

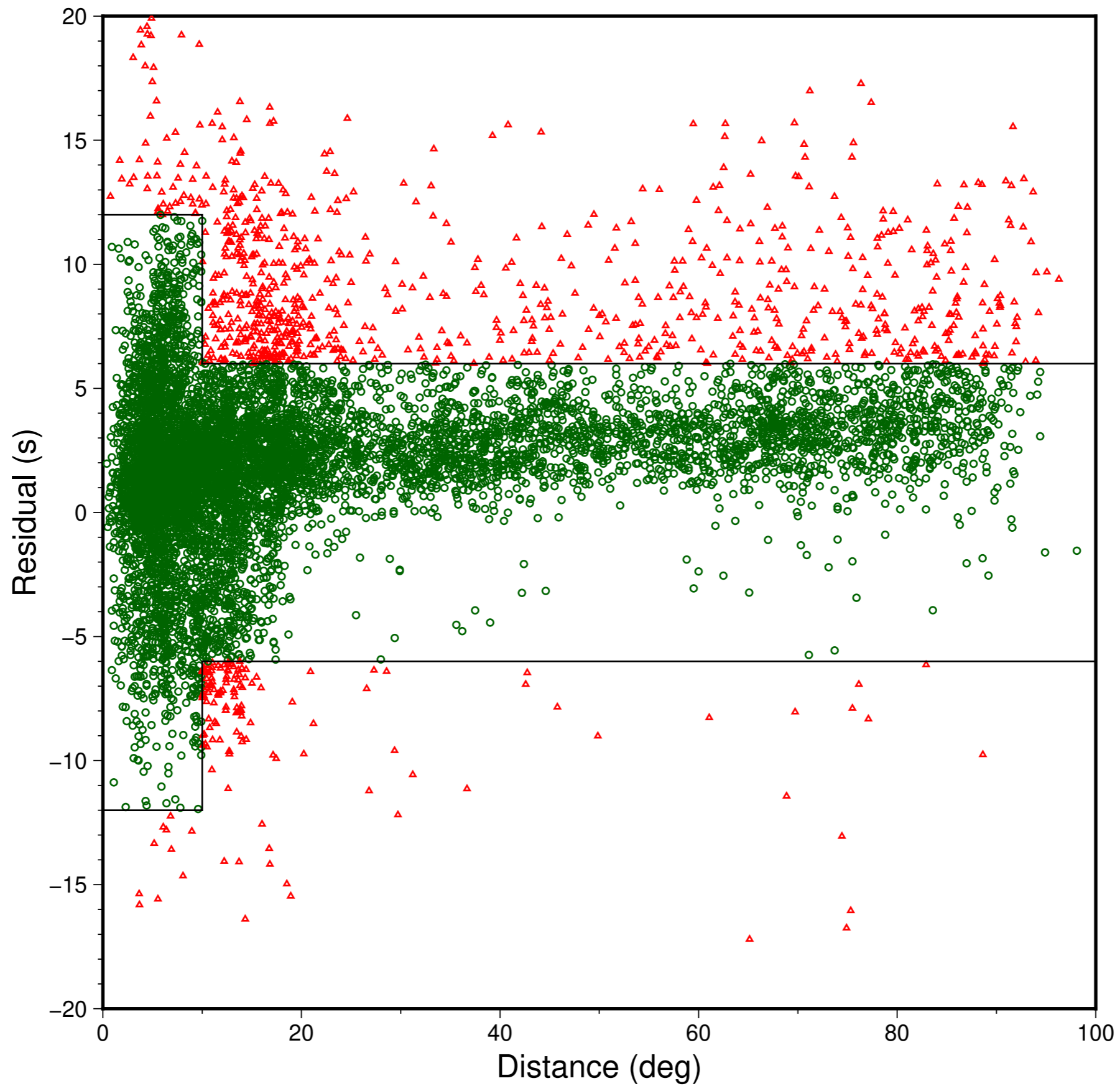
# SPPIM

First report: exploratory inversion experiments

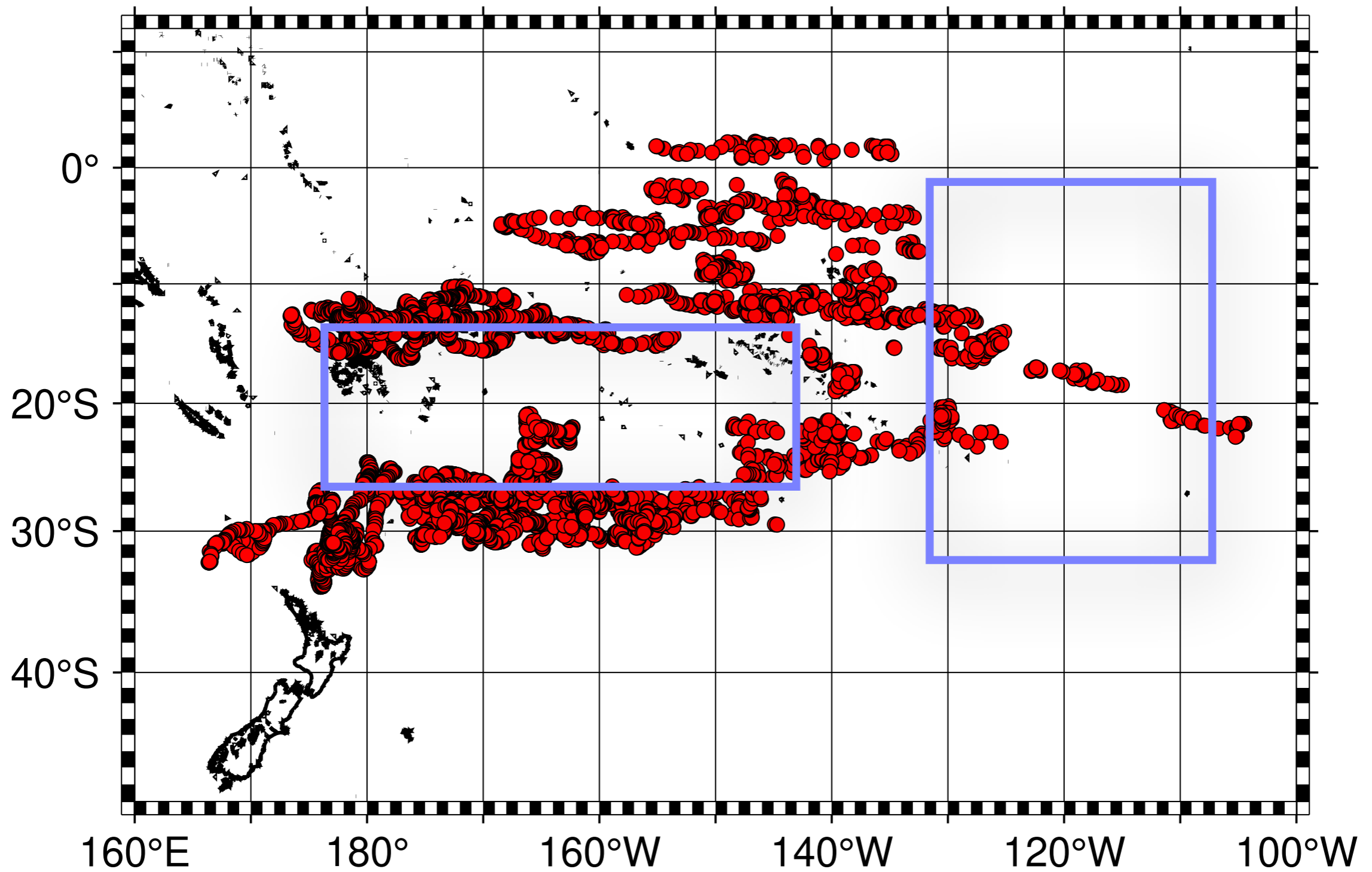
**ESO Meeting**  
**Jan 8, 2024**



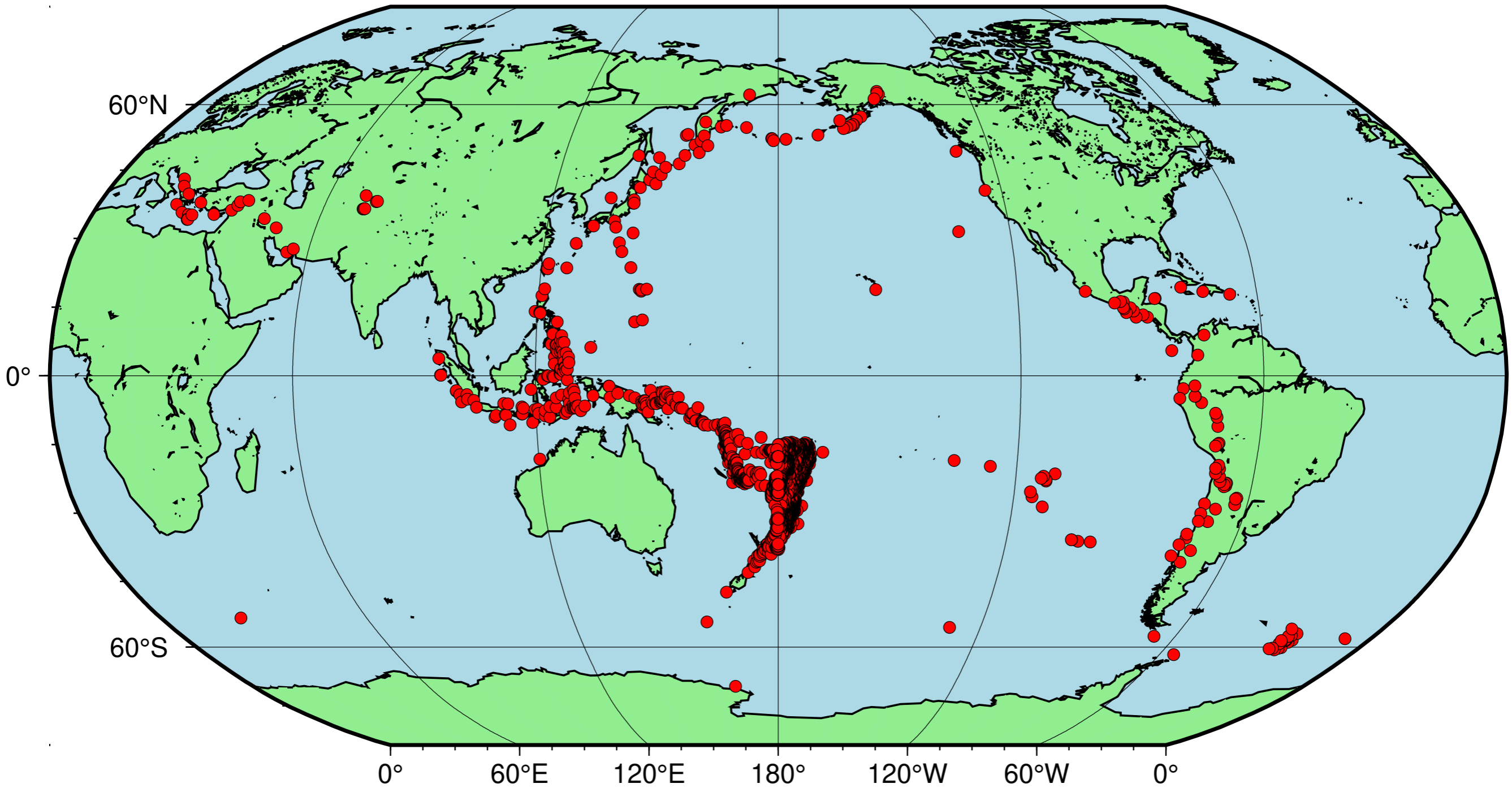
# Prior identification of outliers



# 8102 Seismogram locations



# 2928 event locations: less than 3 seismograms/event...

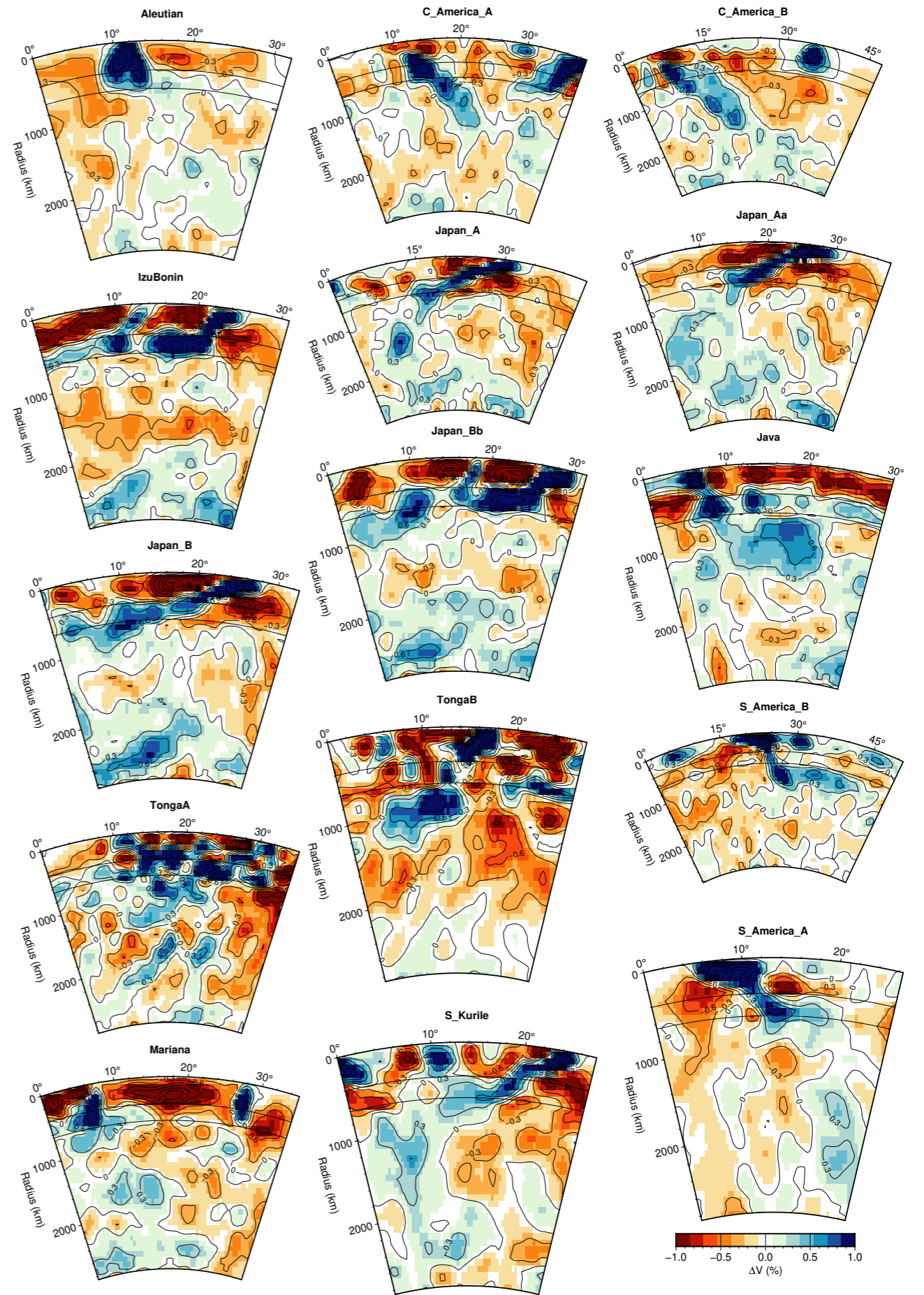
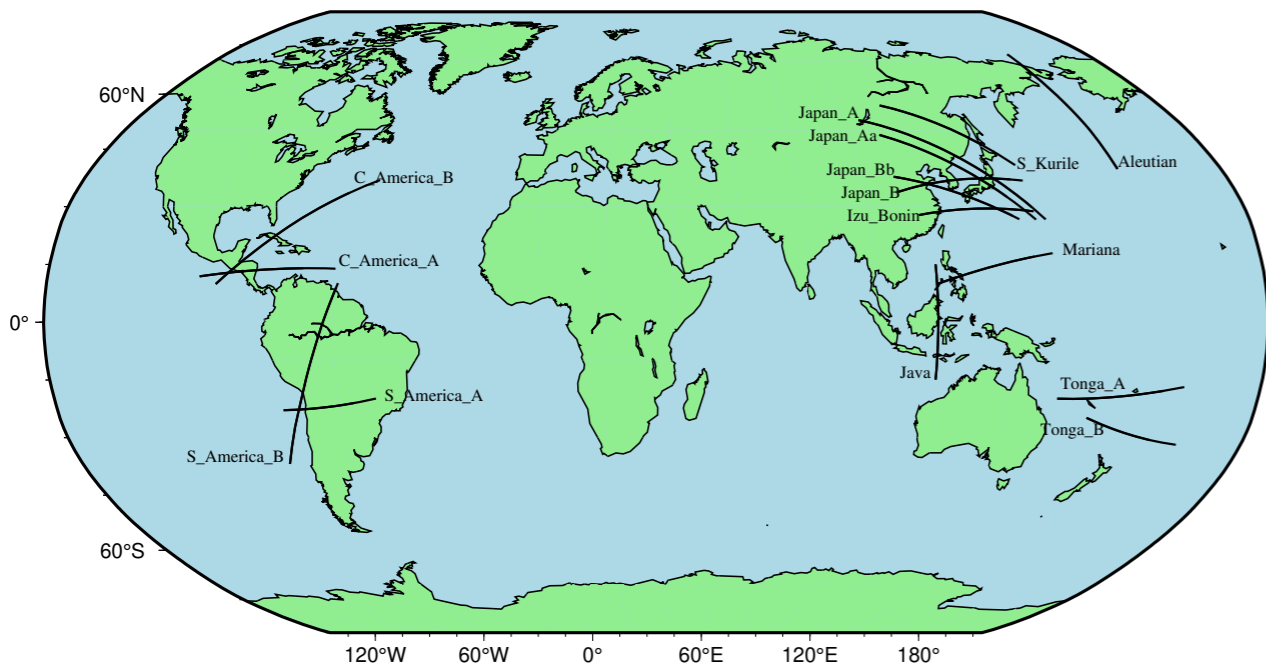




**The SPPIM data are embedded in a global inversion for  $V_p$**

Data set	N
A. SPPIM Mermaid (P,PKP)	8102
B. Associated ISC/NEIC picks (P,Pn)	712132
C. Summary ray ISC-EHB delays (P,PKP)	8601853
D. Finite-frequency delays (P,pP,PP,PcP,PKP)	177603
E. Galapagos Mermaid and local stations (P,PKP)	5278

# The global inversion was successful at the first attempt





# Ray density under the Pacific (log scale)

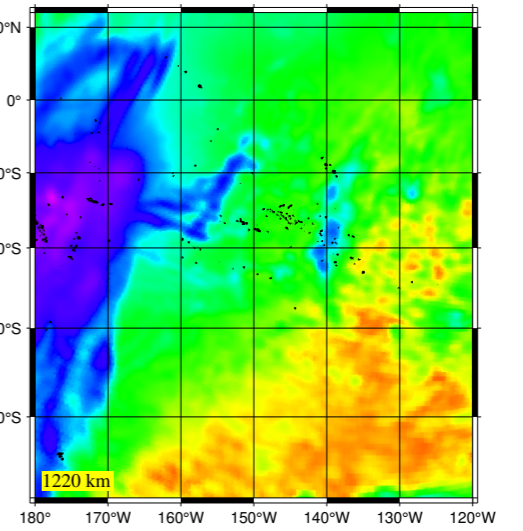
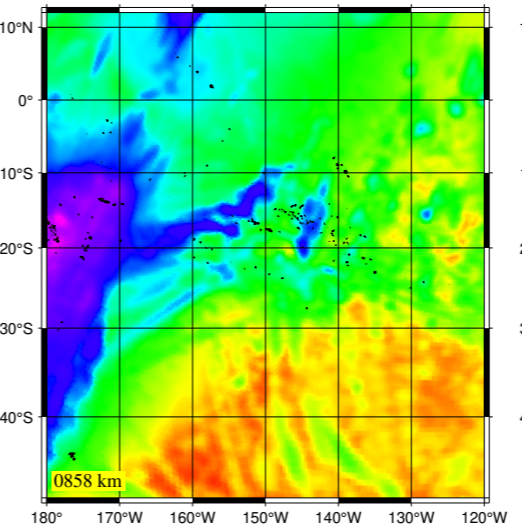
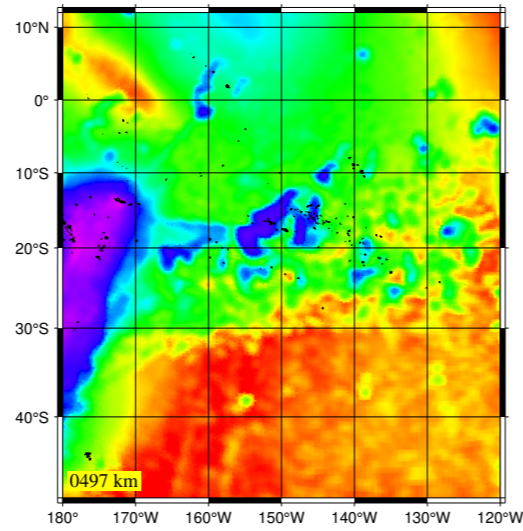
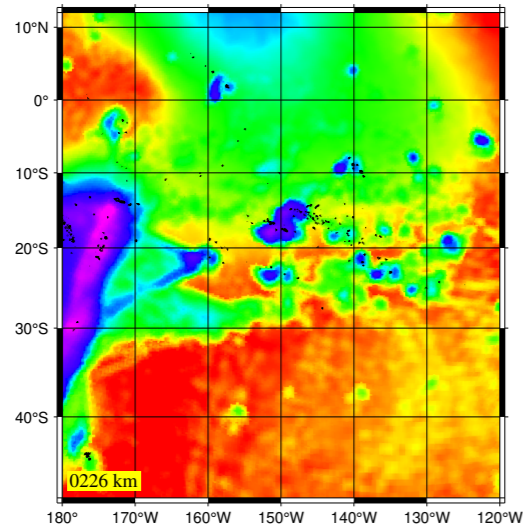
226 km

497 km

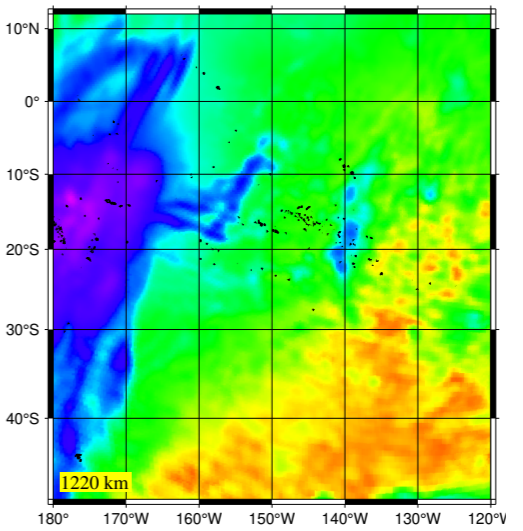
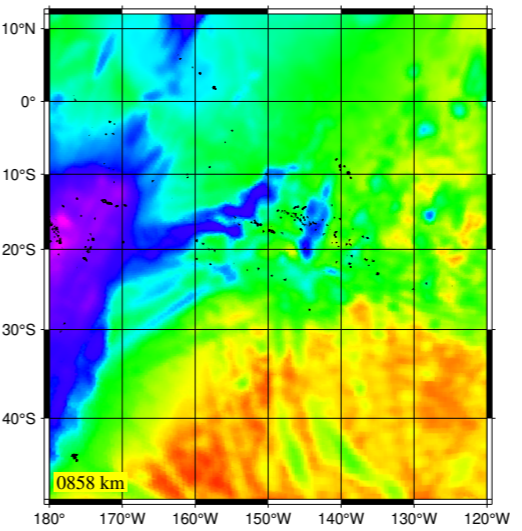
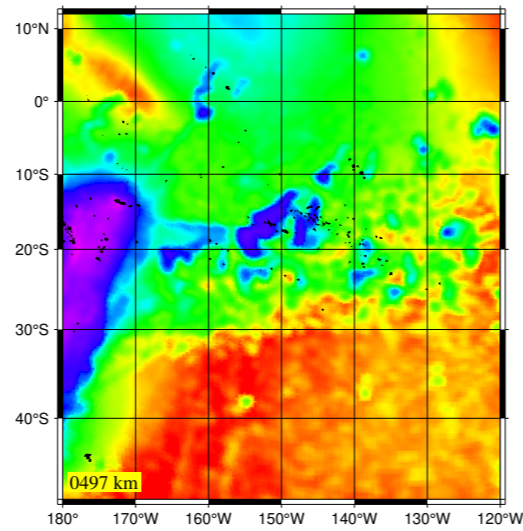
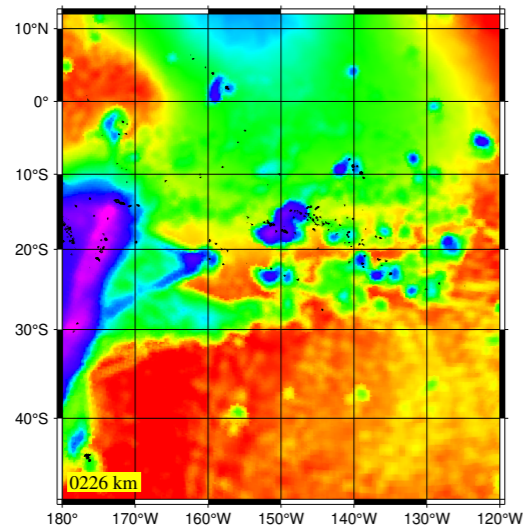
858 km

1220 km

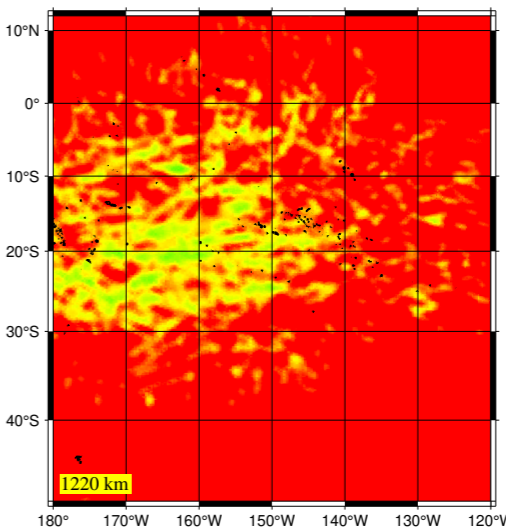
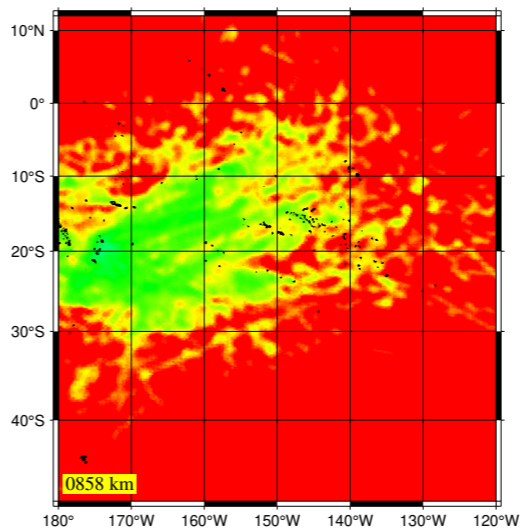
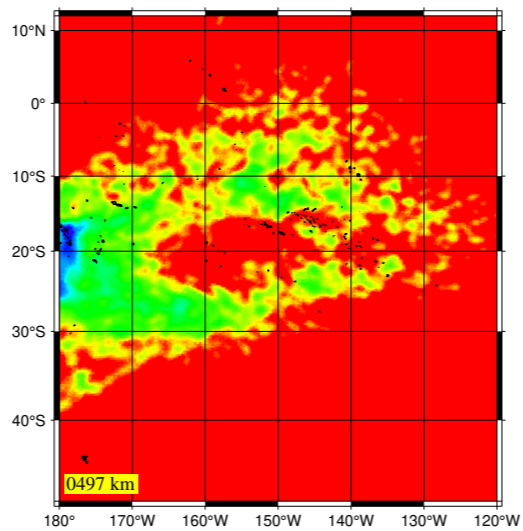
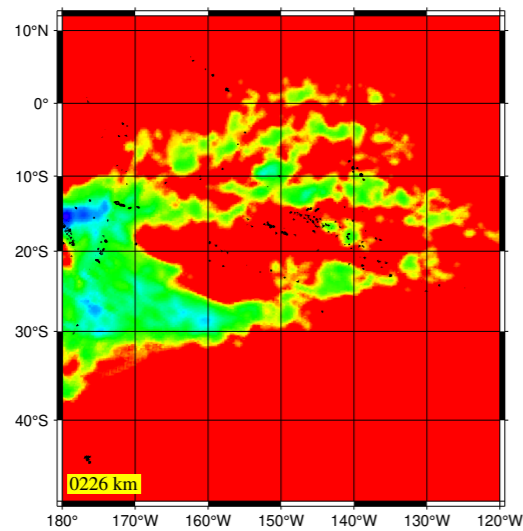
All data



No Mermaids



Mermaids only



# Solutions with (A) and without (B) SPPIM data

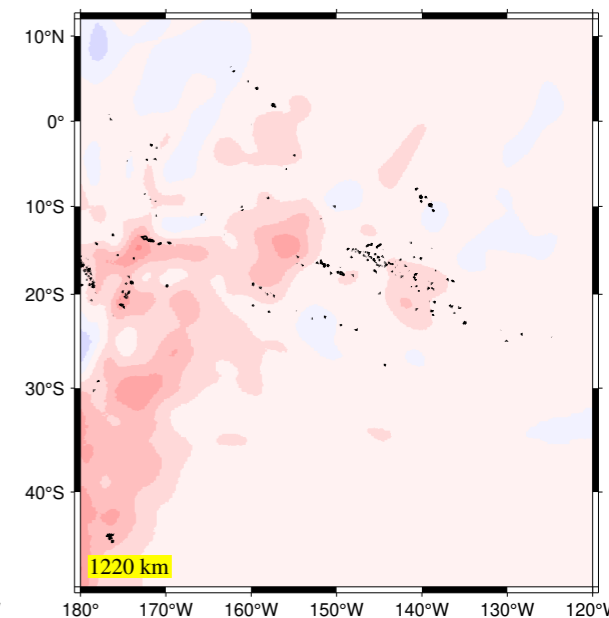
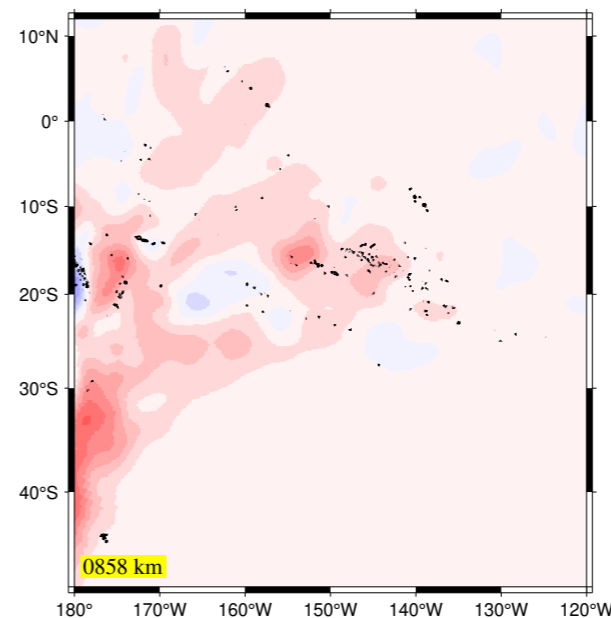
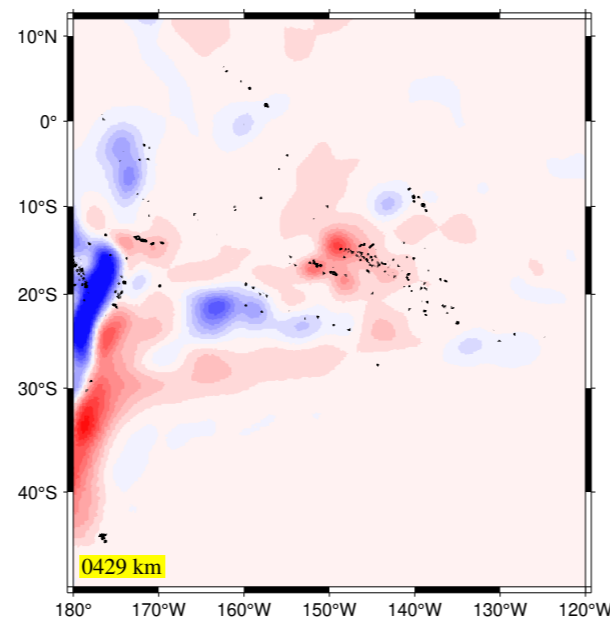
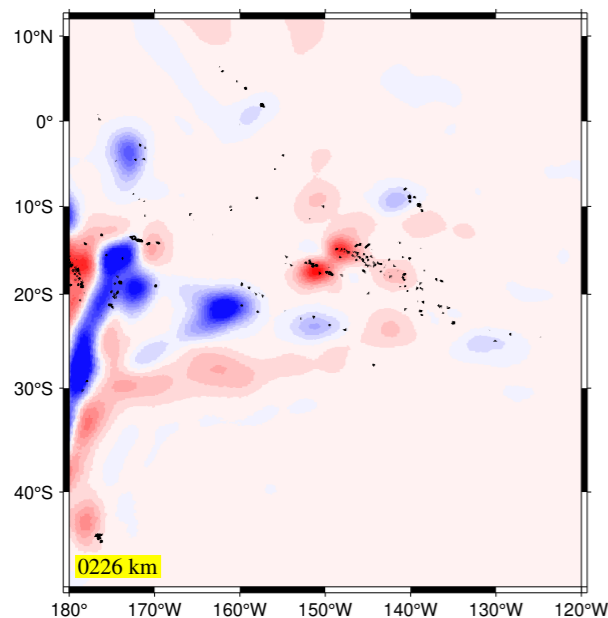
226 km

497 km

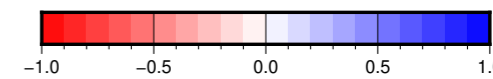
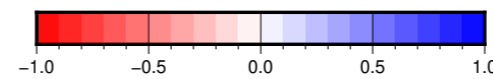
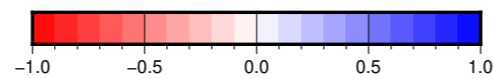
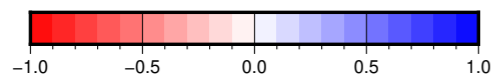
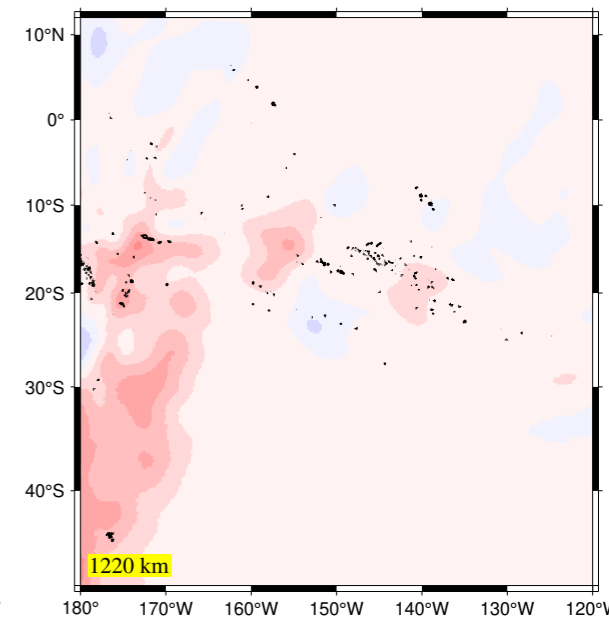
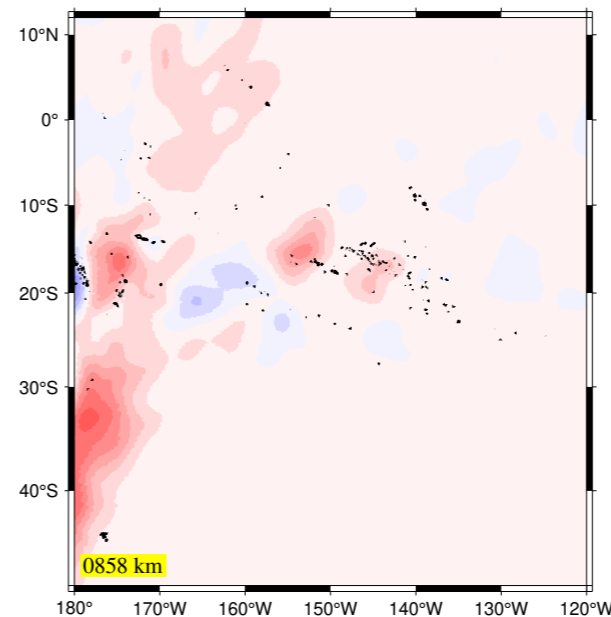
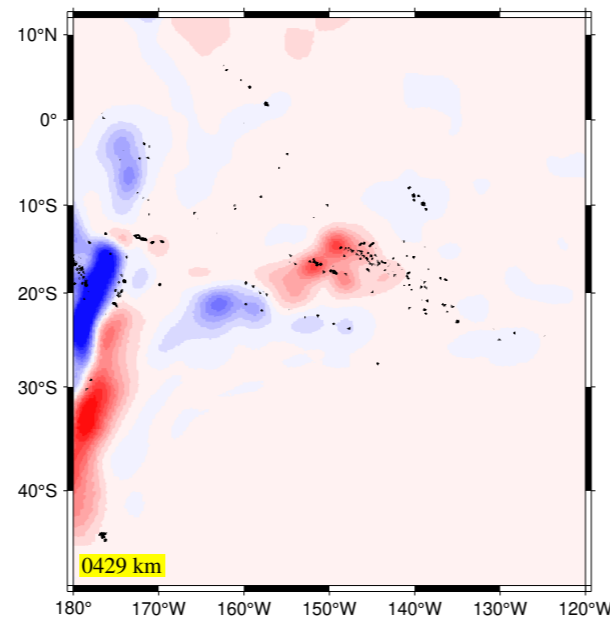
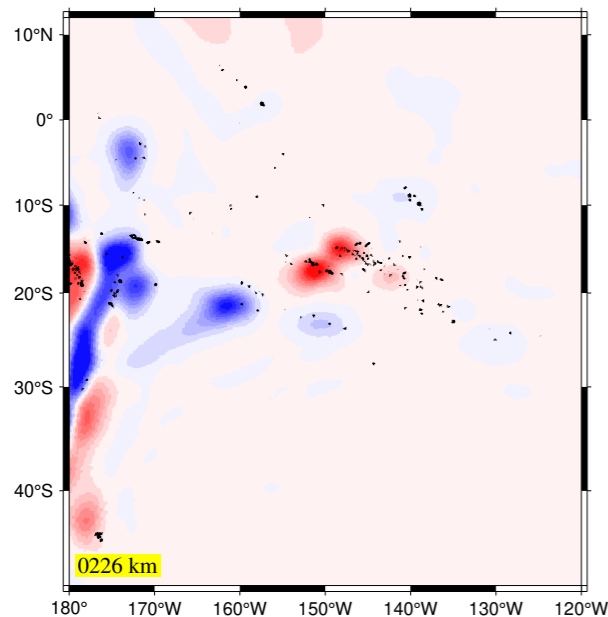
858 km

1220 km

A



B





## Enhancement of SPPIM data?

**LSQR penalty function is based on**

$$N\chi^2 = \sum_{i=1}^N \frac{[Am - d]_i^2}{\sigma_i^2}$$

**Maximum likelihood: equal weights**

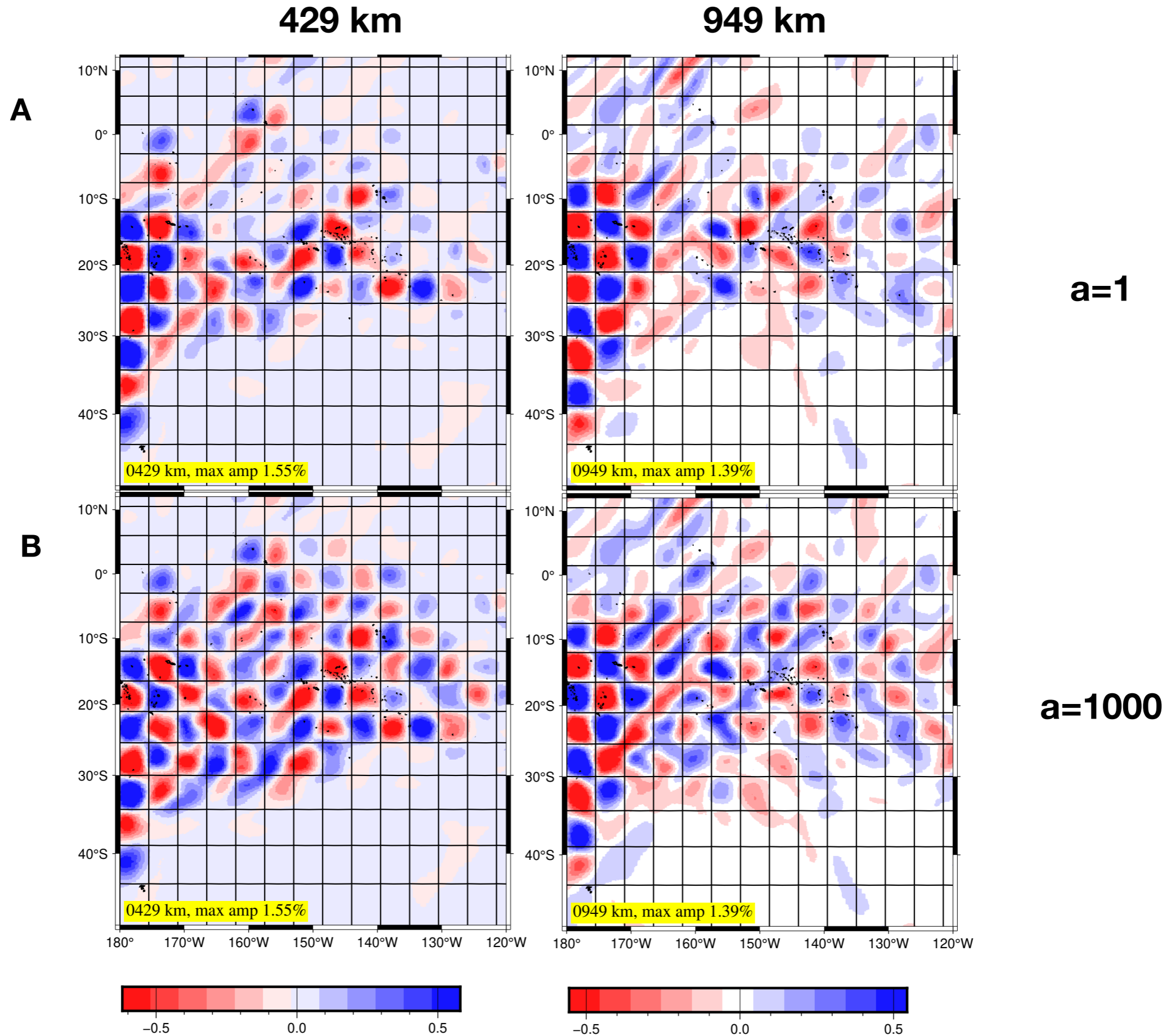
$$J = \chi_{MM}^2 + \chi_{rest}^2$$

**LSQR searches first in the direction of the densest columns (subduction zones)**

**Enhancement by factor a:**

$$J = a \cdot \chi_{MM}^2 + \chi_{rest}^2$$

# Resolution test without (A) and with (B) 1000x enhancement



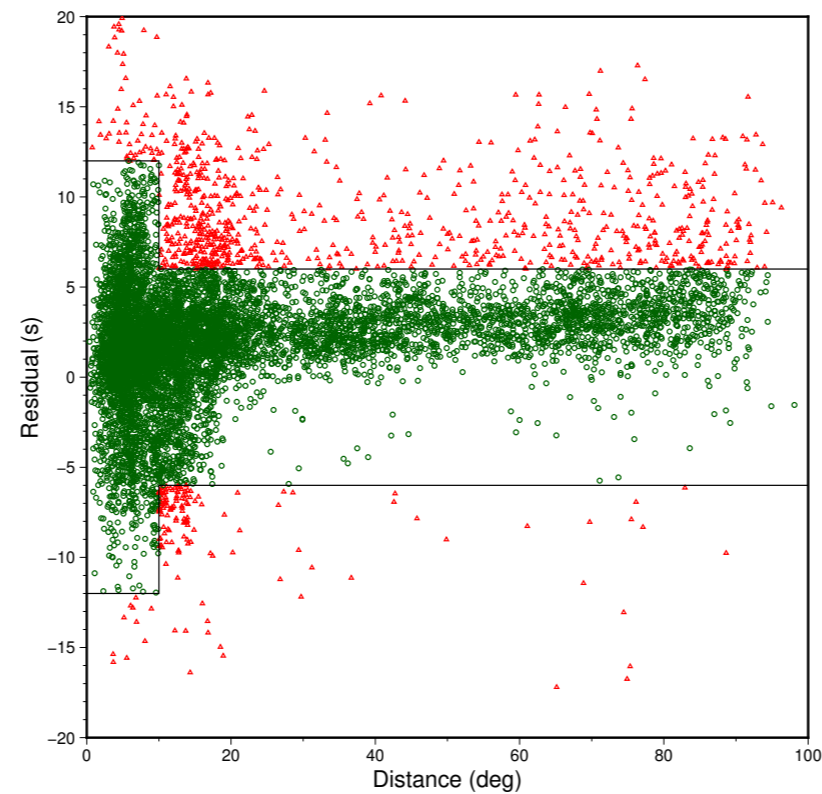




On synthetic data with synthetic noise enhancement seems doable

But is the synthetic noise really representative for the actual errors?

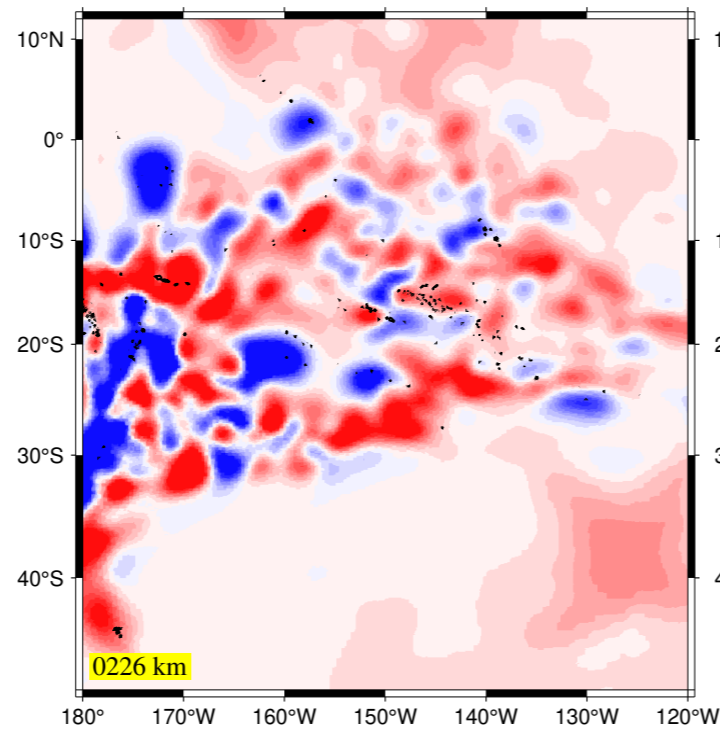
This plot worries me:



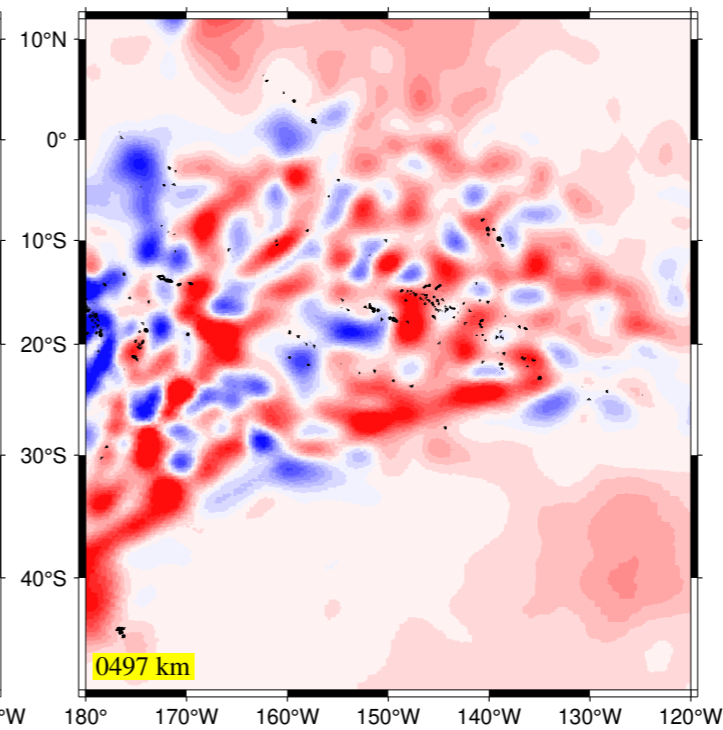


And, yes, enhancement leads to a chaotic image with real data:

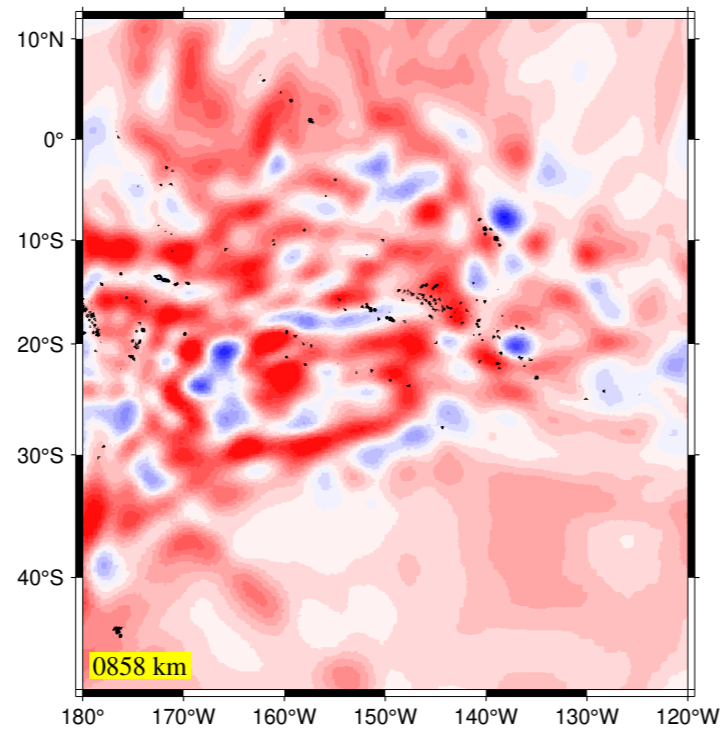
229 km



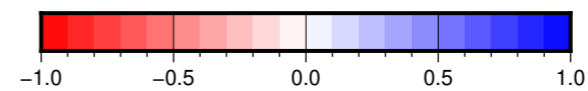
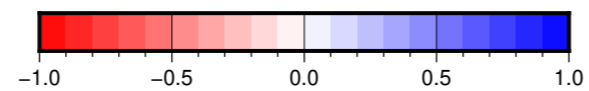
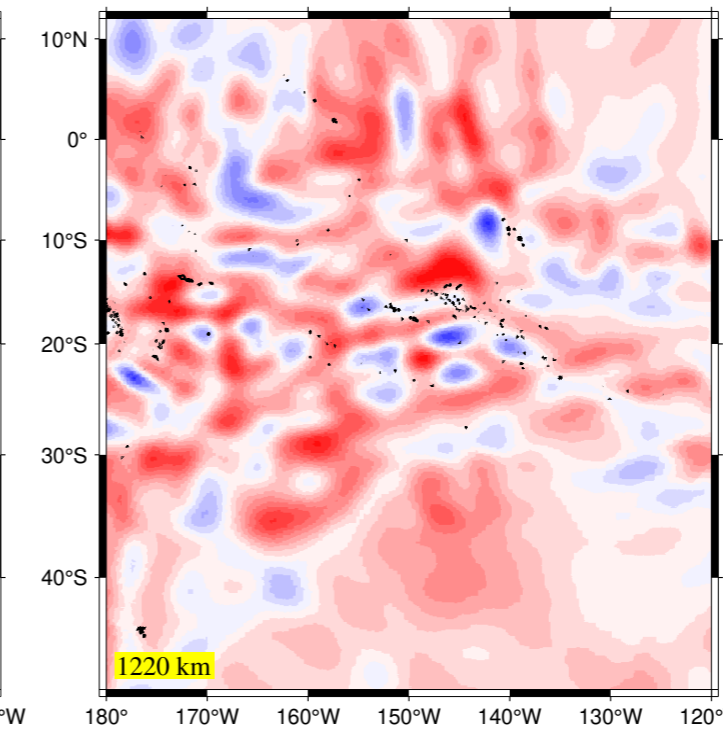
497 km



858 km

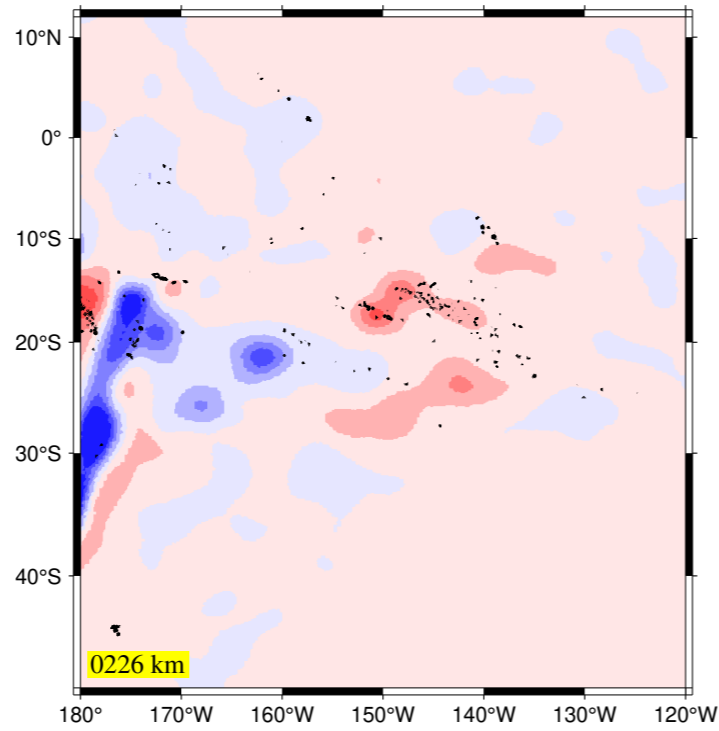


1220 km

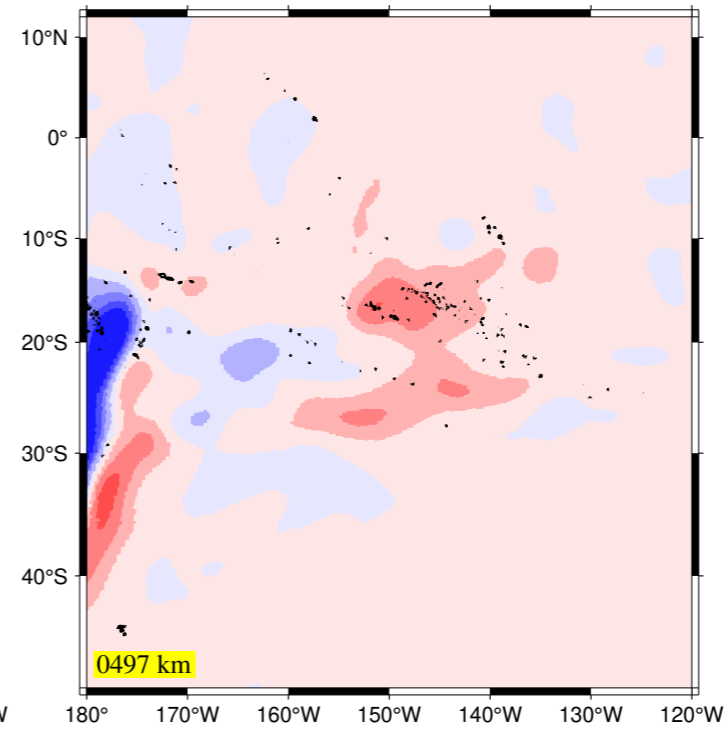


# Reducing $a=20$ and more smoothing helps:

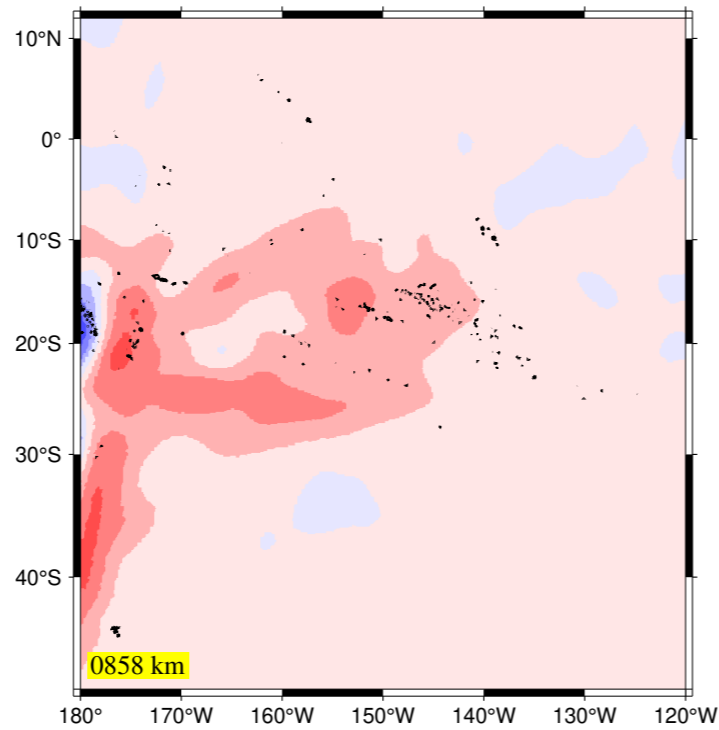
**229 km**



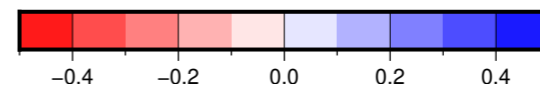
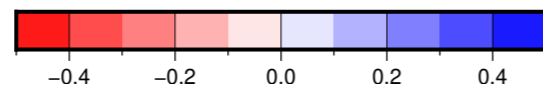
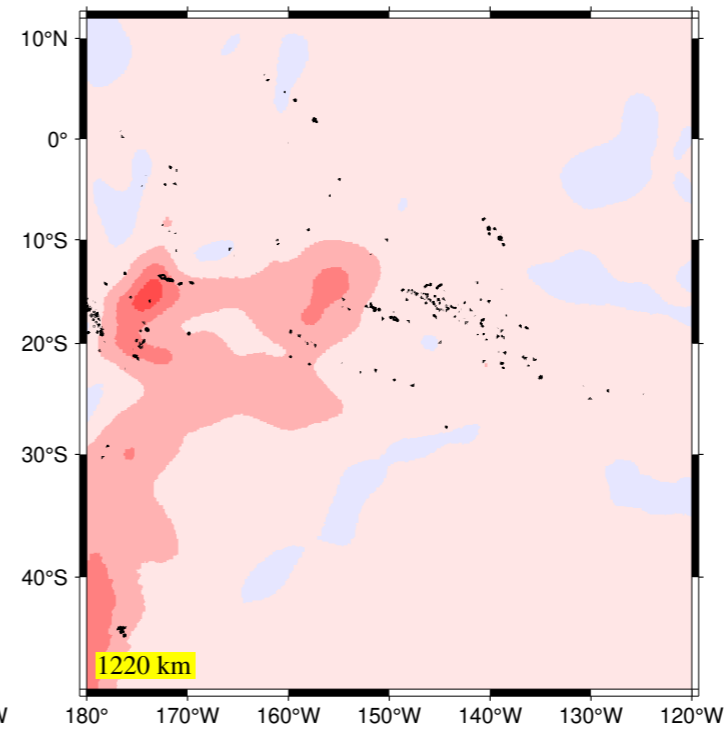
**497 km**



**858 km**



**1220 km**

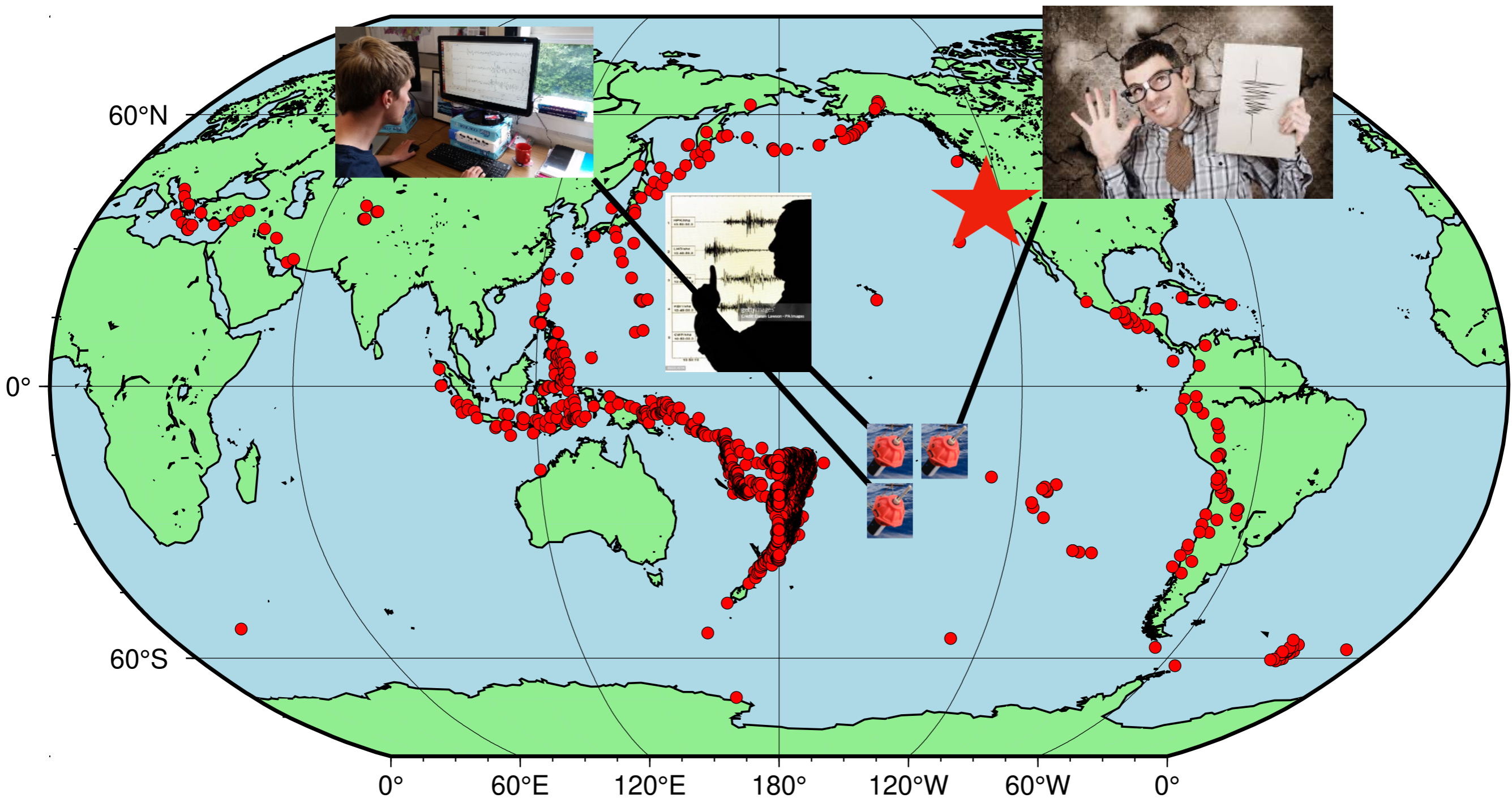




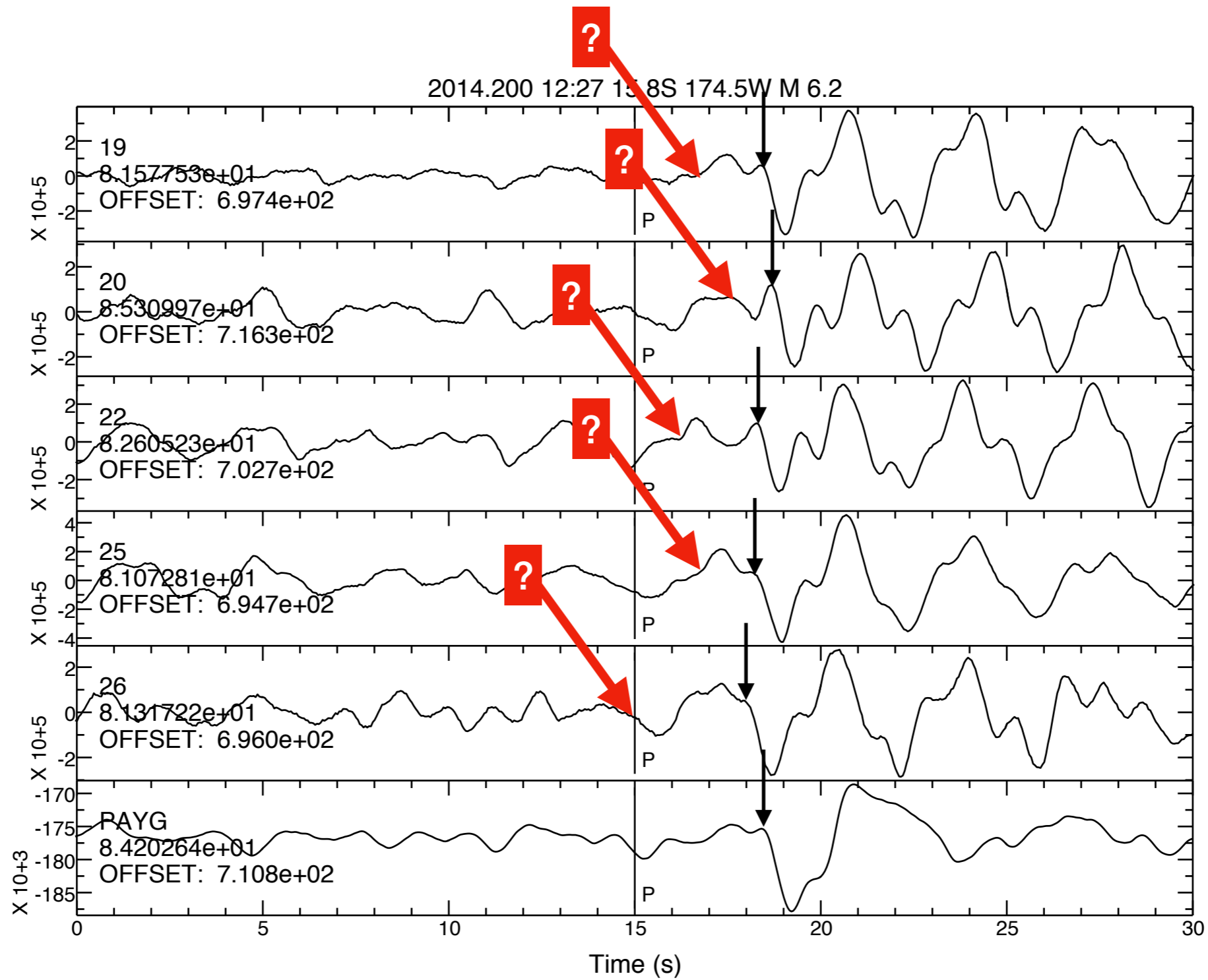
**Why do real data fail where (also noisy) synthetic data succeed ?**

**Answer: *there are too many outliers in the real data***

**My diagnosis: different records of the same event are analysed by different people with no ability to compare**



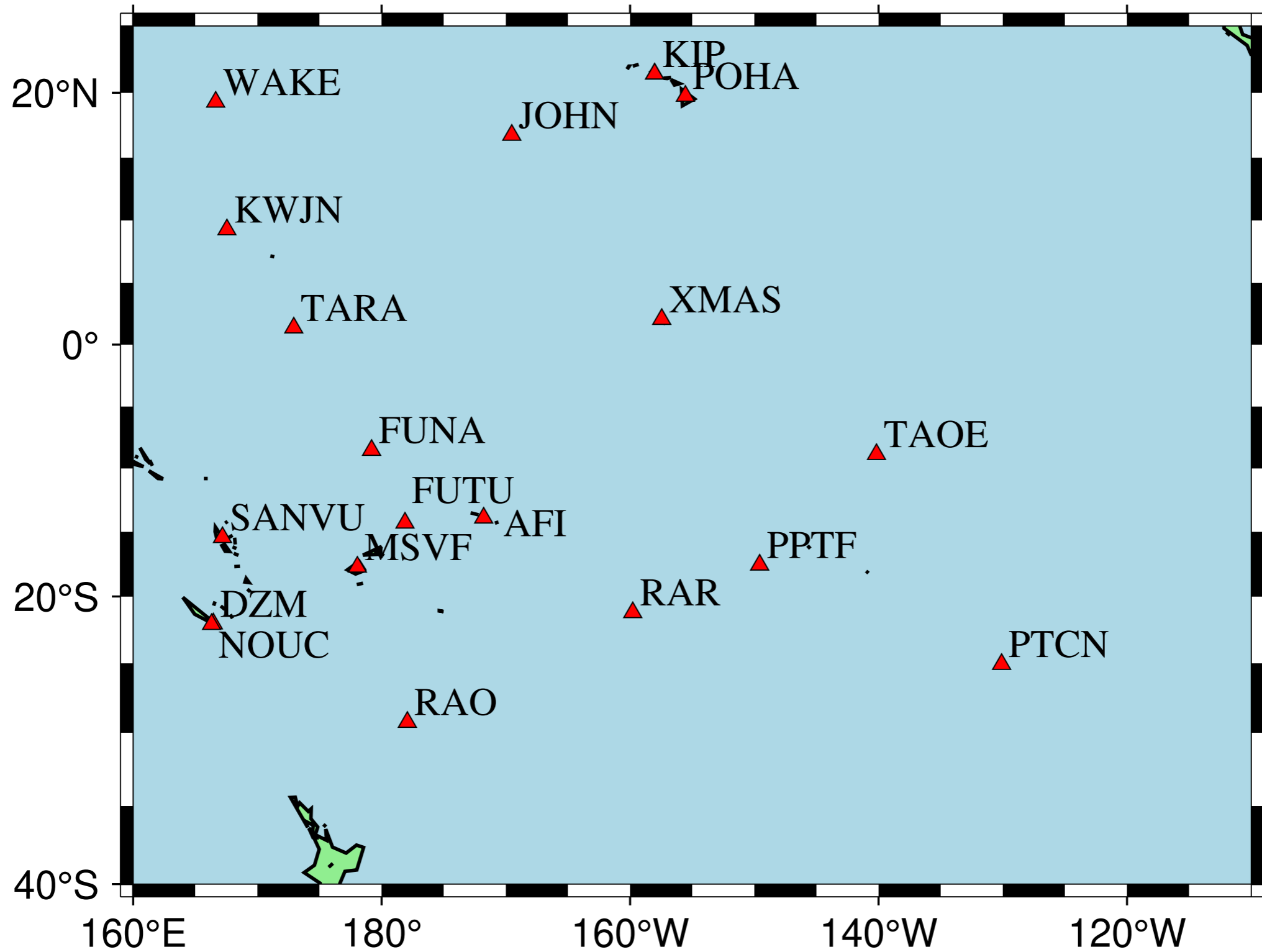
# This is what I did with Galapagos data



**PAYG is  
on (is)land**



**There are also plenty island stations to compare onsets**



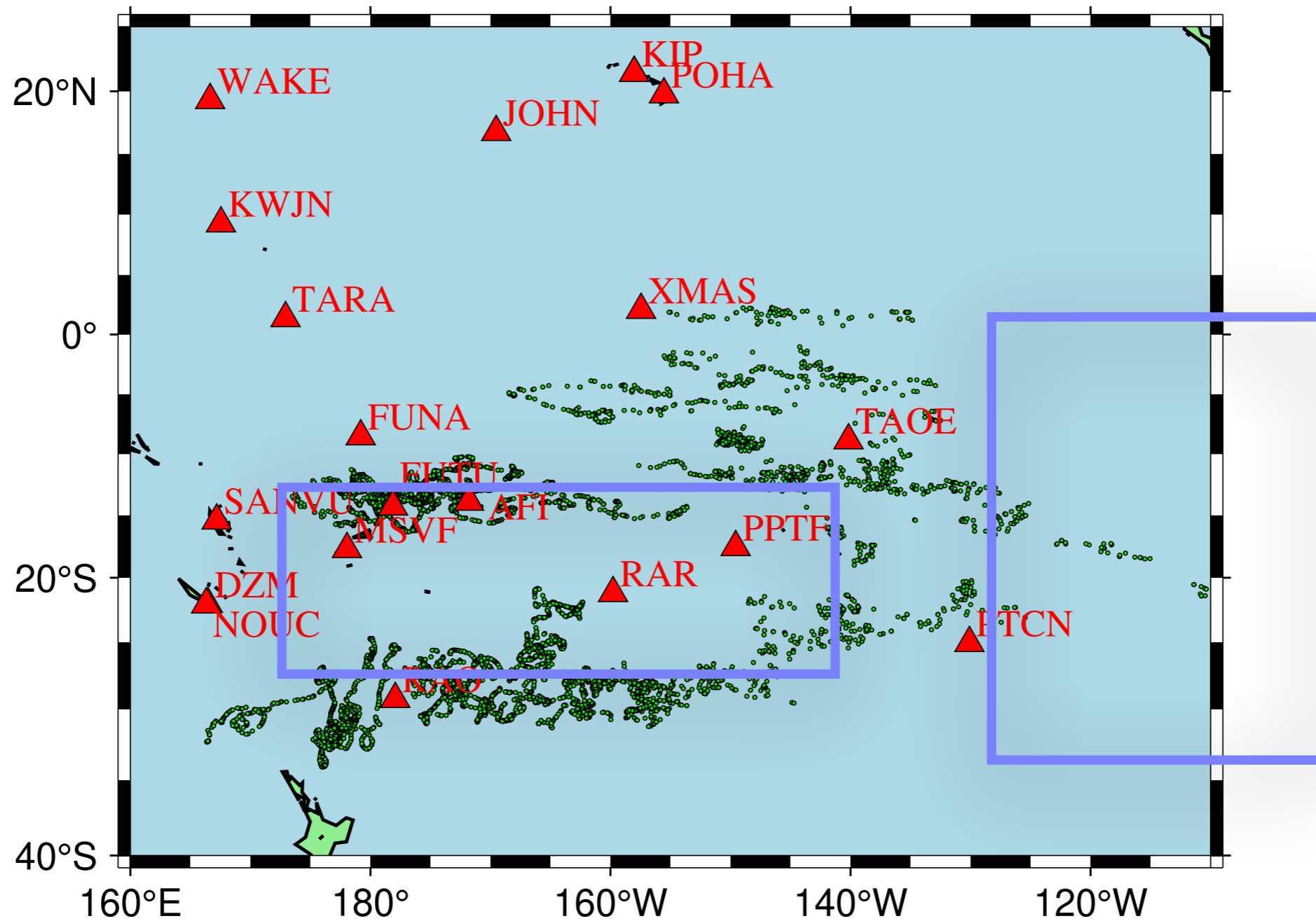
## Recommendation I

- **Group all data by event, if this has not yet been done**
- **Download island station data for all events and  
convert to velocity**
- **Divide the work *by event***
- **Identify the waveform shape and polarity of the onset**
- **Use AIC to zoom in on start of the onset**



# Recommendation II (Masayuki?)

Future missions should focus on the gaps



**Thank you and happy New Year**

