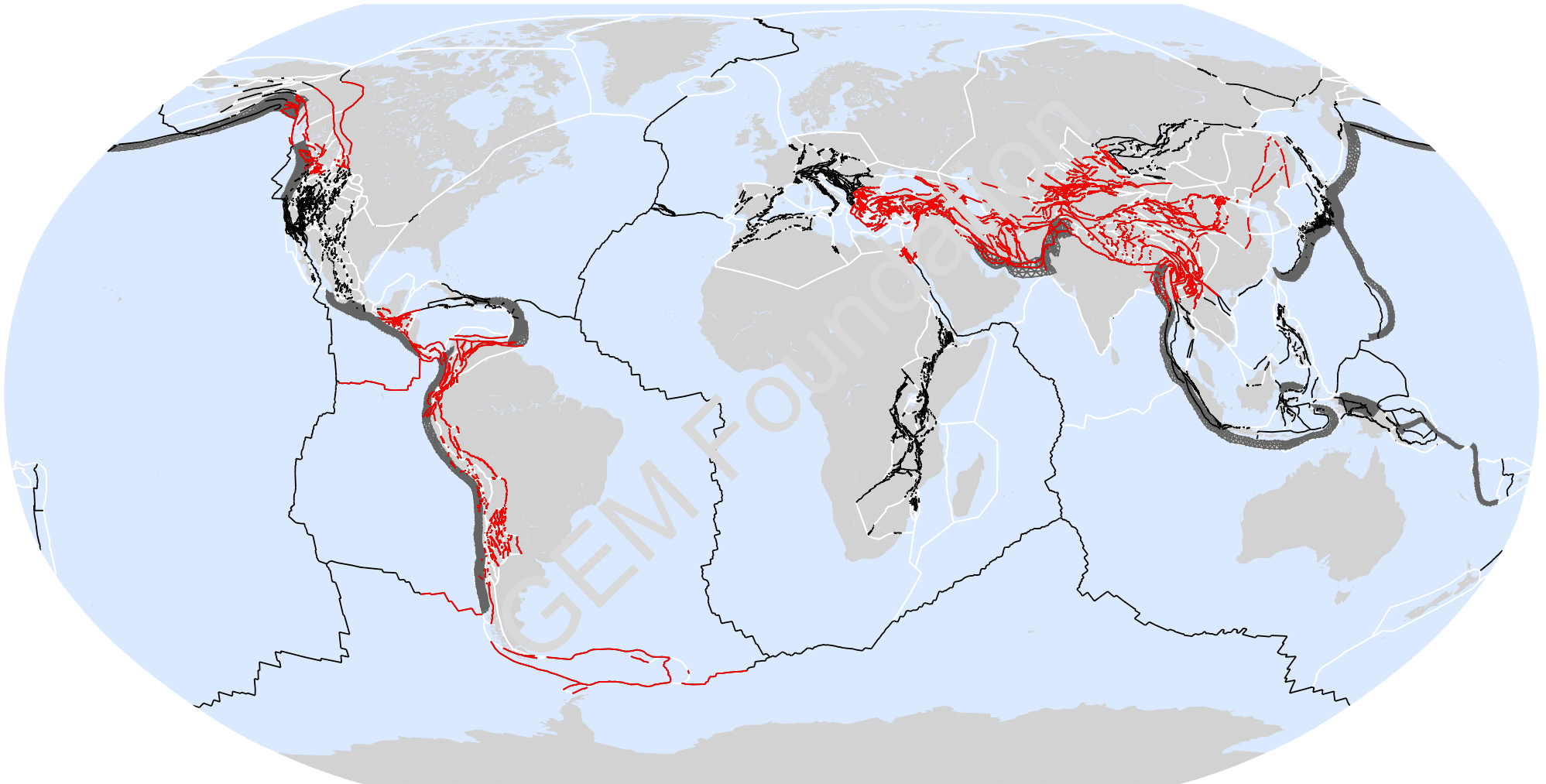


Slip rate results

- Slip rates have log-normal distribution, mode at ~3 mm/yr
 - Ridges and Transforms 10-30x faster
- Median uncertainty (std) is 60% of rate (mean)
- Lengths also log-normal, mode around 50 km



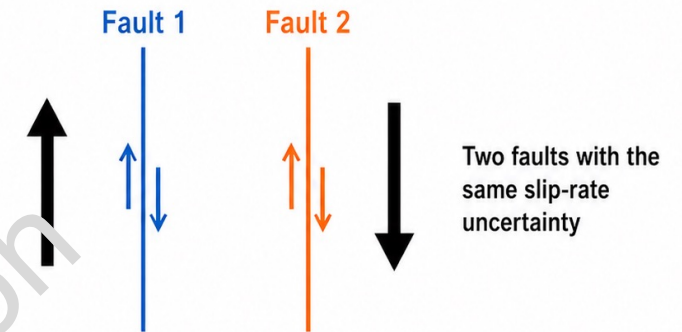
Usage to date in GEM models



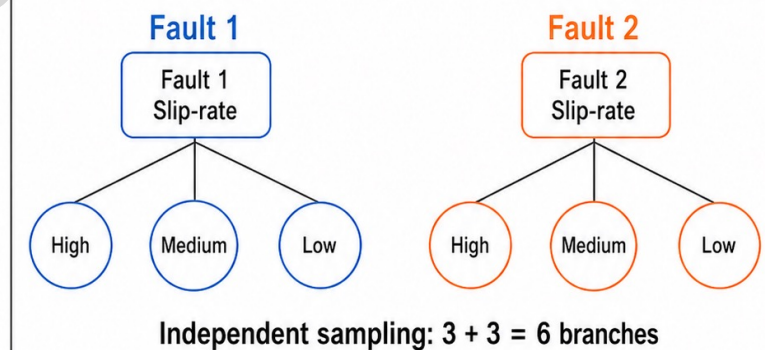
— Used in current or in-progress GEM model

Slip rate (co)variance and model complexity

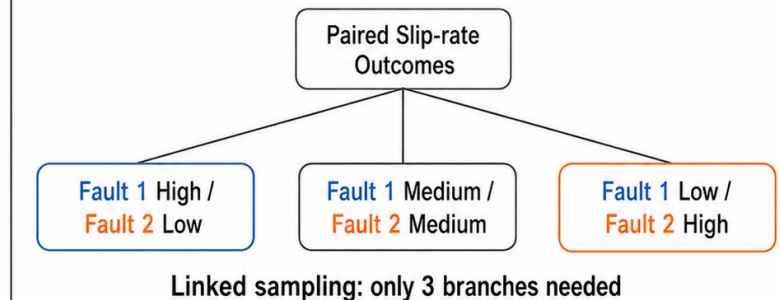
- Increased data precision requires less sampling for single component
- Correlated uncertainties reduce parameter space
- Systematization of faults can greatly increase model precision and decrease model complexity
- Very important as complexity of other components increases exponentially



Scenario A: uncorrelated uncertainties



Scenario B: negatively correlated uncertainties



Data sources and availability

- Many data sources used, but most mapping original linework
- Contributions from other scientists primarily geodetic data (thank you all!)
- Mapping will be complete July 2026
block model modifications through Dec 2026 (initial version!)
- Data and models will be available on Github, GEM products page



Thank you!

Please attribute to the GEM Foundation with a link to:
<https://www.globalquakemodel.org>



Except where otherwise noted, this work is licensed under a
Creative Commons Attribution-NonCommercial-NoDerivatives 4.0
International License (CC BY-NC-ND 4.0)
<https://creativecommons.org/licenses/by-nc-nd/4.0/>



June 23 – 25, 2026

University of Zagreb, Croatia

GEM CONFERENCE

FROM FAULTS TO FUTURE SCENARIOS



GEM Global Flatfile

A global ground-motion homogenisation effort

Christopher Brooks

GEM Hazard Team

June 23 – 25, 2026

University of Zagreb, Croatia

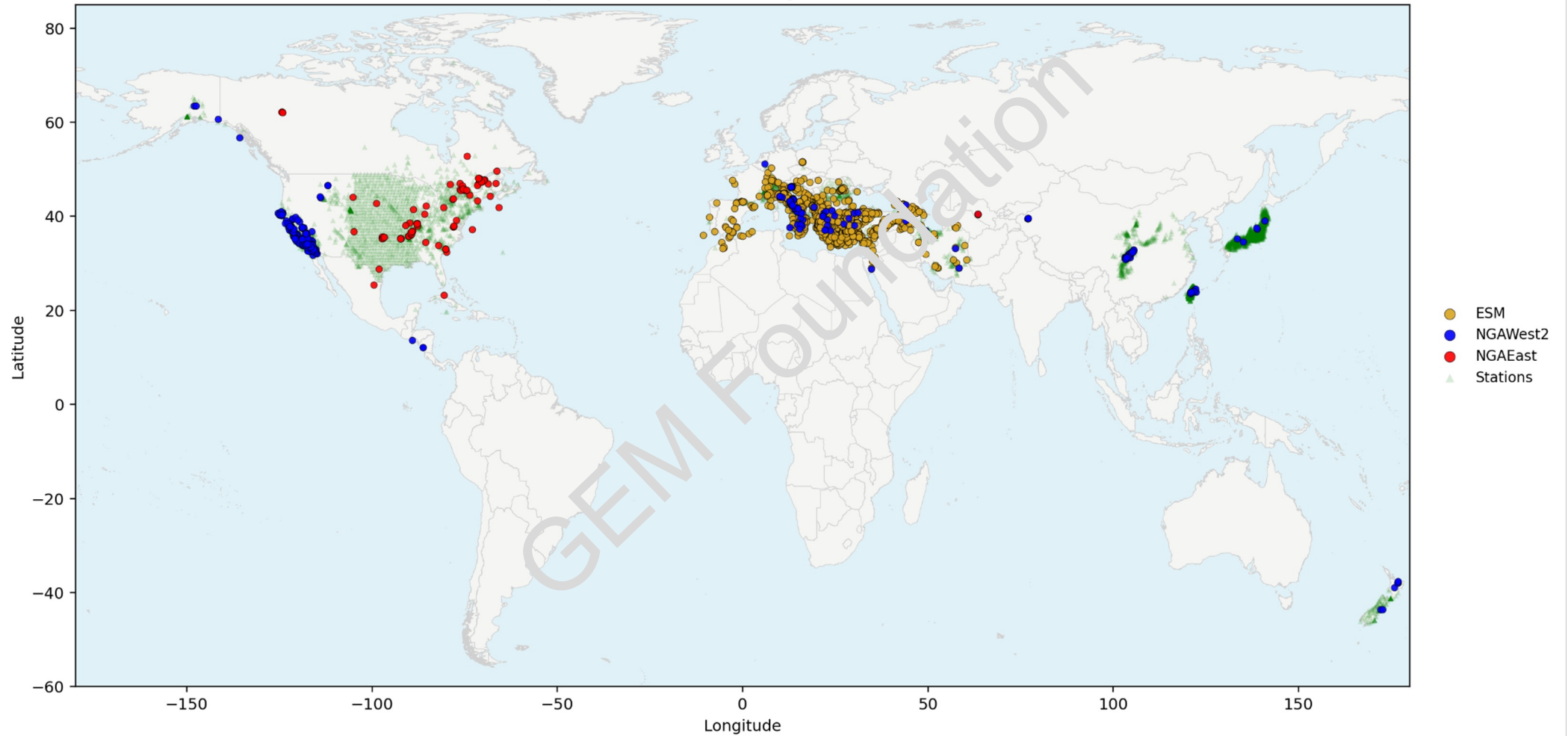
GEM CONFERENCE
FROM FAULTS TO FUTURE SCENARIOS

Flatfiles and the Lack of Homogeneity

- Flatfiles are generally defined as a collection of metadata from processed ground-motion recordings, usually represented as one record per row
- The primary purposes of such datasets are ground-motion model (GMM) development and selection
- Larger datasets and those for previously less well-recorded regions are being published more frequently
- But the lack of homogeneity among these datasets makes them difficult to use
- Ideally, a significant time investment should not be required by the user to parse the dataset into their workflow
- It is also common for flatfiles covering overlapping regions to contain metadata from the processing of recordings from only a **specific** seismic network (or subset of them)
- We are trying to resolve these issues with the **GEM Global Flatfile**

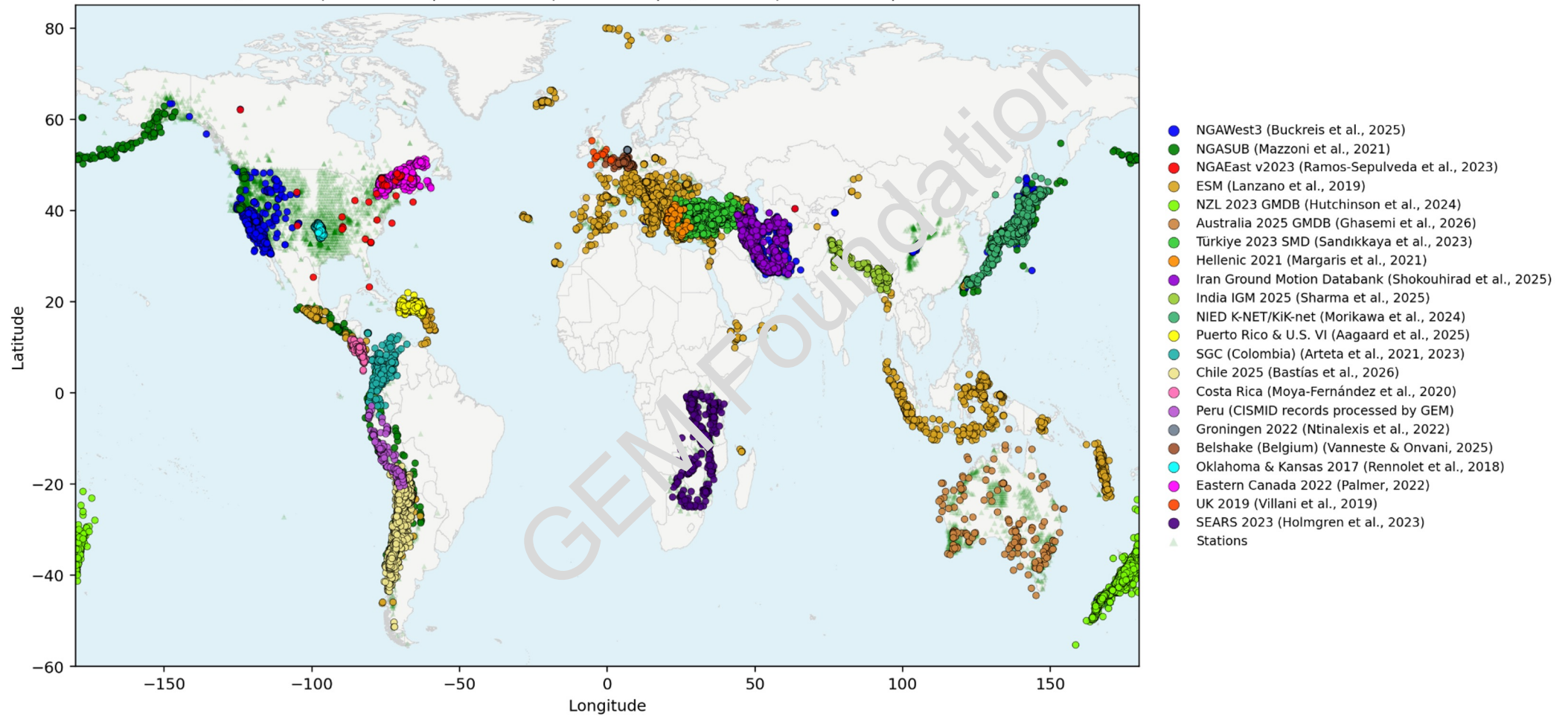
In the 2018 Mosaic Release We Had Only a Few Flatfiles

Publicly Available Flatfiles During 2018 Mosaic Release



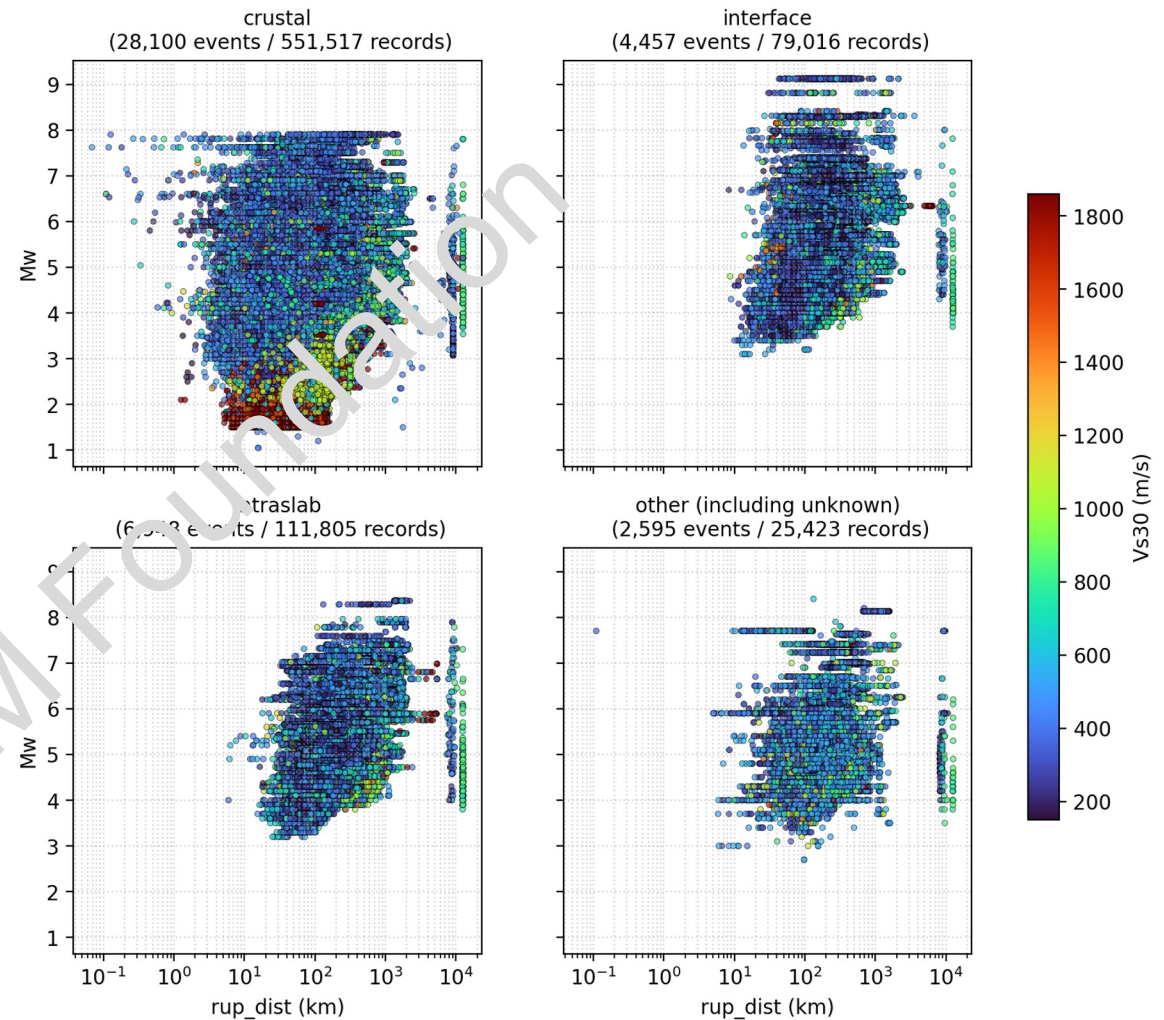
Now We Have Many More...

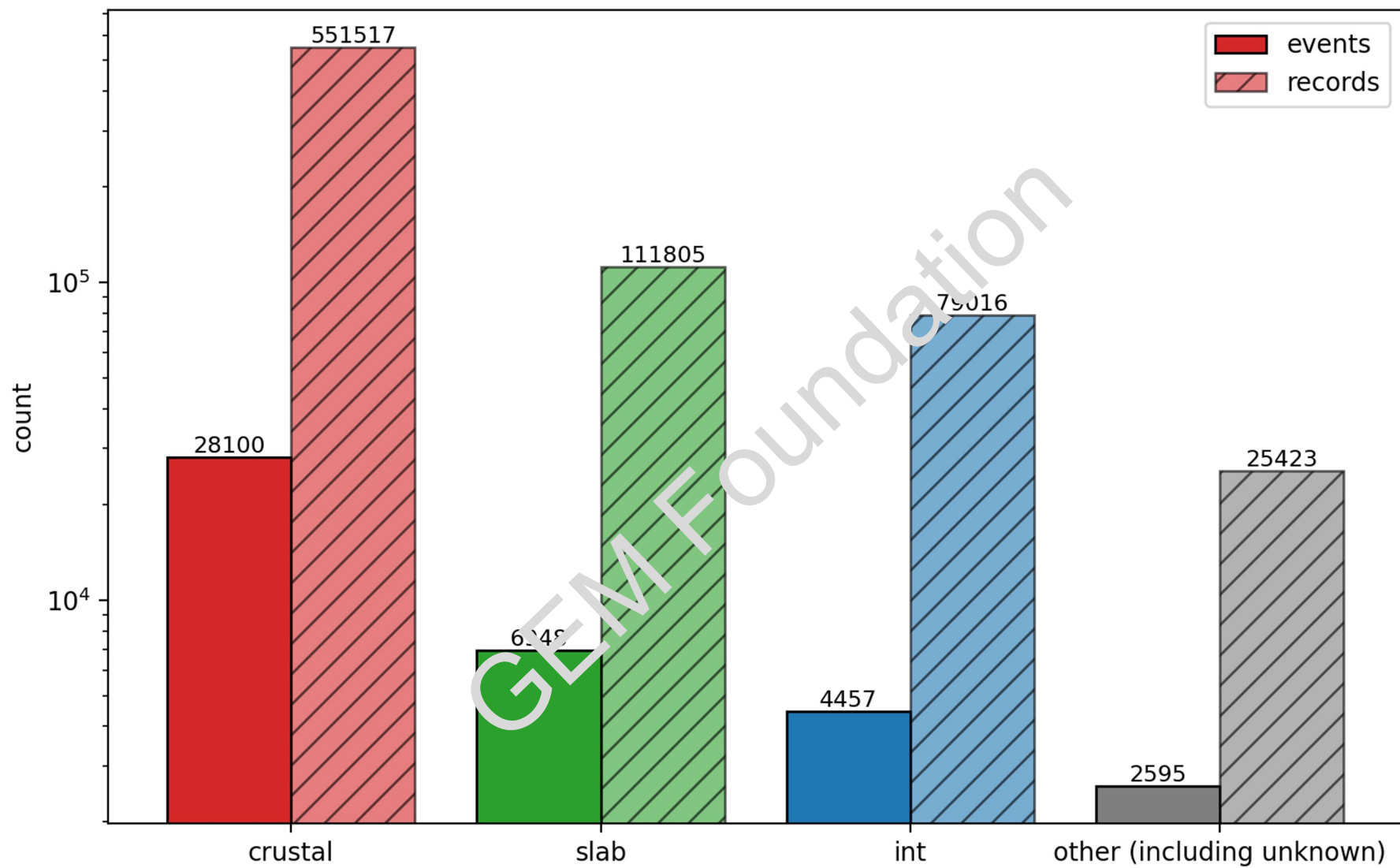
GEM Global Flatfile — coverage by dataset
767,761 unique records | 42,100 unique events | 23,198 unique stations | 22 datasets



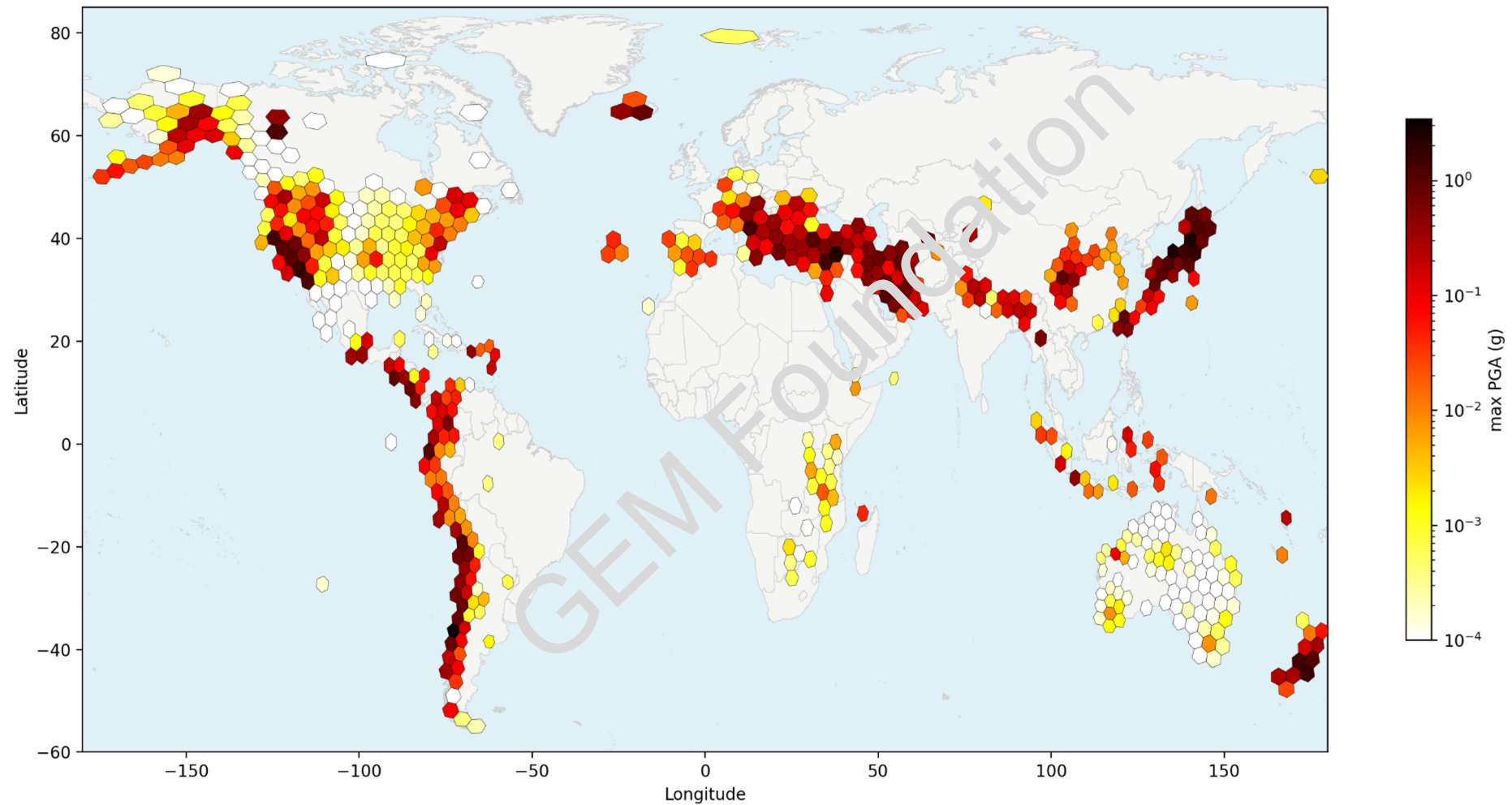
The GEM Global Flatfile

- Fully homogenised = consistent event/station info over different datasets and duplicate records removed
- We supplement the metadata in the incorporated datasets through:
 - 1) Checking the GEM Global Catalogue for missing event information (e.g. nodal plane solutions)
 - 2) Performing tectonic region classification of the events using the OQ-MBTk's classifier tool
 - 3) Recomputing rupture distances following homogenisation of event and station info (if missing)
- The flatfile is readily usable in the OQ-MBTk's **Strong Motion Tools (SMT)** for residual analysis (GMM selection) (as used in the Mosaic)
- Each incorporated flatfile's **license** is fully respected, and the data from each incorporated dataset is acknowledged/tracked

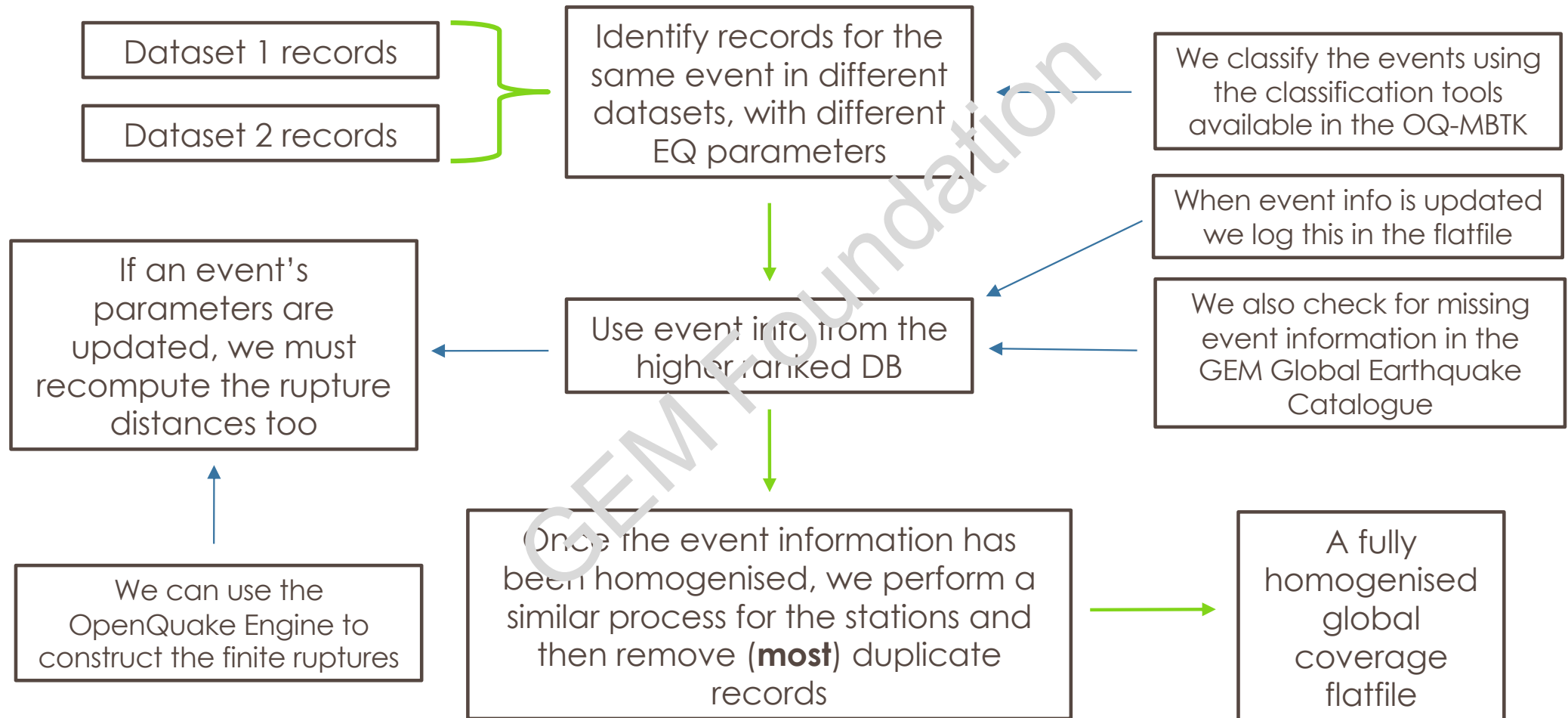




Max Horizontal PGA per H3 Cell of Resolution 2 ($M_w \geq 5$ events only)



Homogenisation Process

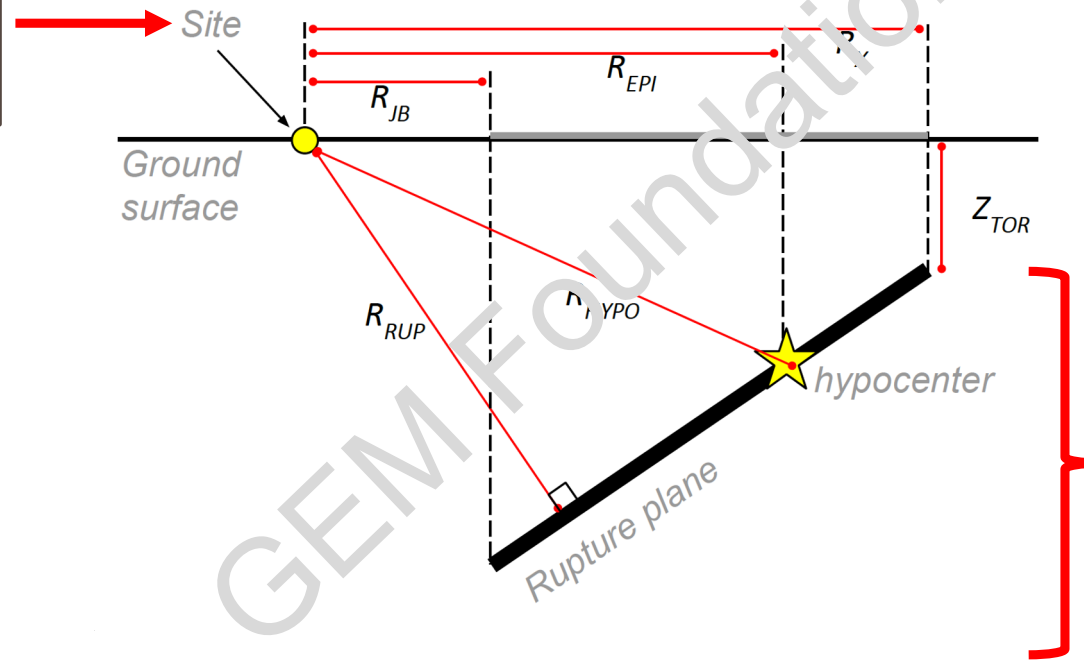


Incorporated Information

Event Info	Station Info	Record Metadata
• Event date/time and ISC event ID • Hypocentre Location • Magnitudes (M_w, M_L, M_s, M_d and m_b) • Nodal plane solutions • Rupture dimensions • Tectonic region type (GEM + original dataset) • Event homogenisation info	• Station location • Station code • Network code • V_{s30} and V_{s30} measurement type • Depths to ref. velocities e.g., $z1pt0$ (1000 m/s), $z2pt5$ (2500 m/s) • Volcanic backarc flag • Station homogenisation info	• Finite rupture distances (e.g., R_{rup}, R_{jb}, R_x, P_y, 0) • Scalar ground-motion intensity values (PGA, PGV, PGD, CAV, IA, SA(T) for T up to 10 sec) • Values are provided for H1, H2, vertical RotD50, RotD00, RotD100 • Shortest and longest usable periods • Record homogenisation info including flag informing if distances were (re)computed using OQ Engine

Using the Flatfile to Reconstruct Ground-Shaking Scenarios

The site in the context of the flatfile is each station recording the given event



The flatfile contains the information required to (best as possible) reconstruct a finite rupture

Usable in the OQ-MBTK's Strong Motion Tools for Residual Analysis

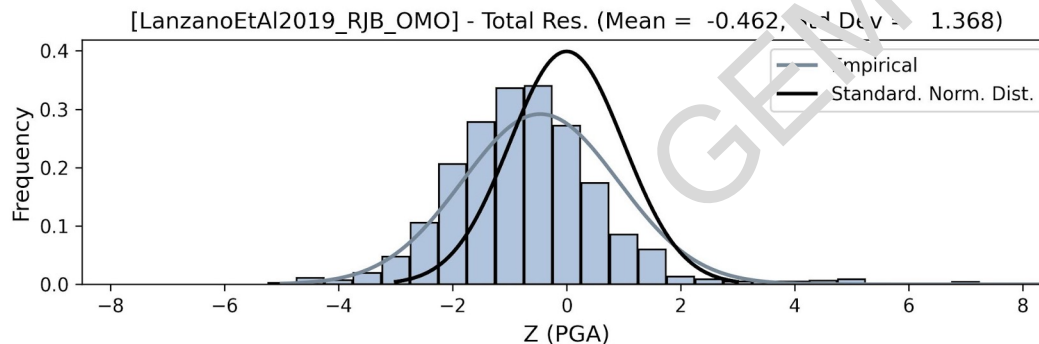
Provide a dataset of processed ground-motion recordings

Parse the dataset into SMT metadata (a database of records including event and station info)

For each record, construct the corresponding **ground-shaking scenario** within the OpenQuake Engine and compute the expected ground-motions

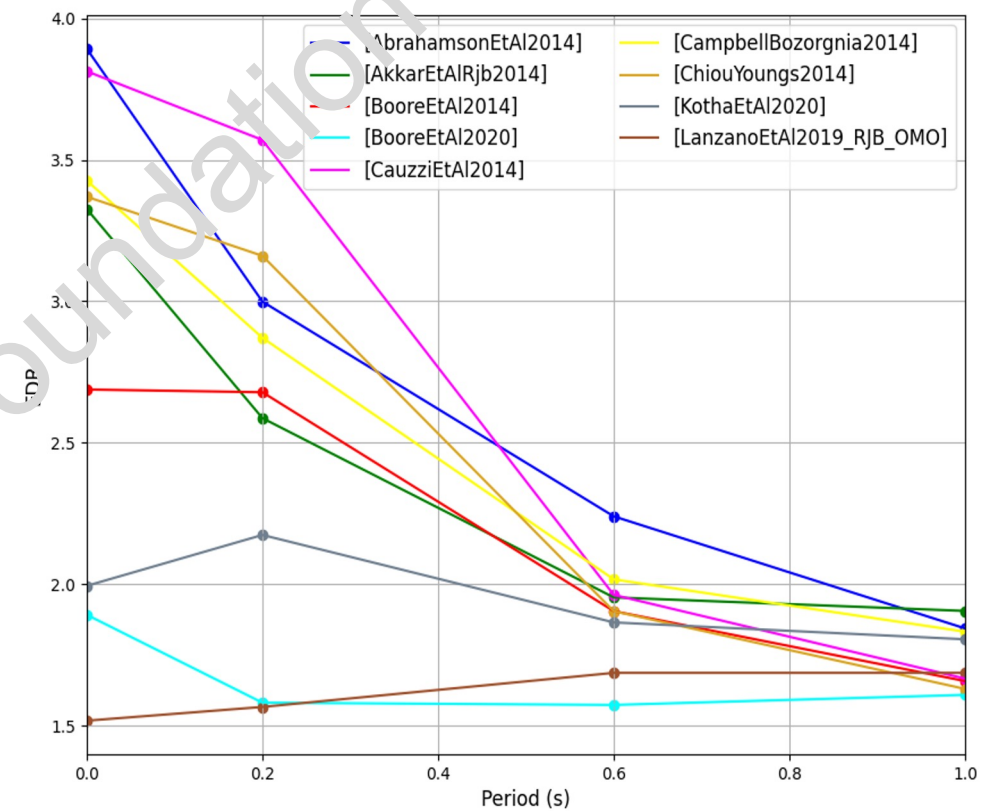
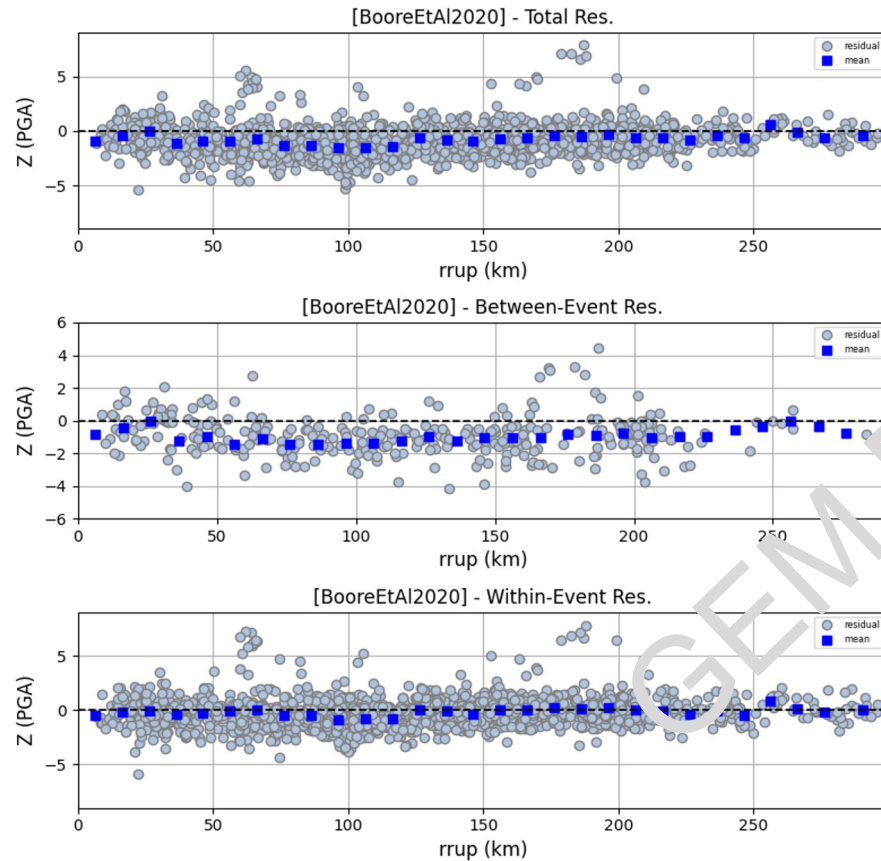
Compute the distributions of **residuals** (one per GMM and per IMT)

$$Z_{normalised} = (\ln(observed) - \ln(predicted)) / GMM \text{ sigma}$$



<https://github.com/GEMScienceTools/oq-mbtk>

Fully Supported in the OQ-MBTK's Strong Motion Tools Module



Use Case: Validation of the Global Risk Model with GEESE

- The Global Earthquake Scenarios (GEESE) rupture matching algorithm can retrieve an appropriate finite rupture for an admitted event
- It can also **check the GEM Global Flatfile** for station data for the queried event
- GEESE converts it to OQ format for use in the **conditioning** of ground-motion fields (GMFs)
- This capability is being used to support the Risk Team's validation of the Global Risk Model
- Leveraging the flatfile as a supplementary input for GEESE is therefore providing a capability to rapidly retrieving station data for many significant and instrumentally recorded events

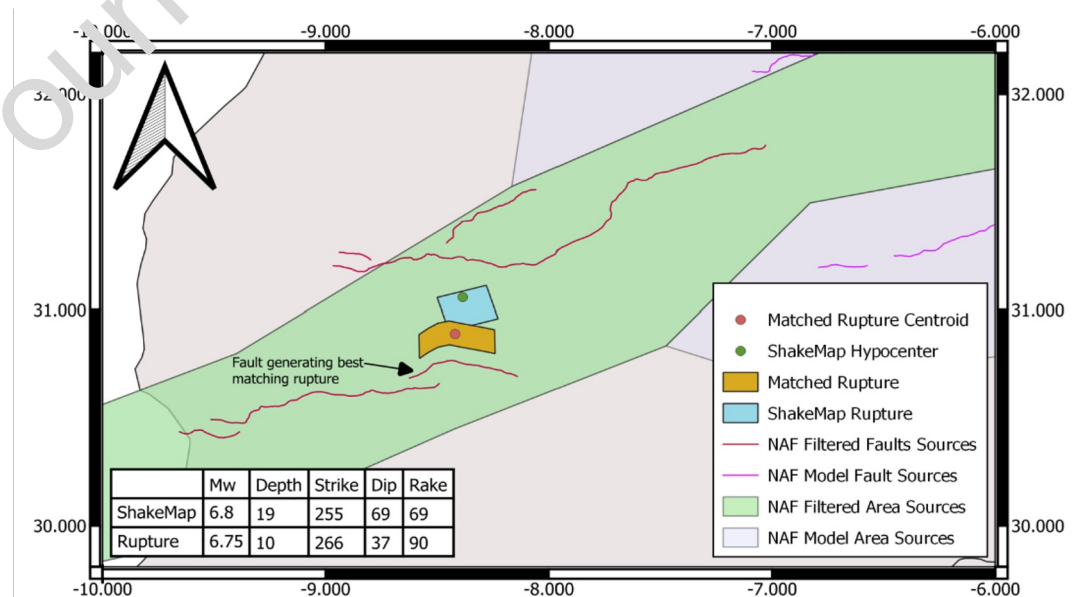
doi:10.26443/seismica.v4i2.1654

SEISMICA

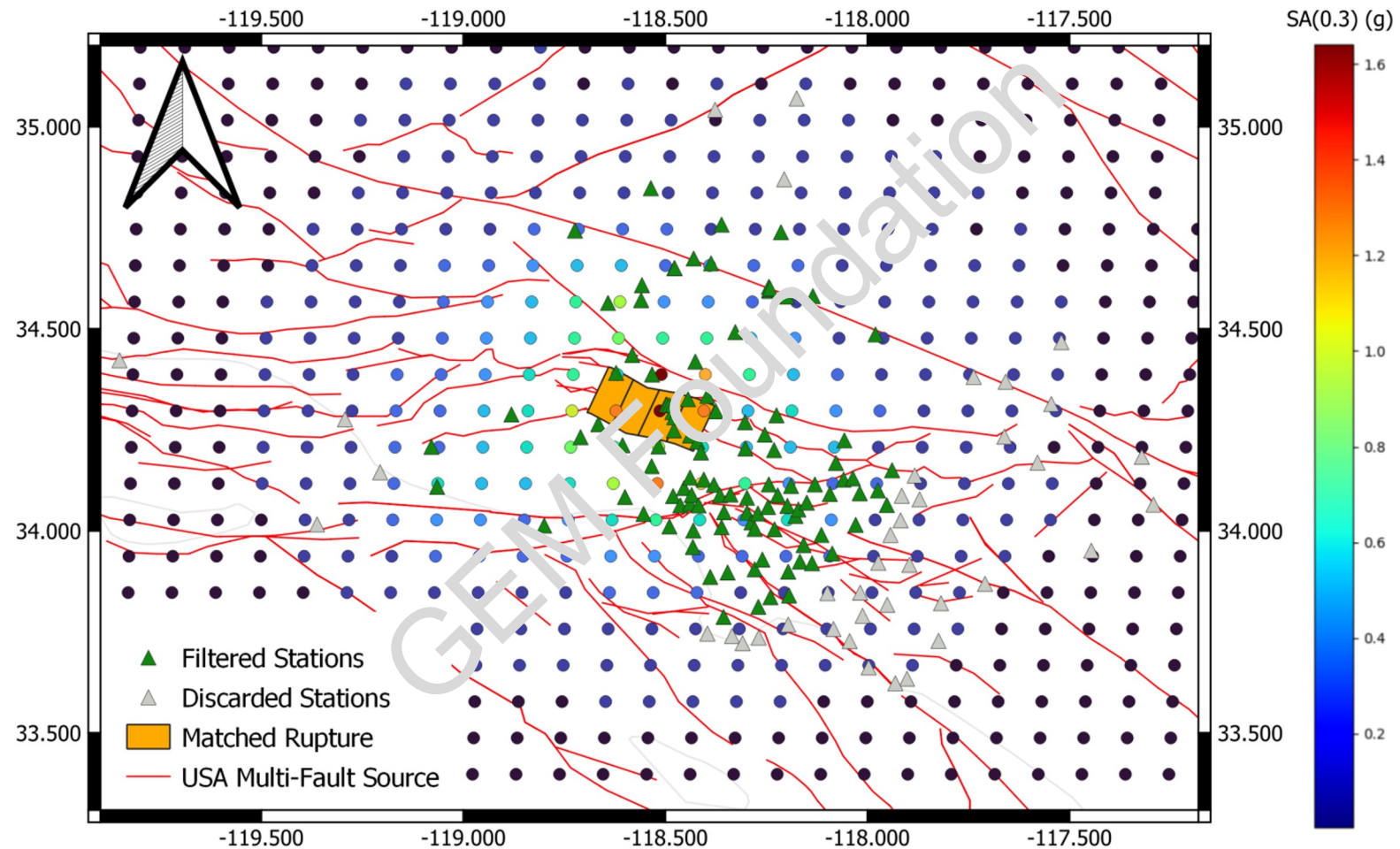
Global Earthquake Scenarios (GEESE): An OpenQuake Engine-Based Rupture Matching Algorithm and Scenarios Database for Seismic Source Model Testing and Rapid Post-Event Response Analysis

Christopher Brooks ¹, Marco Pagani ^{1,2}, Manuela Villani ^{1,3}, Kendra Johnson ¹, Richard Styron ¹, Kirsty Bayliss ¹

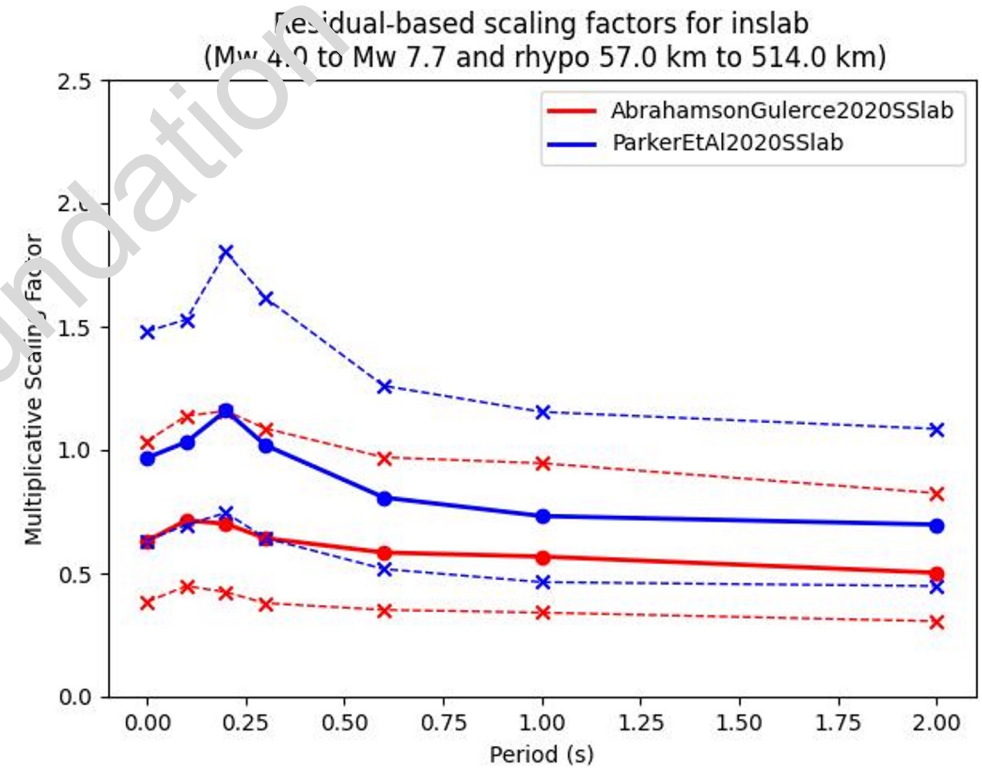
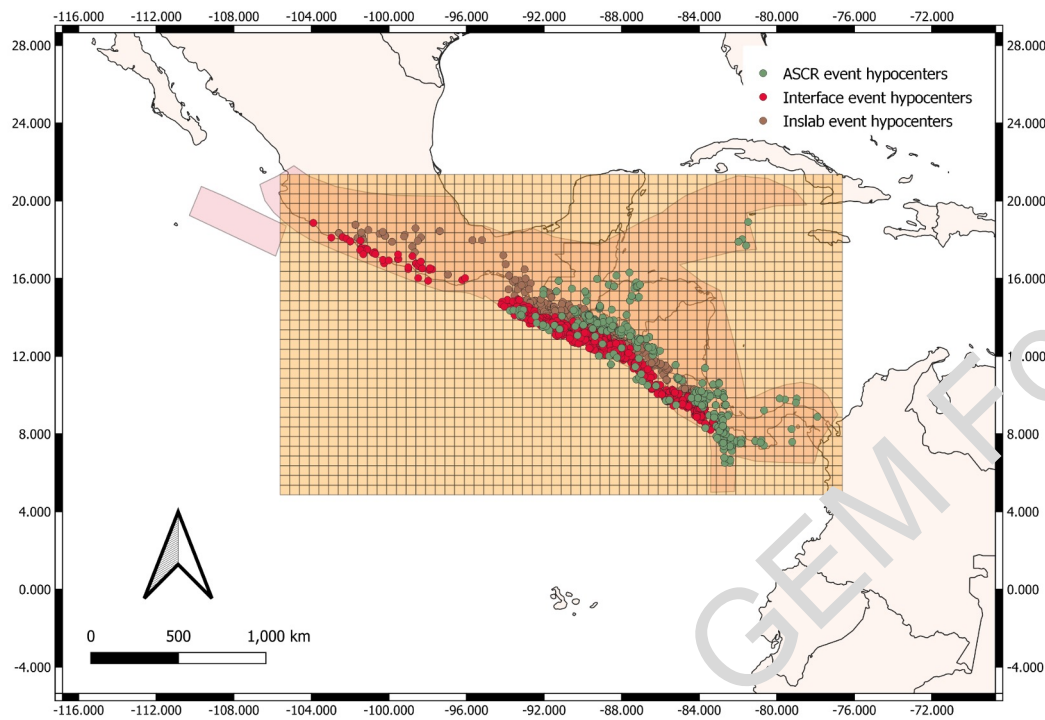
¹Global Earthquake Model (GEM) Foundation, via Ferrata 1, 27100, Pavia, Italy, ²Adjunct Professor, Institute of Catastrophe Risk Management, NTU, Singapore, ³Imperial College London, United Kingdom



Conditioning of GMFs for a GEESE Rupture for the 1994 Northridge EQ

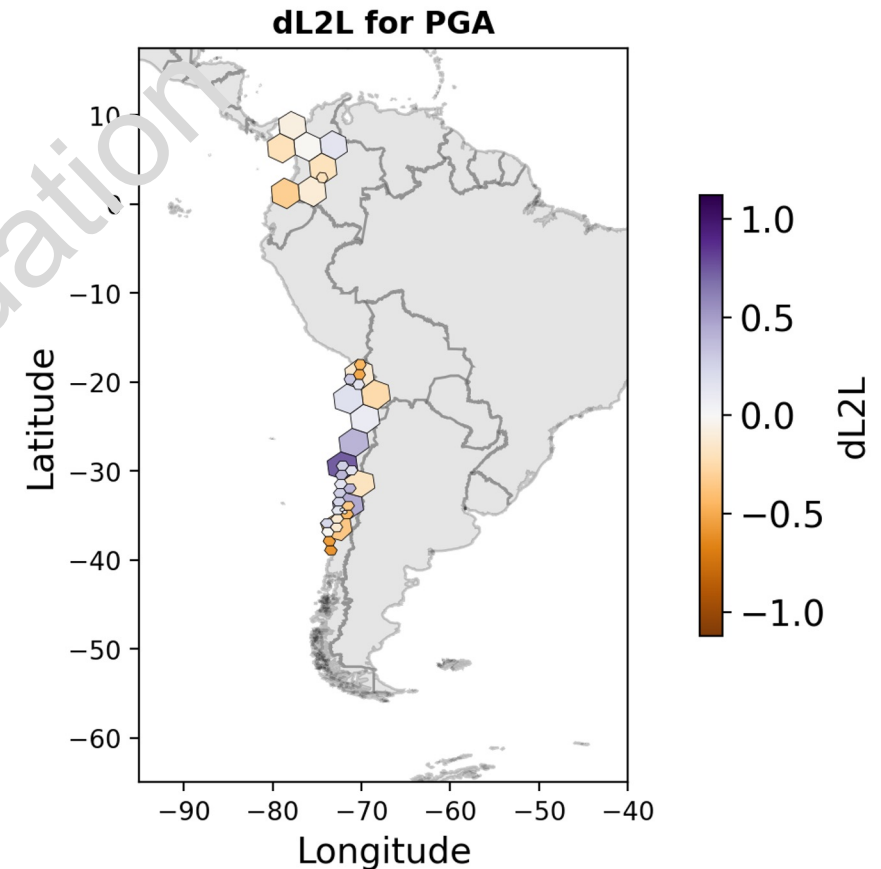


Use Case: Developing Hybrid Backbone GMCs for Central America



Use Case: Partially Non-Ergodic GMM for next South America (SAM) model

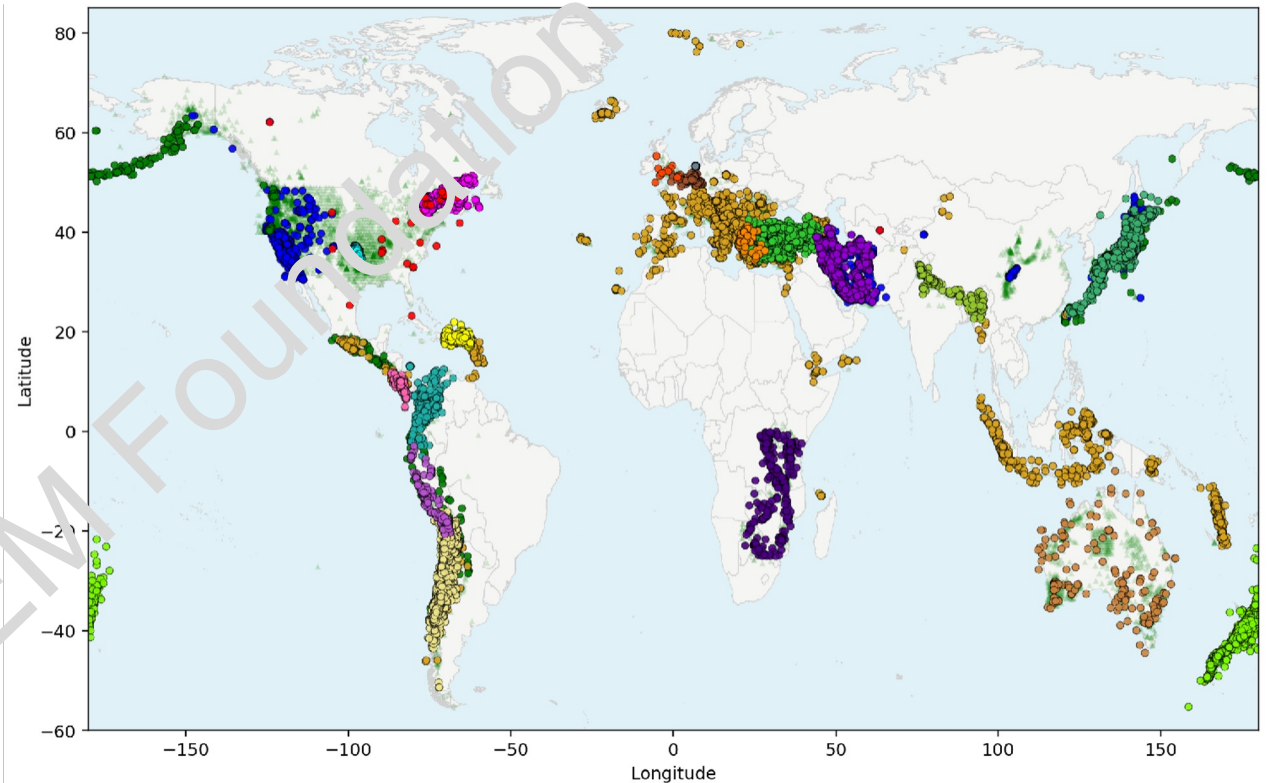
- Work is currently being undertaken for a partially non-ergodic GMM for each TRT in our next South America (SAM) model
- Given this is a continent-spanning PSHA model, the GEM Global Flatfile is being leveraged extensively
- Precomputing residuals using the SMT
- Jointly solving for source, site and path corrections using a constrained least-squared approach
- The non-ergodic adjustments are allocated to a densifying H3 grid
 - Enough data points above a given threshold = move to a finer resolution with a more robust correction
 - Cannot always densify for **all** cells in parent H3 cell, so retain parent cell even if densification occurs
 - And then use the parent cells **only** for events/sites which don't lie in the finer resolution cells



Location-based source effect (locality-to-locality) adjustment for **active crustal** events

Future Plans for the GEM Global Flatfile

- The flatfile will be available soon (this fall) on GEM's Zenodo page
- We plan to document the dataset in a journal article
- We are always interested in improving the database
- Suggestions on additional information to include, or datasets to be incorporated are always welcome
- We recommend using the OQ-MBTk's SMT module to fully leverage the benefits of this dataset



Thank you!

Please attribute to the GEM Foundation with a link to:
<https://www.globalquakemodel.org>



Except where otherwise noted, this work is licensed under a
Creative Commons Attribution-NonCommercial-NoDerivatives 4.0
International License (CC BY-NC-ND 4.0)
<https://creativecommons.org/licenses/by-nc-nd/4.0/>



June 23 – 25, 2026

University of Zagreb, Croatia

GEM CONFERENCE

FROM FAULTS TO FUTURE SCENARIOS



Stochastic EvEnt Sets At Worldwide Scale (SEESAWS)

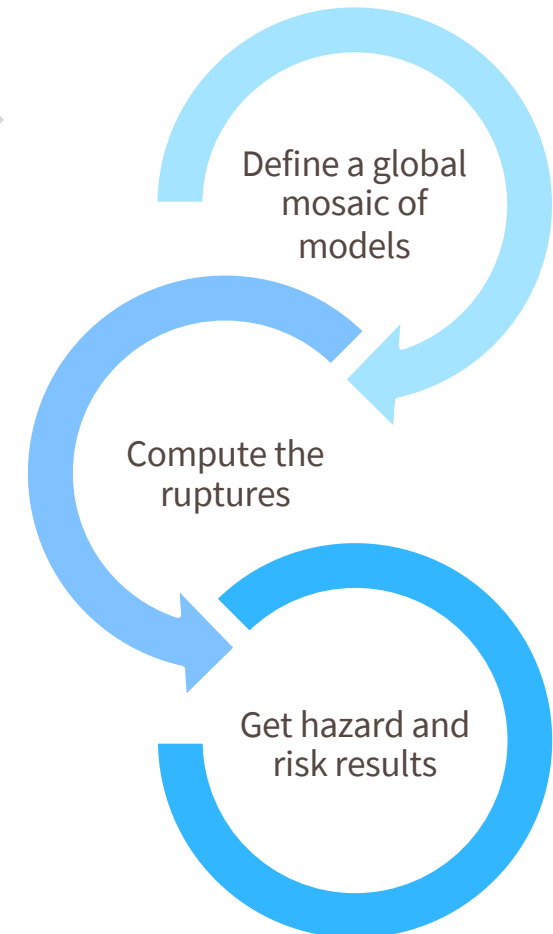
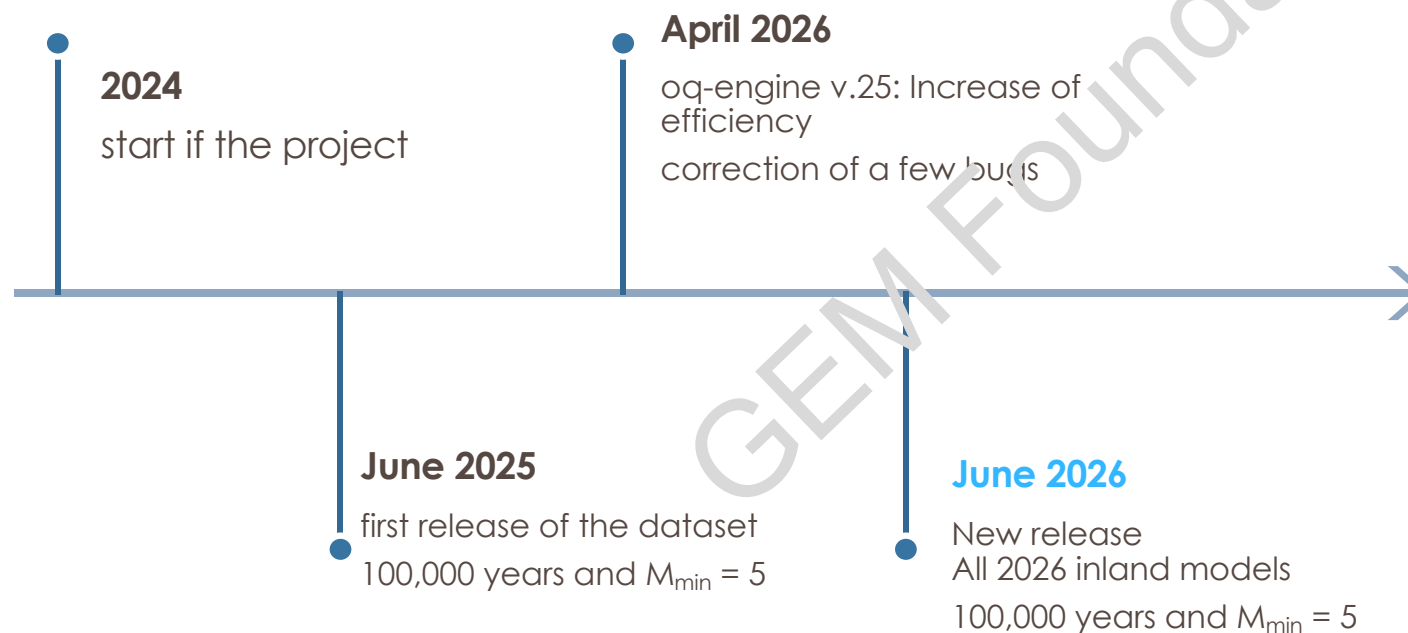
Manuela Villani, GEM Hazard team and Michele Simionato
GEM Hazard Team

June 23 - 25, 2026
University of Zagreb, Croatia

GEM CONFERENCE
FROM FAULTS TO FUTURE SCENARIOS

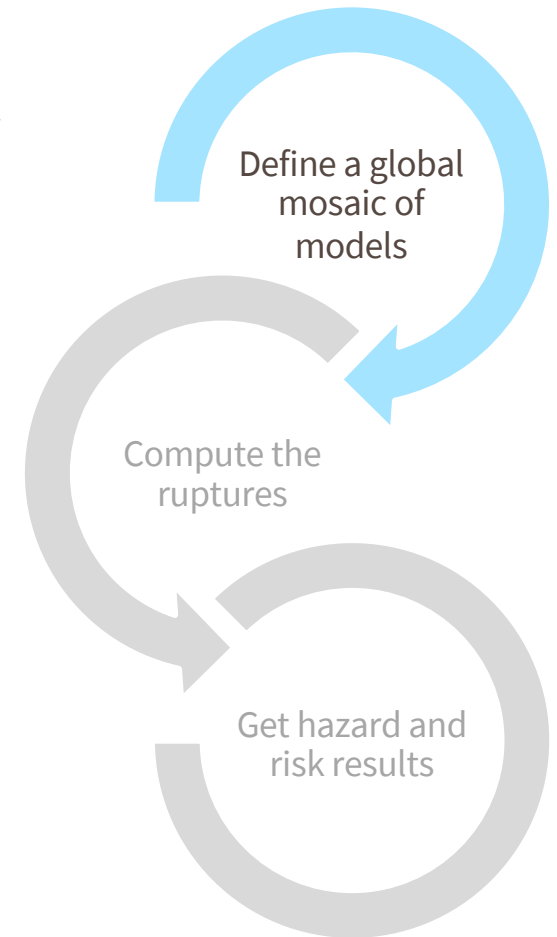
The project

Aim: creating a global dataset of ruptures that could be used in hazard and risk calculations to generate ground motion fields (GMFs) at a global scale.

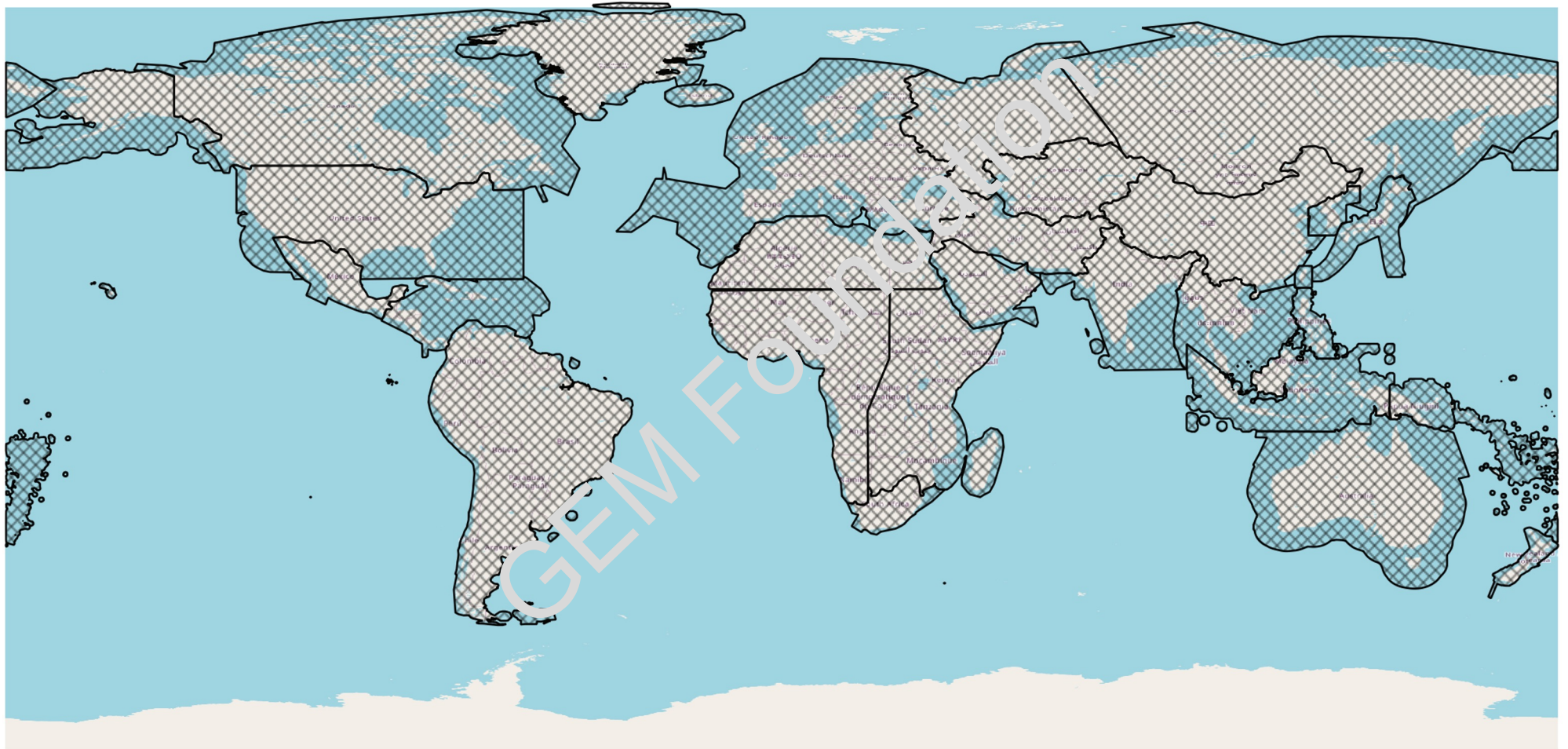


The mosaic of models

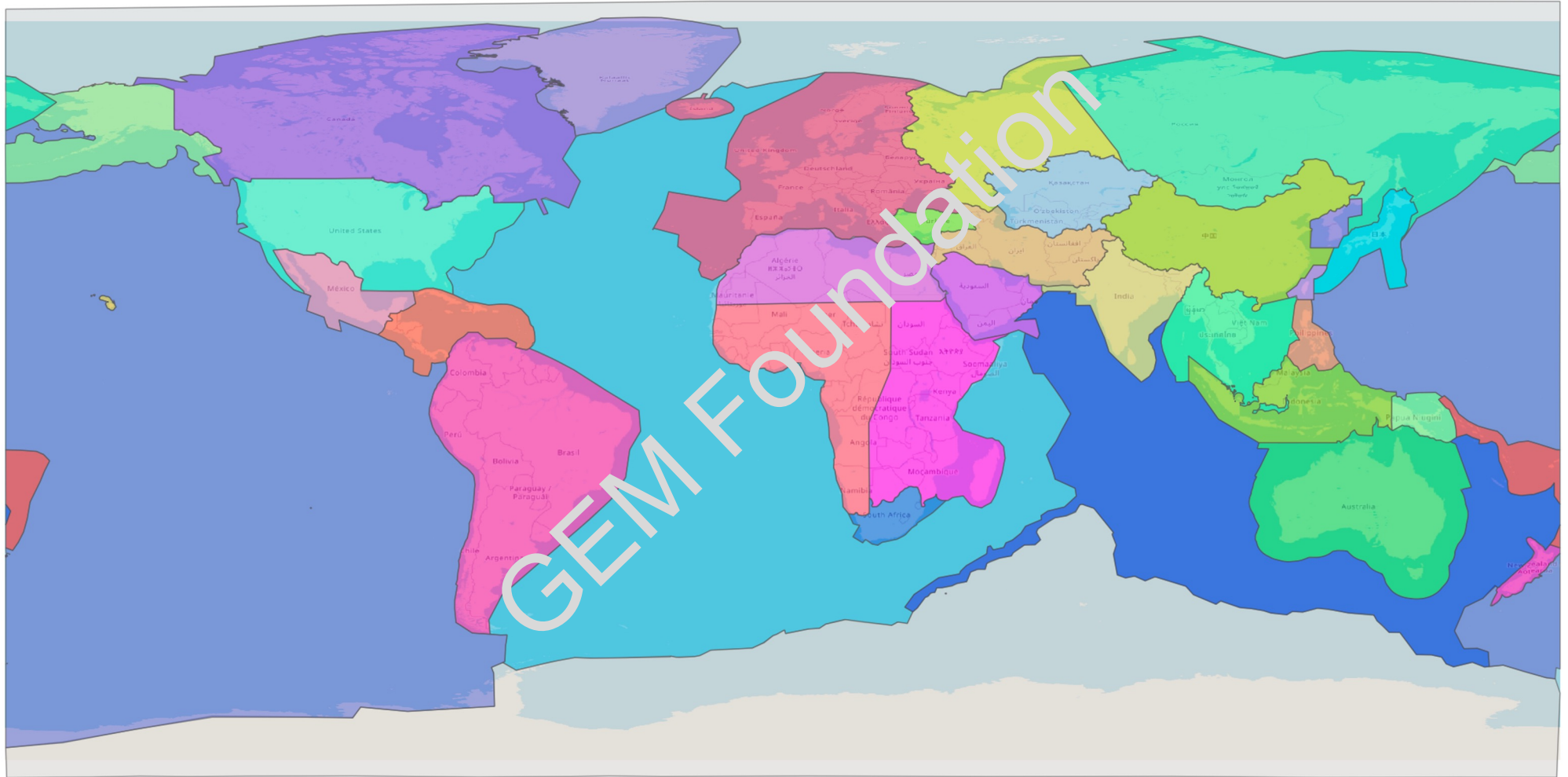
- How we choose the ruptures



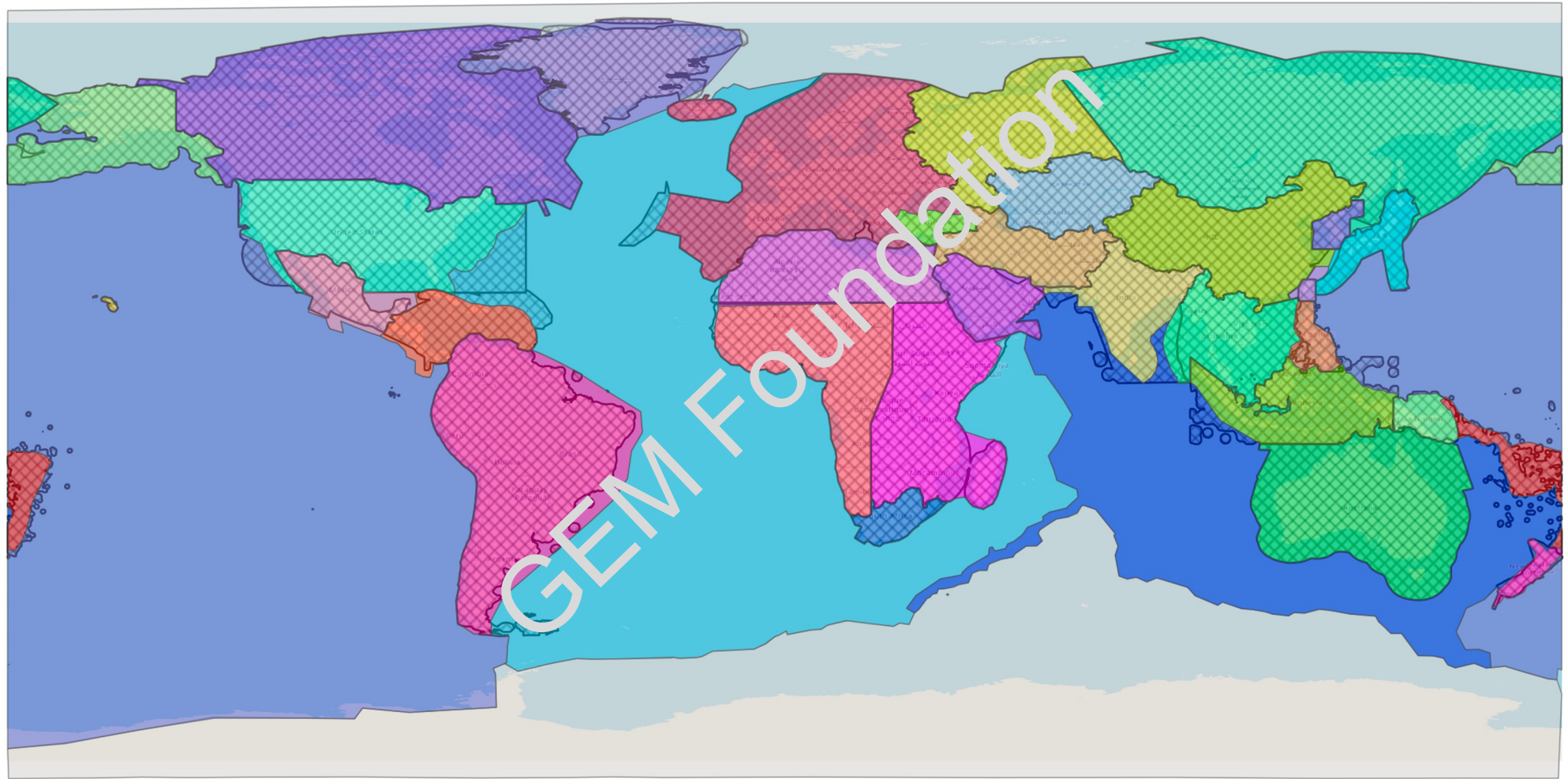
SEESAWS Inland - June 2025



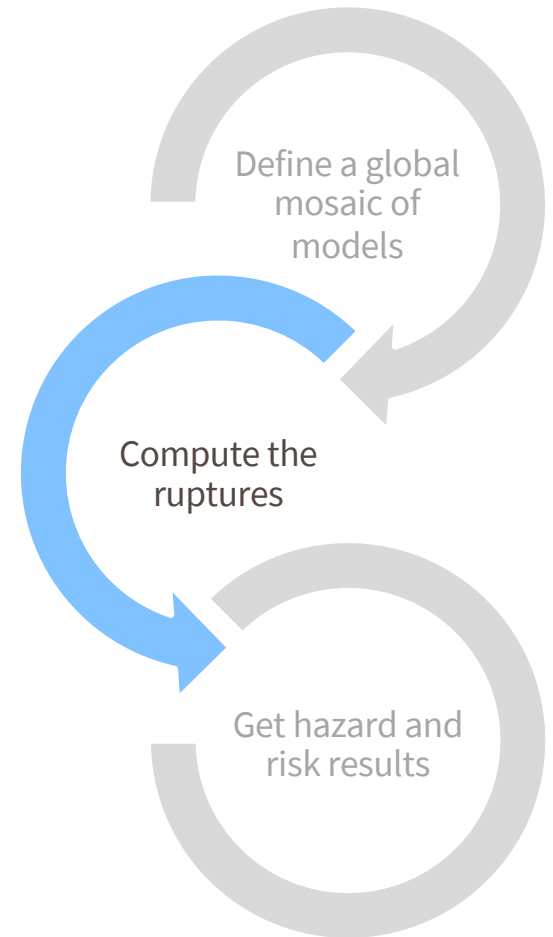
SEESAWS Global - June 2026



SEESAWS - June 2025 vs. June 2026

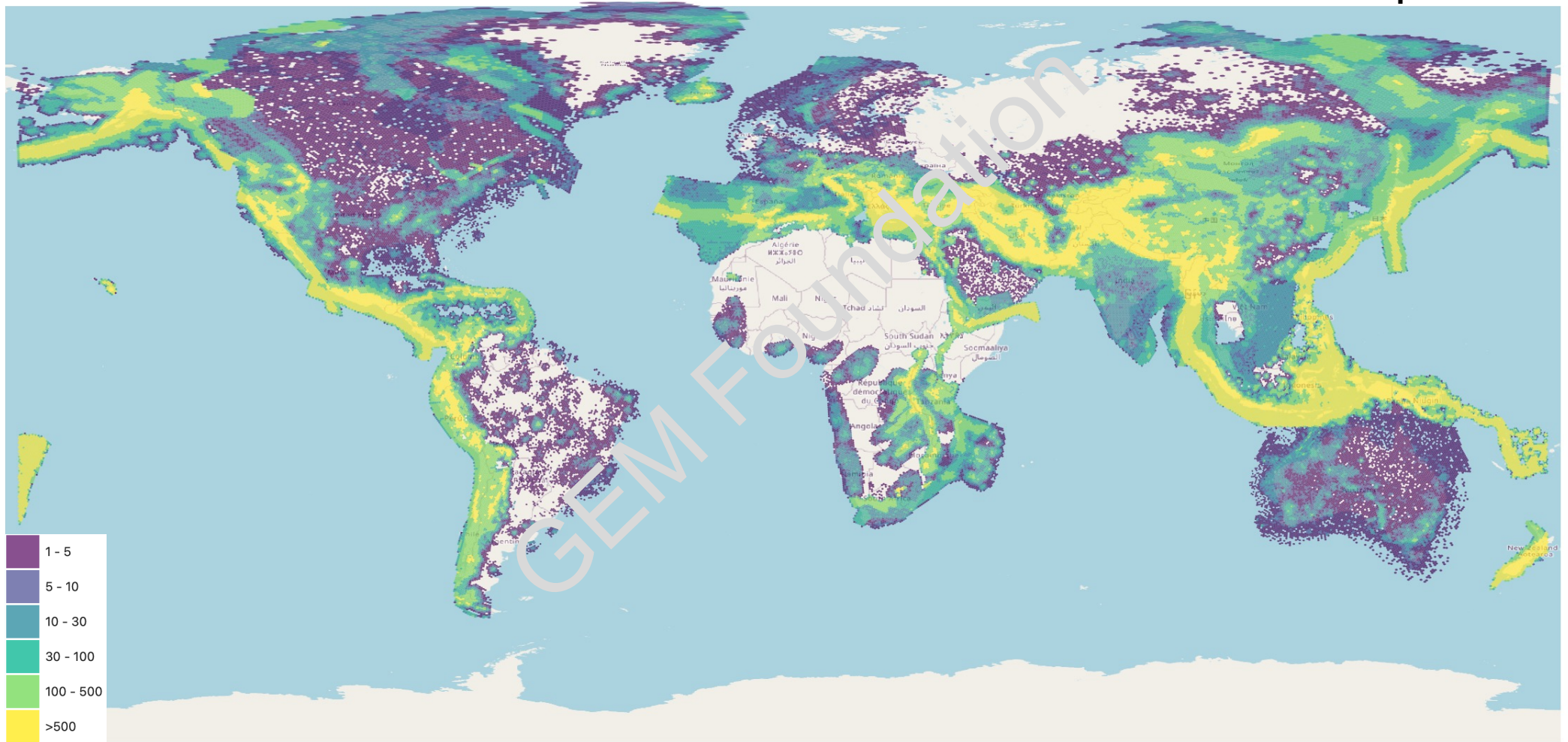


The dataset of ruptures



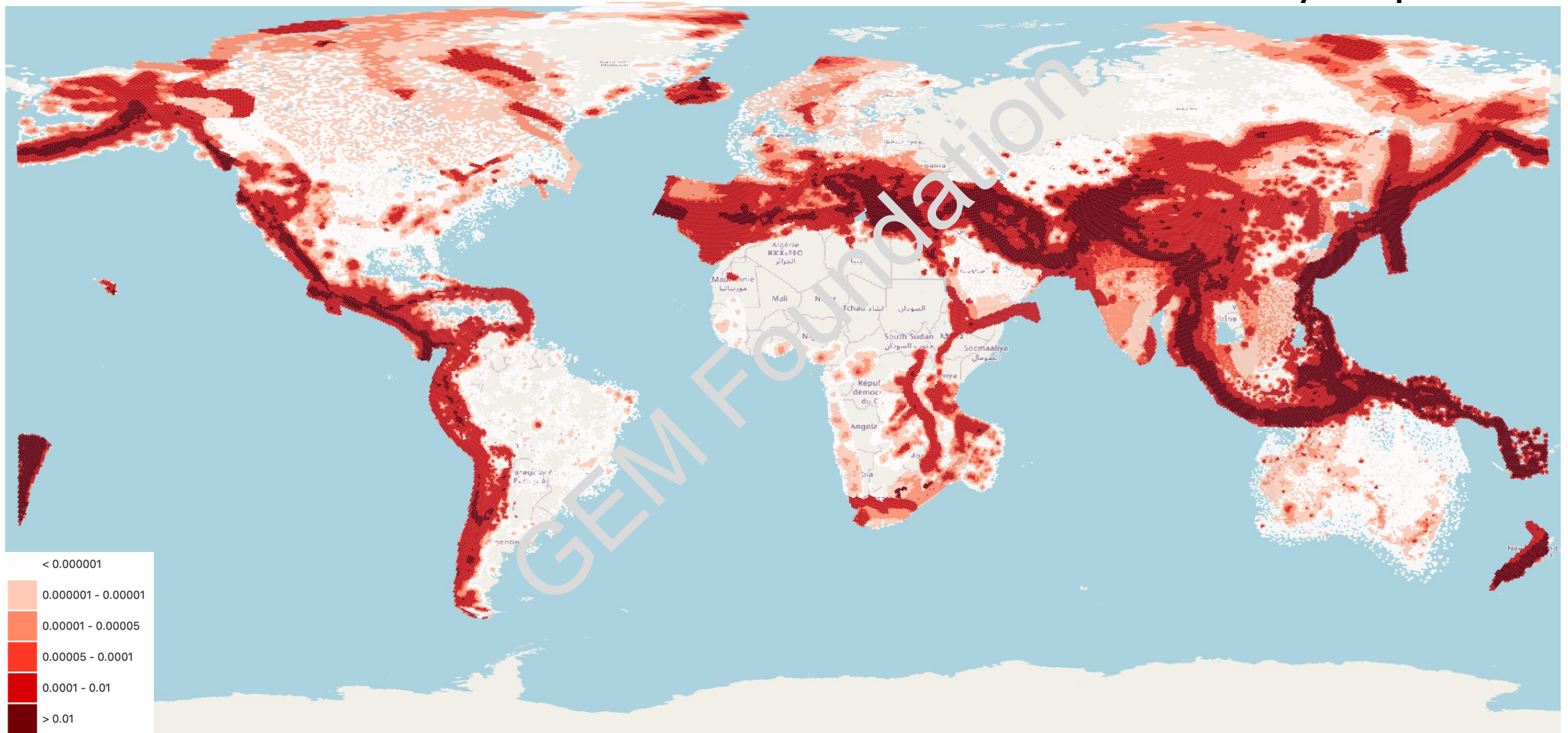
Ruptures dataset v2026 ($M_{\min} = 5 - 100,000$ years)

Number of events per h3 cell



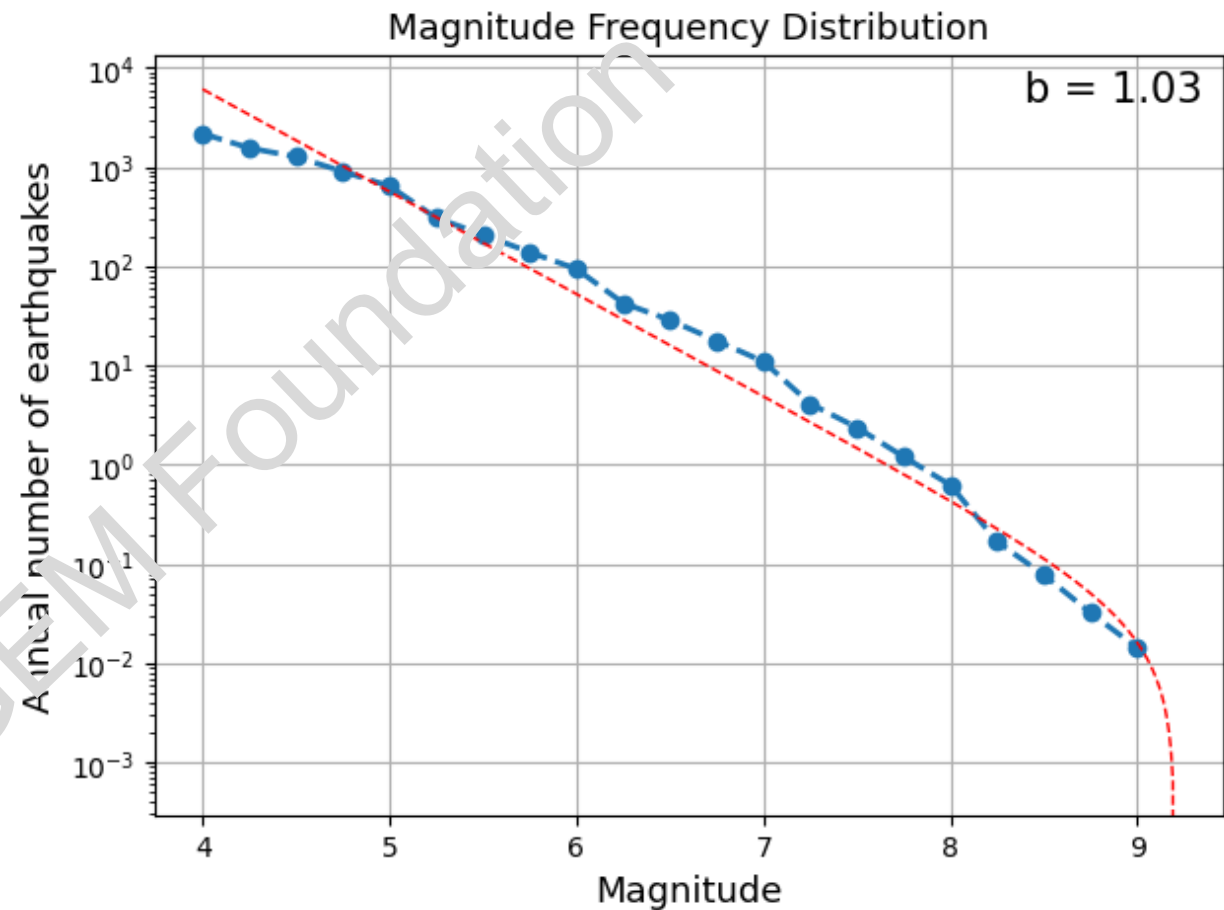
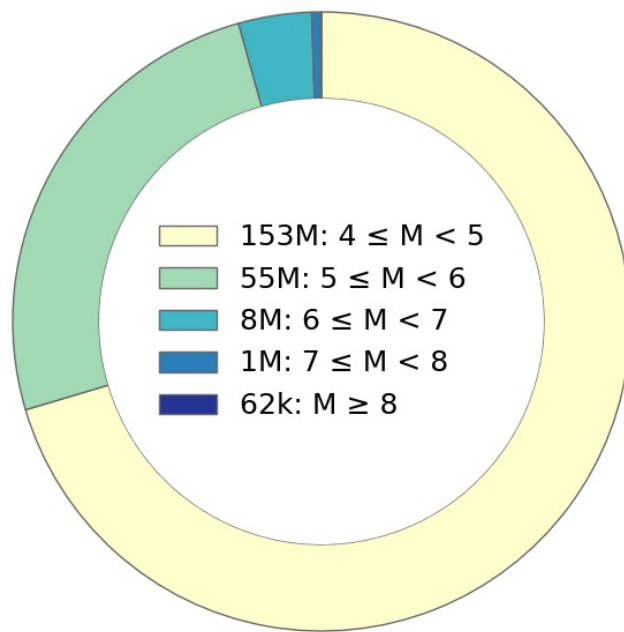
Ruptures dataset v2026 ($M_{\min} = 5 - 100,000$ years)

Activity rate per h3 cell



Global dataset of ruptures ($M_{\min} = 4 - 100,000$ years)

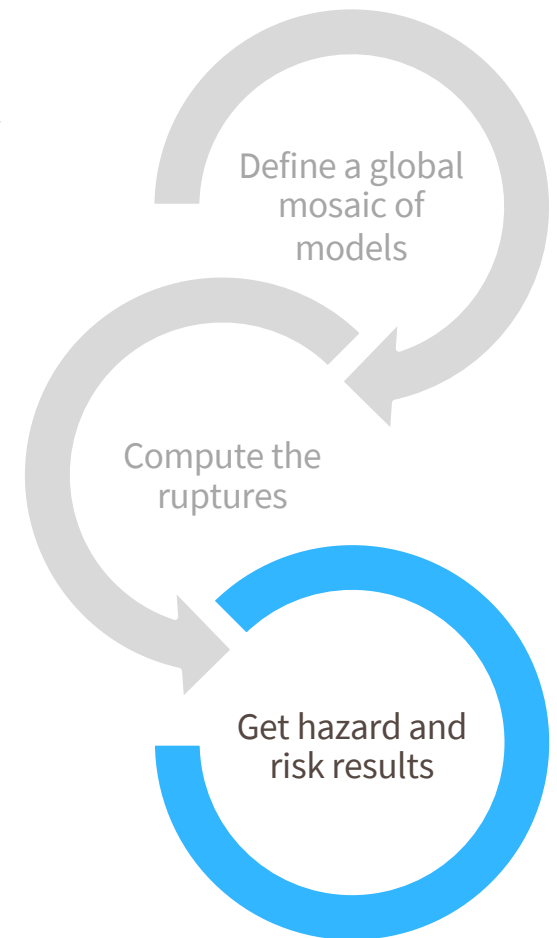
Ruptures in 100,000 years SES



Dataset > 50Gb

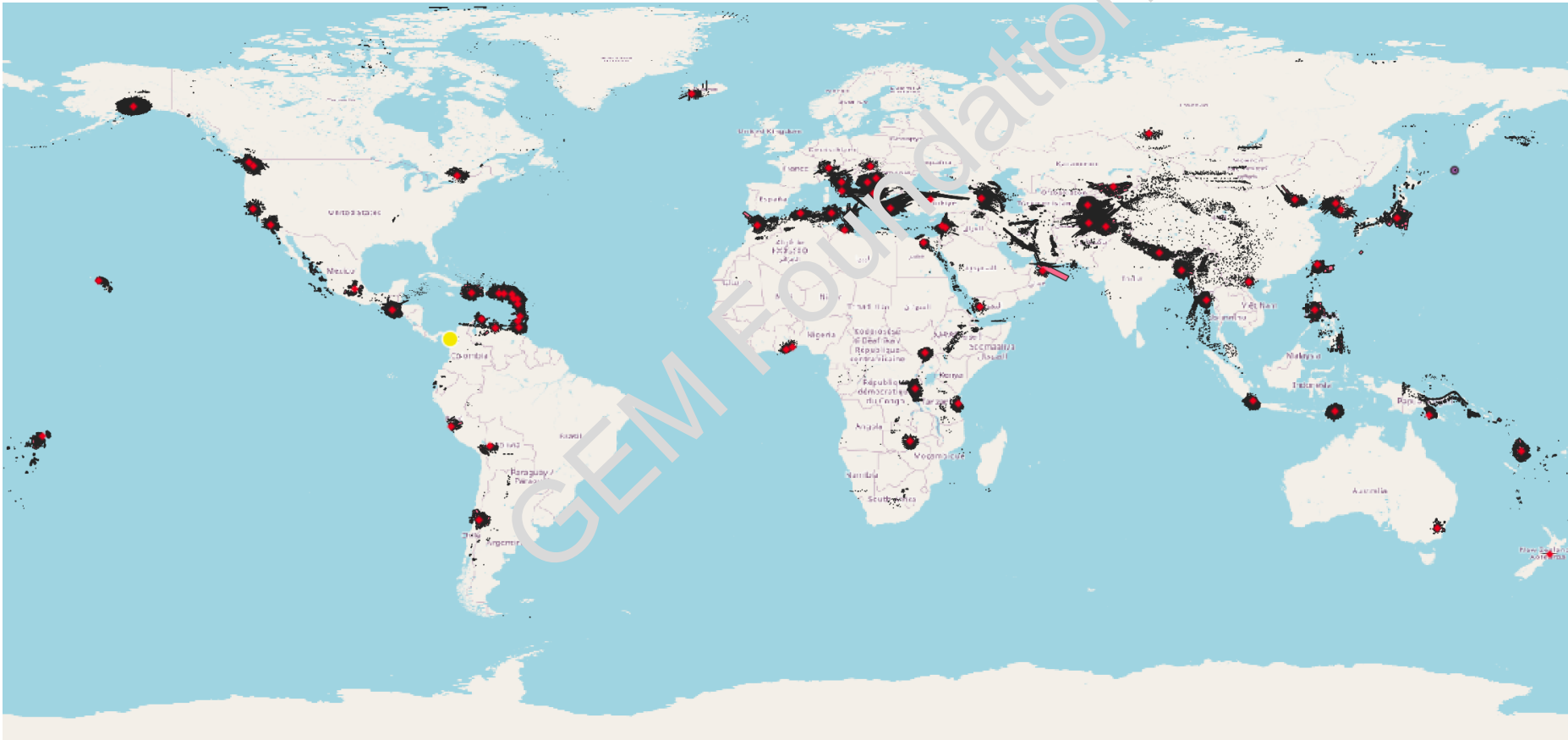
Applications

- Sites from multiple models
- Global seismic hazard map
- Global seismic risk map

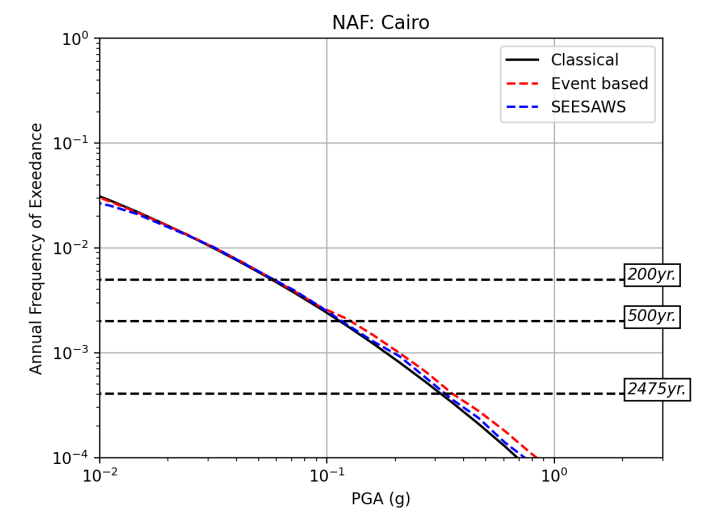
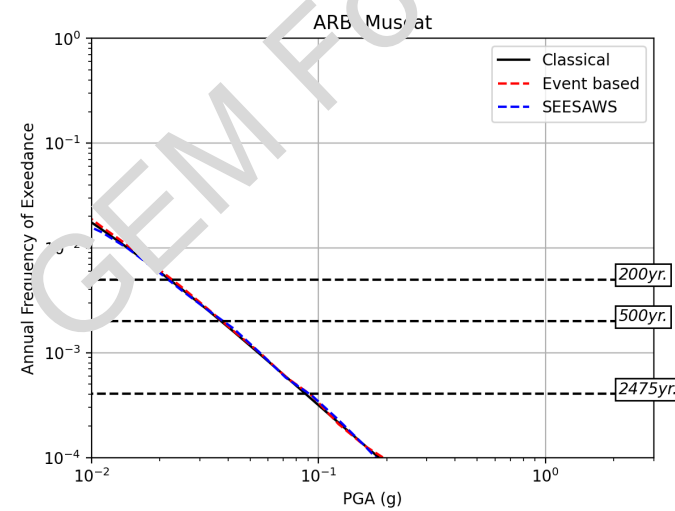
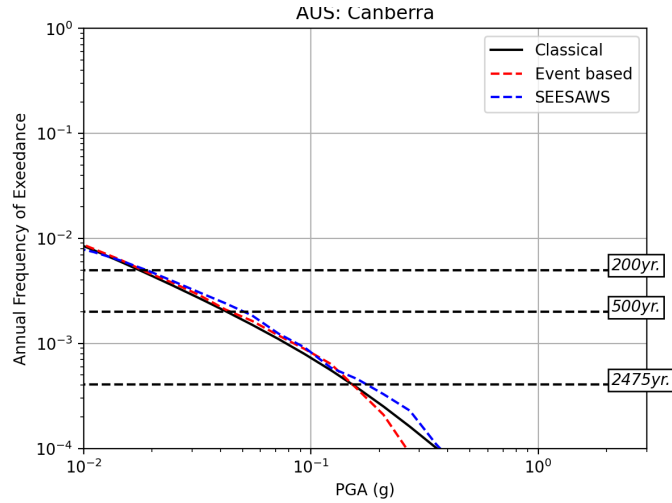
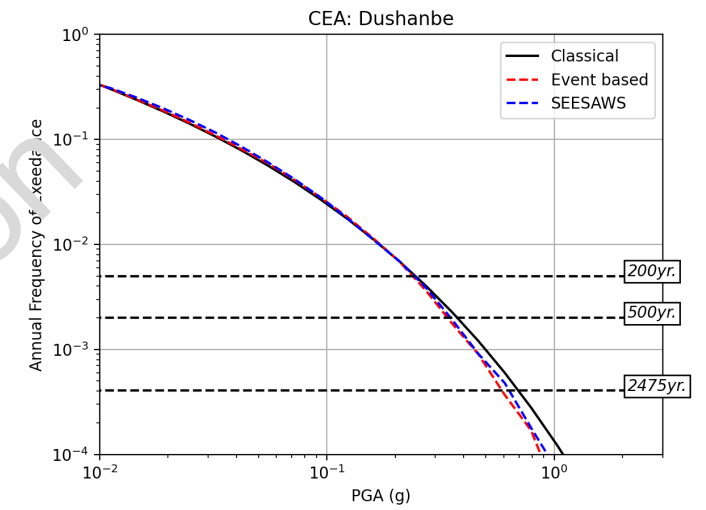
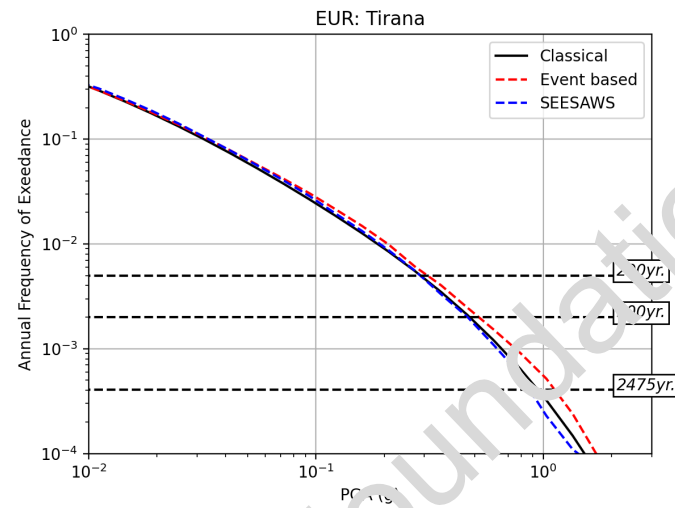
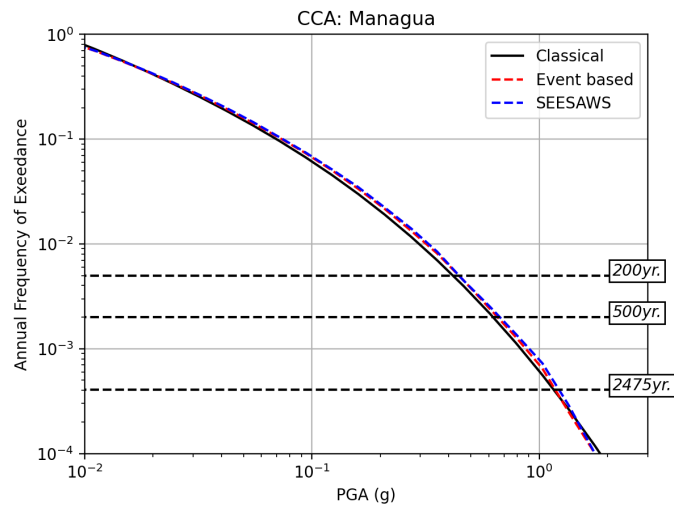


Run sites from multiple models

A key advantage of the SEESAWS dataset is its ability to support the computation of the hazard (and consequently risk) in a single hazard calculation for sites distributed globally or for sites belonging to different hazard models.

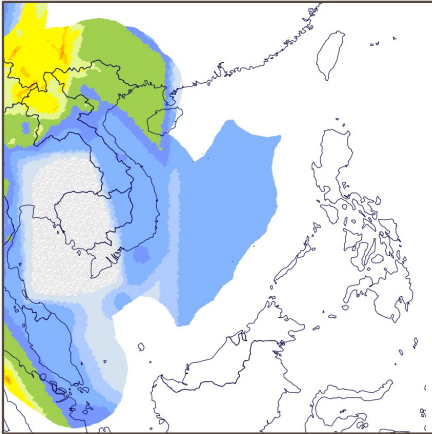


Verification

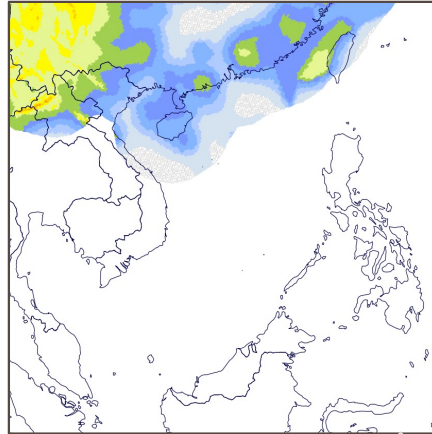


SEESAWS in the global seismic hazard map

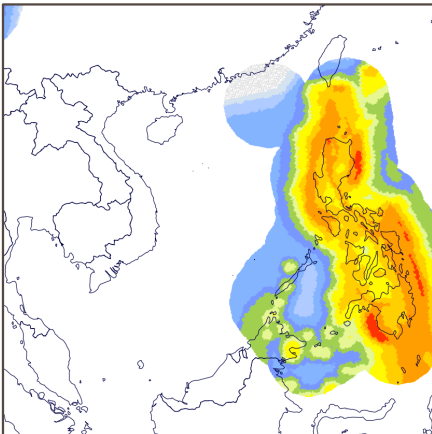
SEA



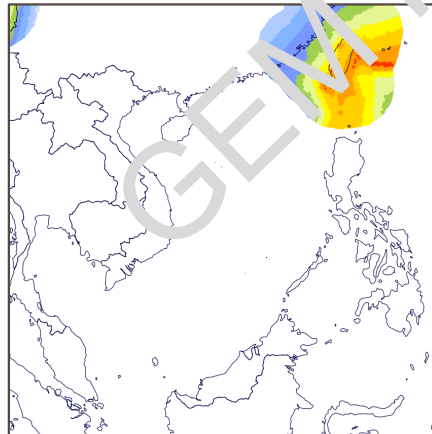
CHN



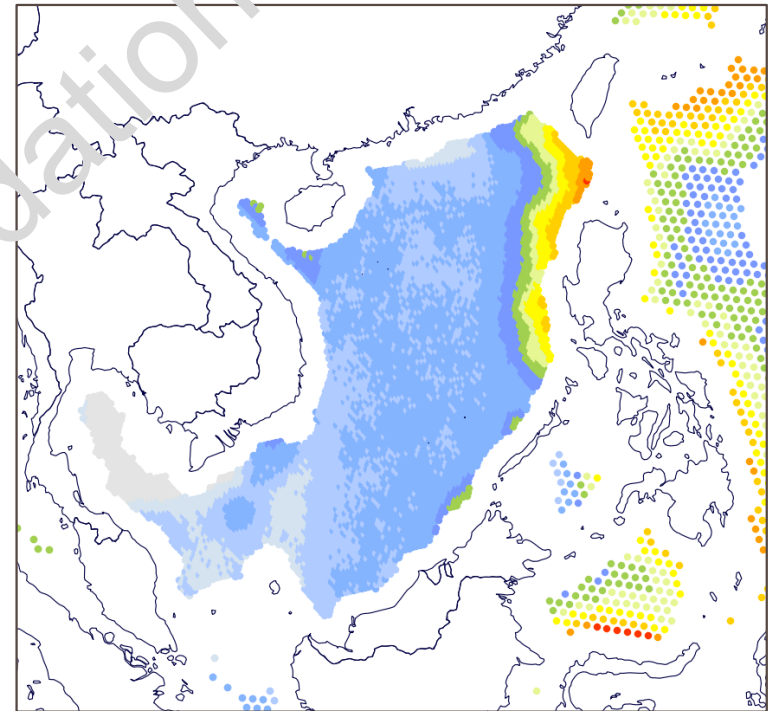
PHL



TWN



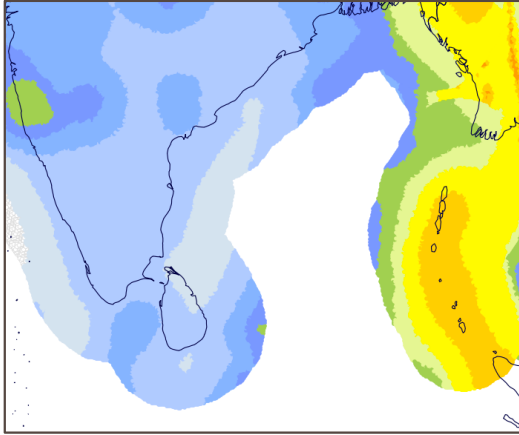
PGA resulting from SEESAWS



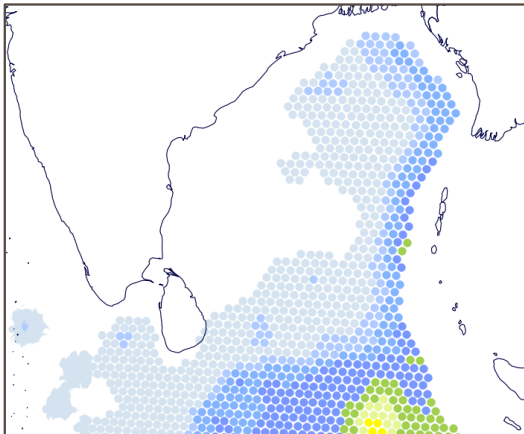
SEESWAS computations account for ruptures from different models providing harmonized hazard results.

SEESAWS in the global seismic hazard map

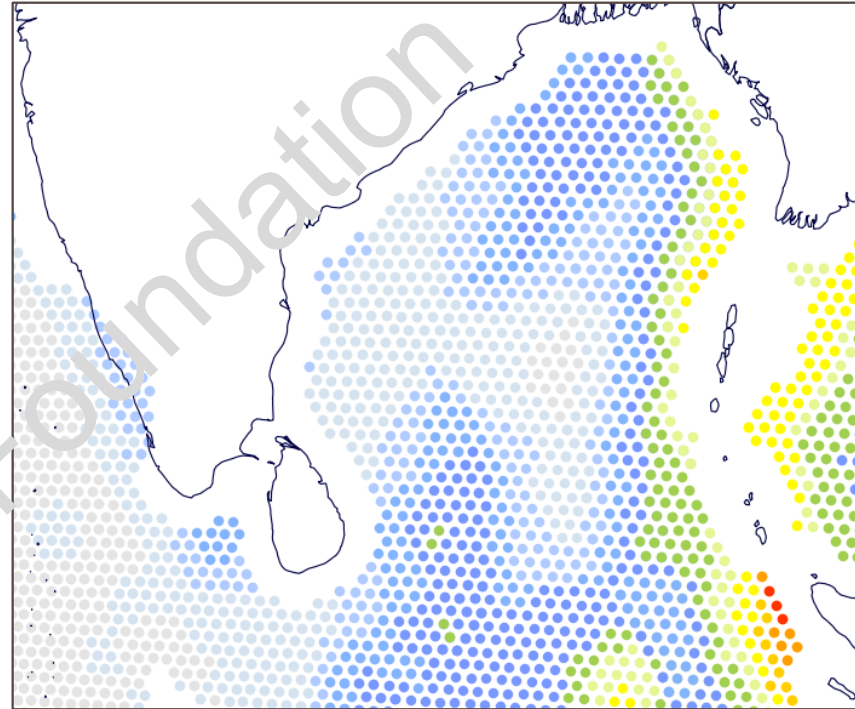
IND



OIN



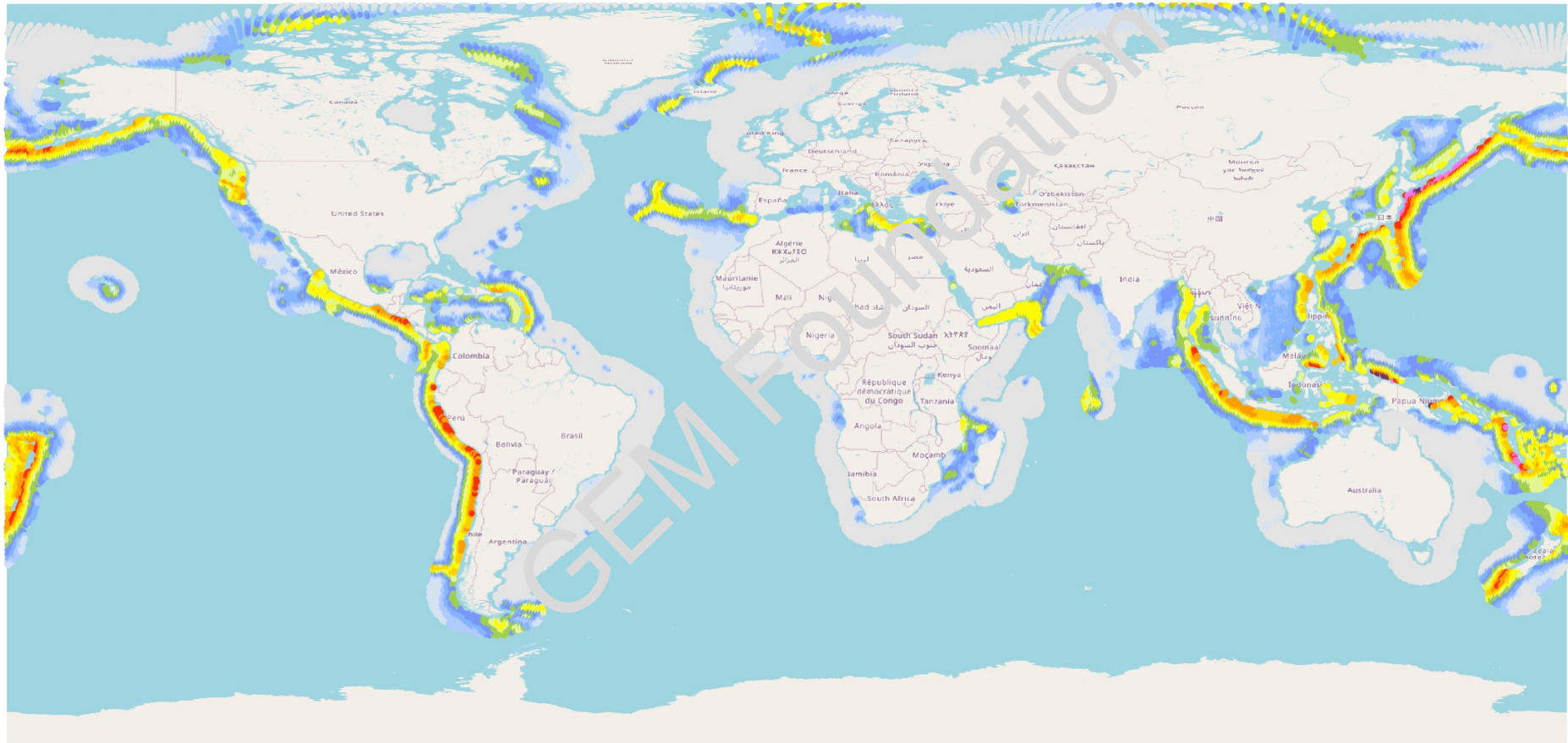
PGA resulting from SEESAWS



The oceans models were built, along the coast, to complement the inland models. SEESWAS computations allow to include all contributions in the transition between offshore and inland

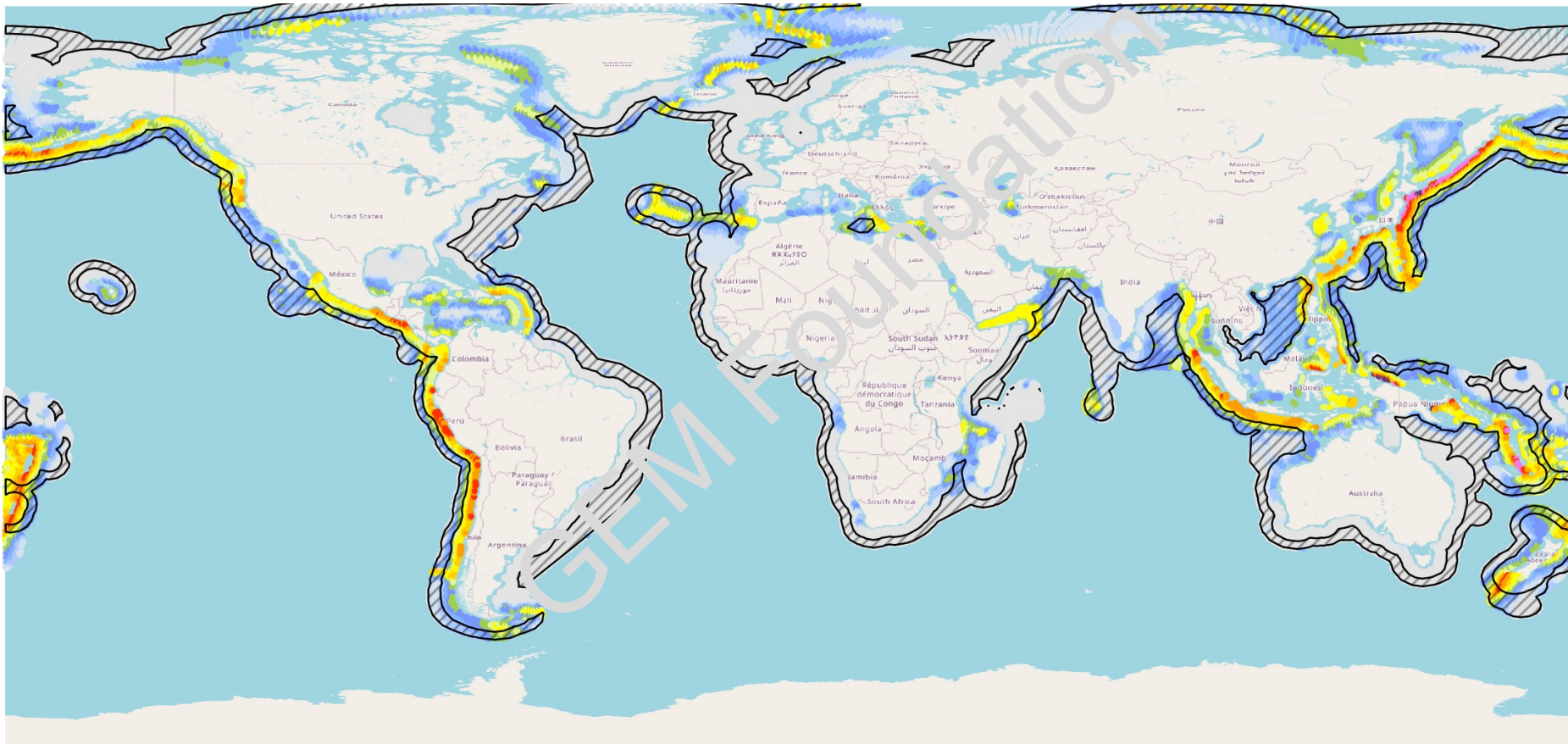
SEESAWS in the global seismic hazard map

SEESAWS was adopted for “tricky” offshore areas and all offshore sites between 300-500 km



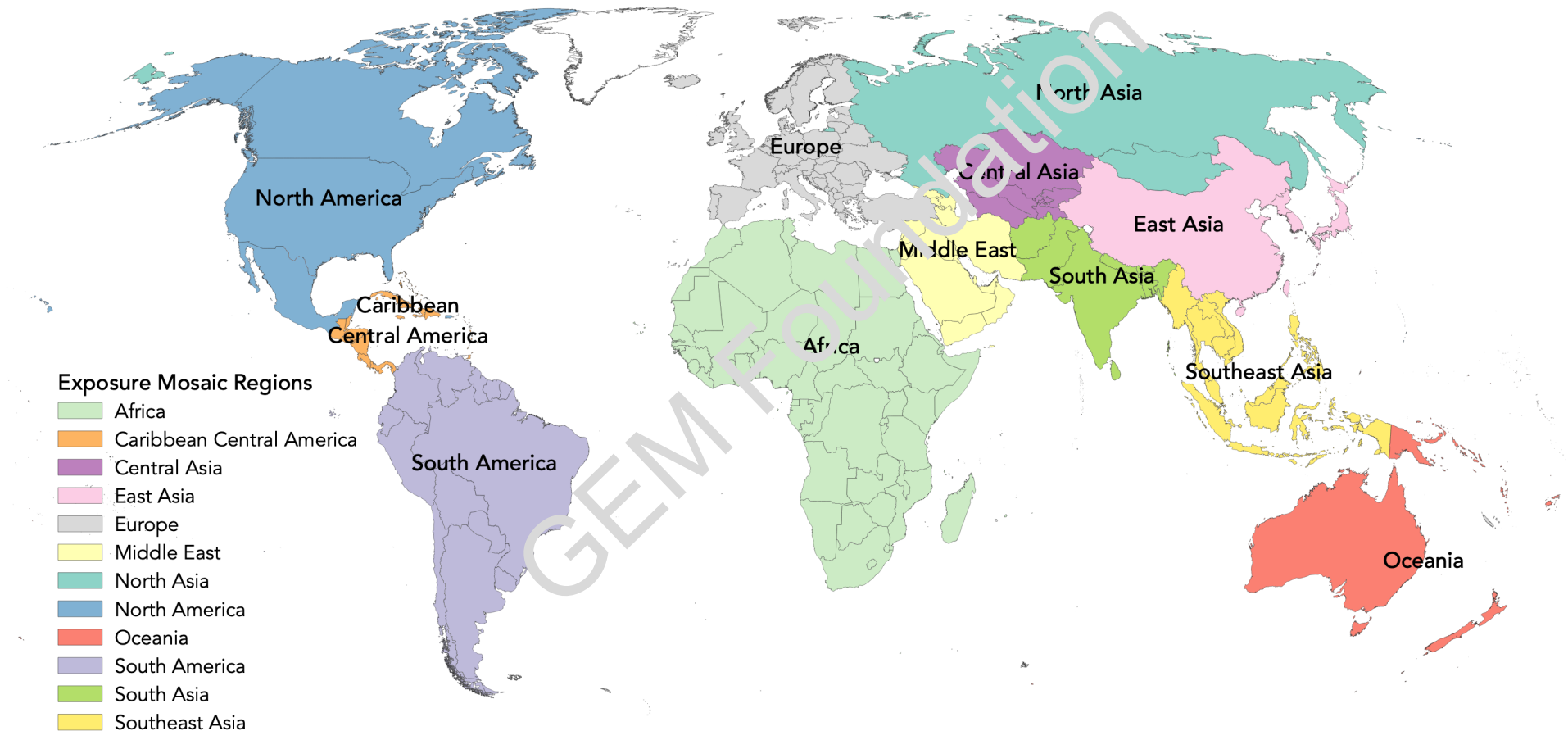
SEESAWS in the global seismic hazard map

SEESAWS was adopted for all “tricky” areas and all offshore sites between 300-500 km



SEESAWS in the global seismic risk map

SEESAWS was adopted to run risk using regions, disentangling the process from any specific model.



Available now

- New ruptures dataset for all inland models fully consistent with the 2026 release
- `global_ses.py`: creation of a ruptures dataset with command:

```
oq shell openquake.engine.global_ses --number-of-logic-tree-samples N --ses-per-logic-tree-path S --minimum-magnitude M mosaic_dir ses.hdf5
```
- Post processing tools (jupyter notebooks): create a geojson file, filter and analyse the ruptures dataset, analyse a Event-based datastore created with SEESAWS.
- Run an event based calculation with input the ruptures dataset through the

```
oq engine --run job.ini --hc="ruptures_dataset.hdf5"
```

What's next?

- Improve the generation of ruptures along the border between models (e.g., buffer within which ruptures from models are considered)

Thank you!

Please attribute to the GEM Foundation with a link to:
<https://www.globalquakemodel.org>



Except where otherwise noted, this work is licensed under a
Creative Commons Attribution-NonCommercial-NoDerivatives 4.0
International License (CC BY-NC-ND 4.0)
<https://creativecommons.org/licenses/by-nc-nd/4.0/>



June 23 – 25, 2026

University of Zagreb, Croatia

GEM CONFERENCE

FROM FAULTS TO FUTURE SCENARIOS

