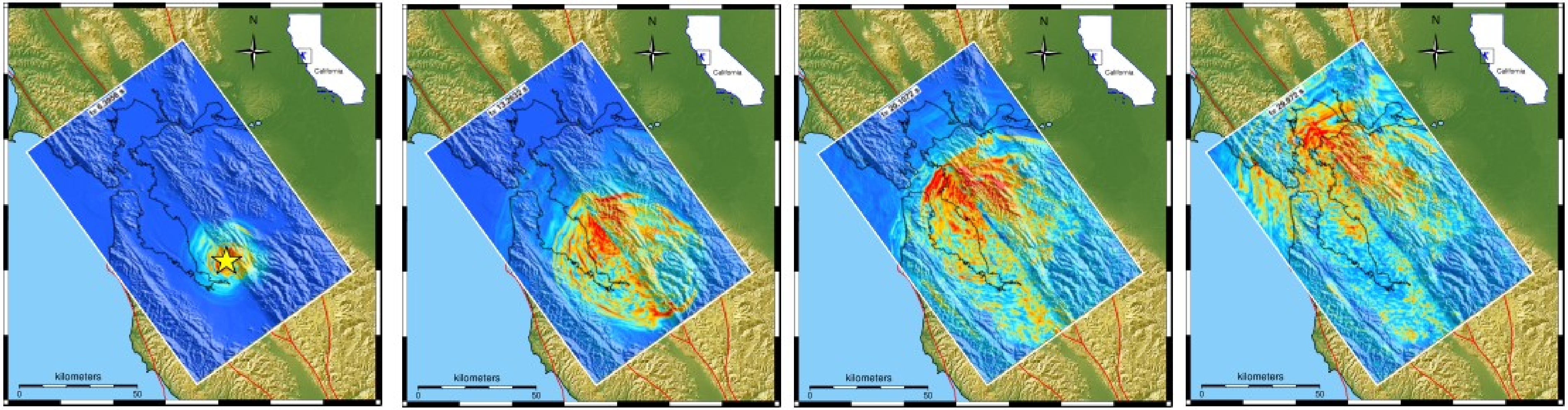


Performance evaluation of the velocity model developed by USGS for the San Francisco Bay Area

A PEER – LBNL workshop

January 18-19, 2024

Camilo Ignacio Pinilla Ramos
Critical Infrastructure Program
Energy GeoSciences Division
Lawrence Berkeley National Laboratory

BERKELEY LAB

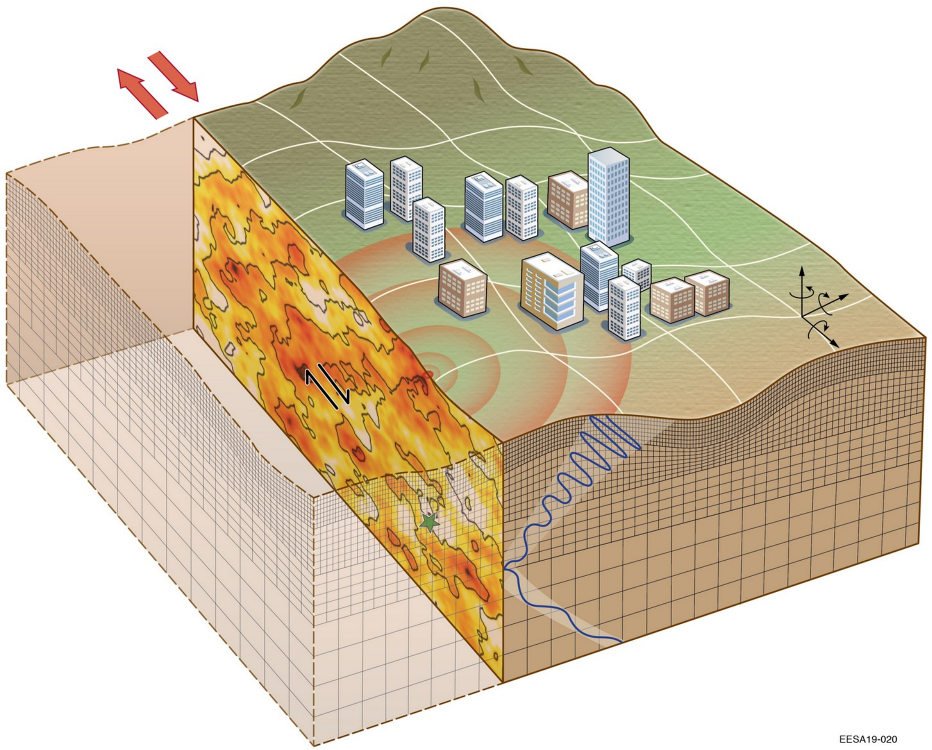
U.S. DEPARTMENT OF ENERGY

Office of Cybersecurity, Energy Security, and Emergency Response

EXASCALE COMPUTING PROJECT

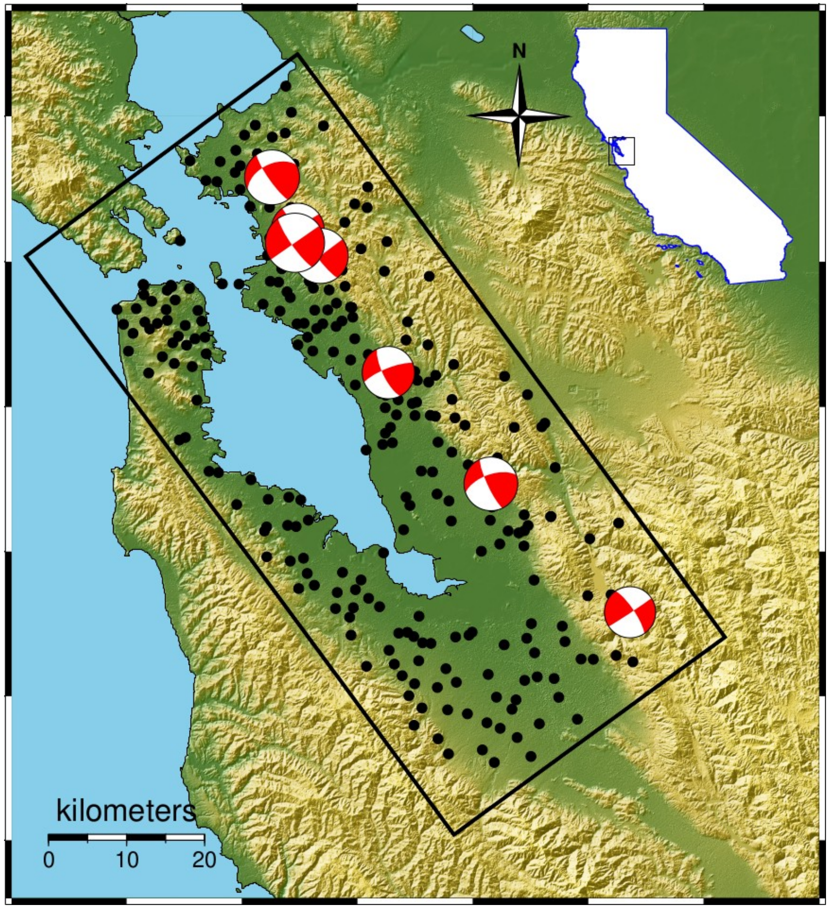
Introduction

- In this study, we focused in the **evaluation of the VM developed by the USGS** for the SFBA using an innovative methodology.
- The outline of this presentation is:
  - Setup of the simulations.
    - Source characterization and vel. model.
  - Results of our methodology:
    - Qualitative analysis.
    - Overall model performance.
    - Spatial distribution of spectral res.
    - Spatial distribution of duration res.
  - Discussion and conclusions.



Simulation setup

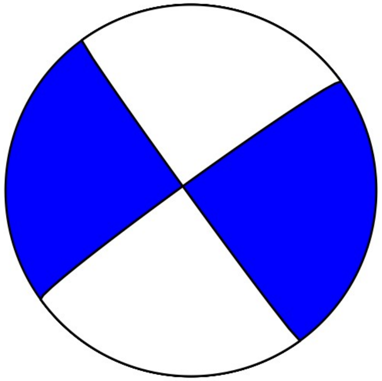
- 7 small Bay Area earthquakes **(M_W from 3.8 to 4.4)** simulated.
- **High-resolution simulations with SW4:** 5 Hz with a V_{Smin} of 250 m/s.
- The source: **double-couple point-source mechanism**.
- **Linear elastic constitutive model**, assuming negligible non-linear effects due to the small magnitudes of the earthquakes.



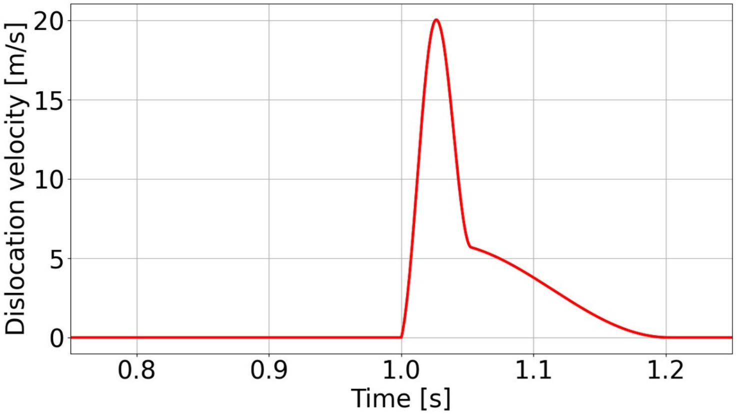
Map of stations and epicenters.

Source characterization

- The source was parameterized using external models and information, minimizing any **trade-off between the source characterization and the VM evaluation**.
- We adopted the focal mechanisms from the NCEDC.
- We modeled the spectral illumination of the source using **Liu's source time functions** (Liu et al. (2016)).
- **Corner frequency** from the regression and/or model developed by Trugman and Shearer (2018), who used earthquakes from the SFBA.

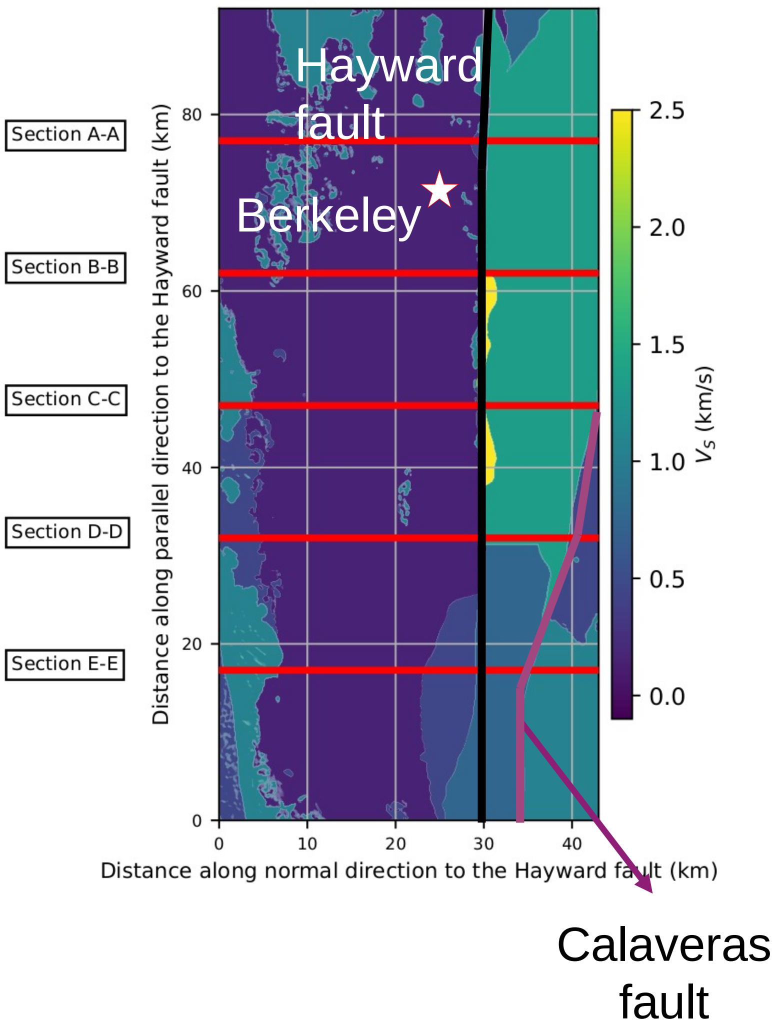


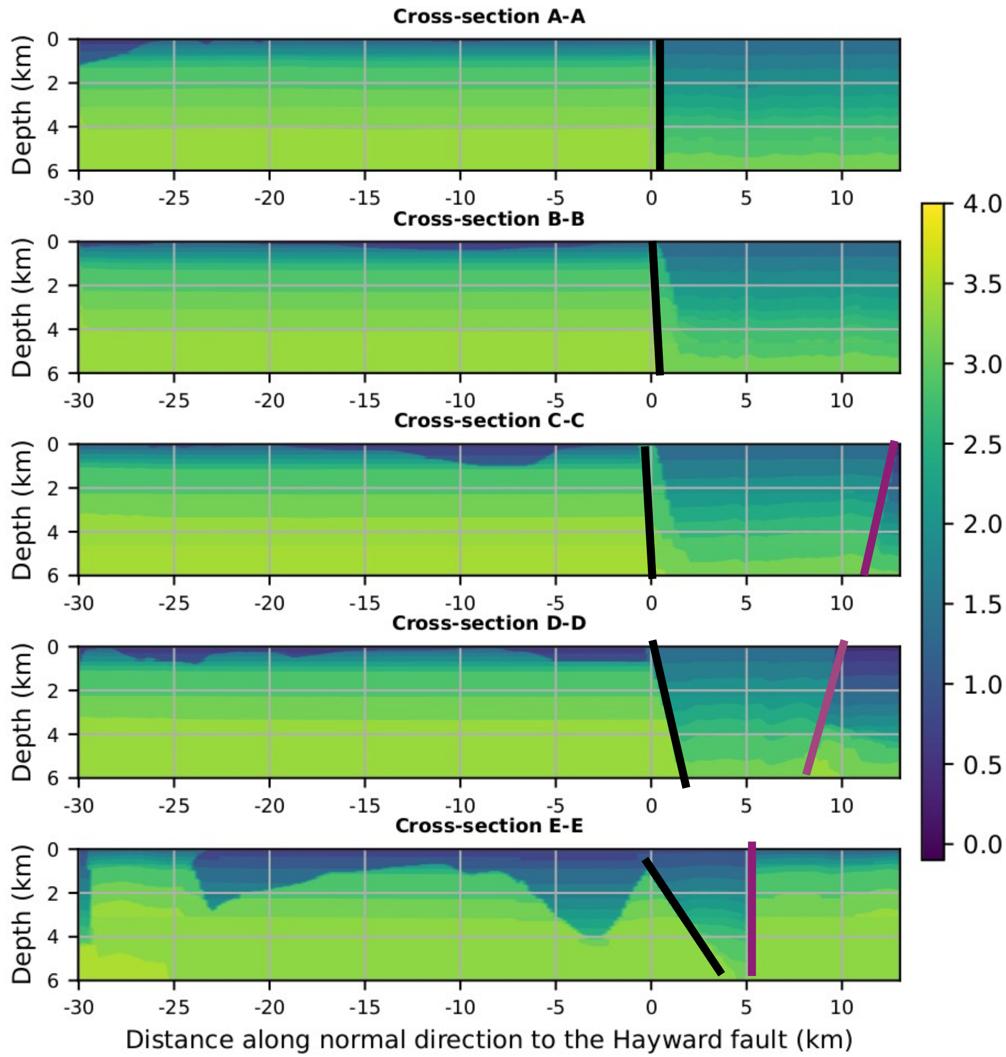
Classical shape of the strike-slip focal mechanism used in the simulations



Example of a Liu Source Time Function.

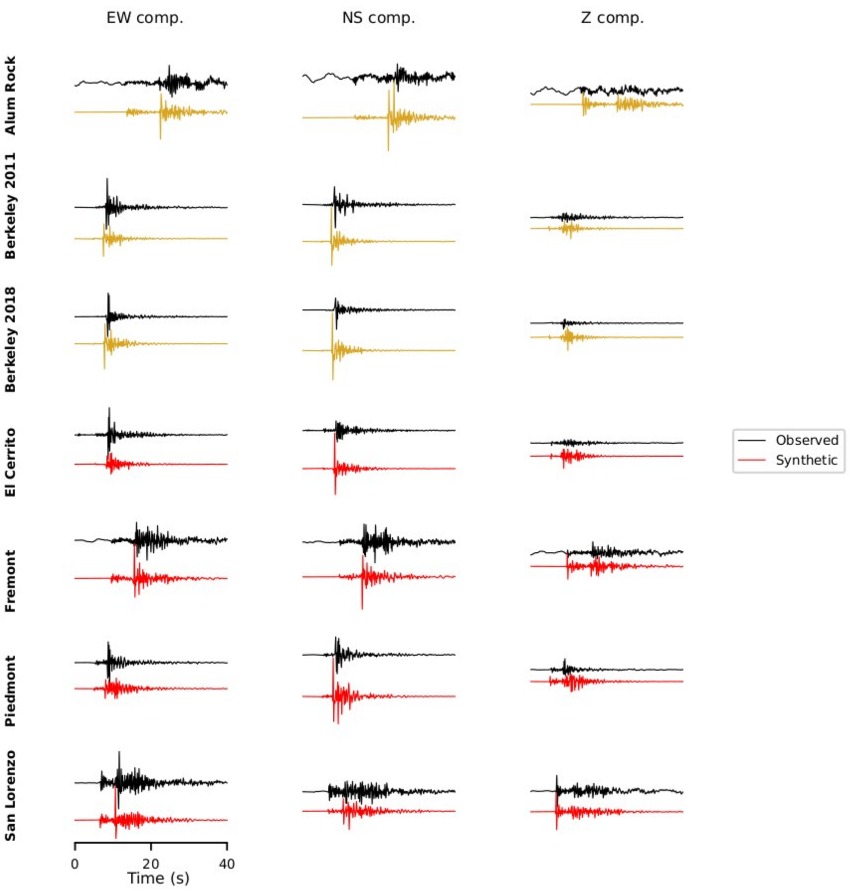
Velocity model cross-sections





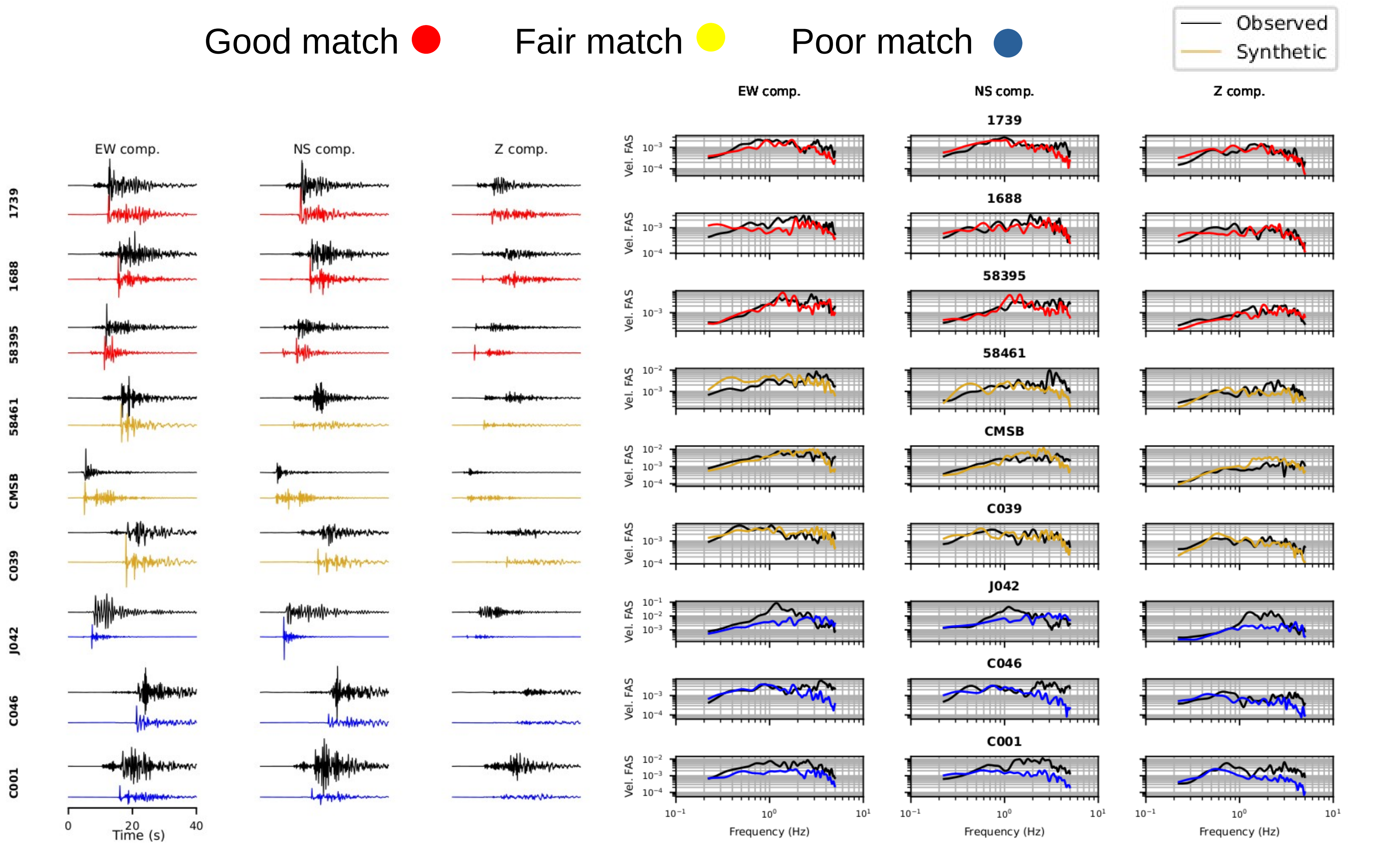
Qualitative analysis – Definition

- This analysis characterizes waveforms' fit in terms of amplitude and wave phases.
- They are based on **three criteria**:
  - Similarity of wave phases.
  - Similarity of wave amplitude.
  - Phase and amplitude fit considering all three components.
- We proposed **three classifications**:
  - Good waveform fit: Overall, the records satisfy all three criteria.
  - Fair fit: Most two of the three criteria are satisfied.
  - Poor fit: Instances in which two or more criteria are not satisfied



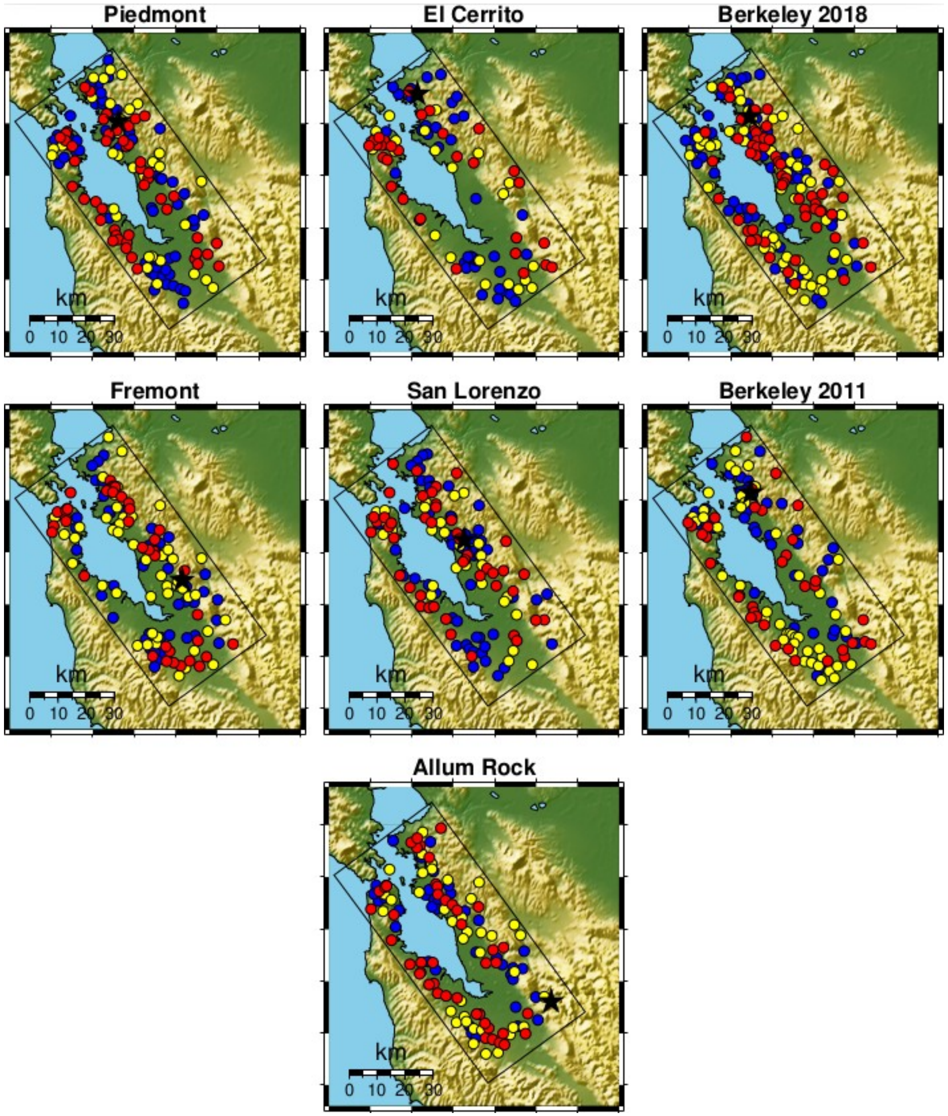
Waveforms at station J056 for the seven earthquakes. Red: good match. Yellow: Fair match

Qualitative analysis – Example of 9 stations for the 2018 M_W 4.4 Berkeley earthquake



Qualitative analysis – Spatial distribution of the classifications

- Spatial distribution of classifications by event.
- We can identify path effects: Some areas show good matches for some events and fair and poor matches for others.
- For example San Jose / Santa Clara region (south of our domain).



Good match

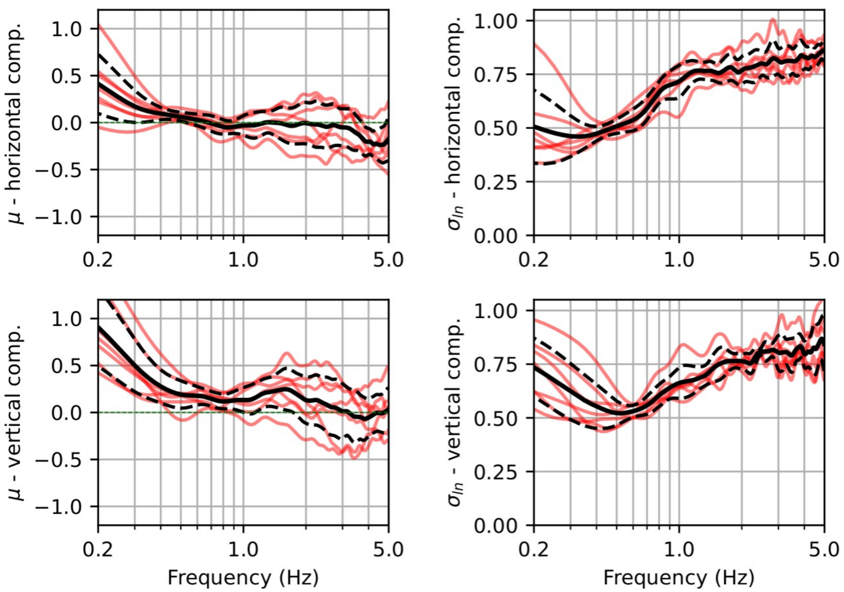
Fair match

Poor match

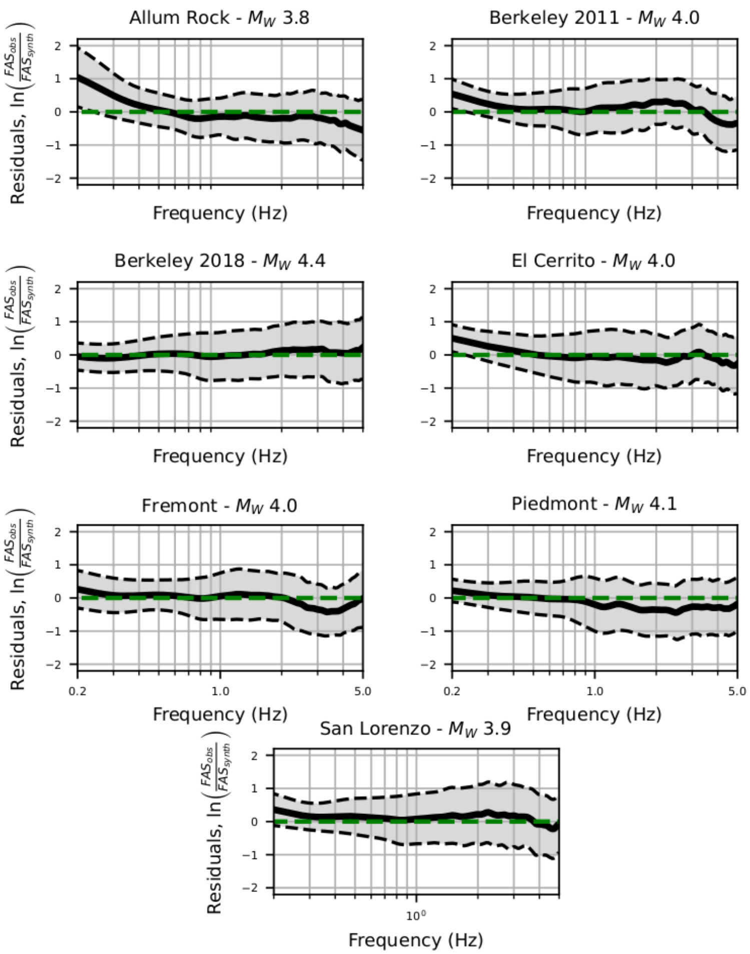
Overall model performance

- Residuals in the Fourier domain aid in evaluating the overall model performance.
- The mean FAS residuals for the seven events are centered around 0 up to 3.5 Hz, with a slight over-prediction of 0.2 LN units.

Global mean and standard deviation

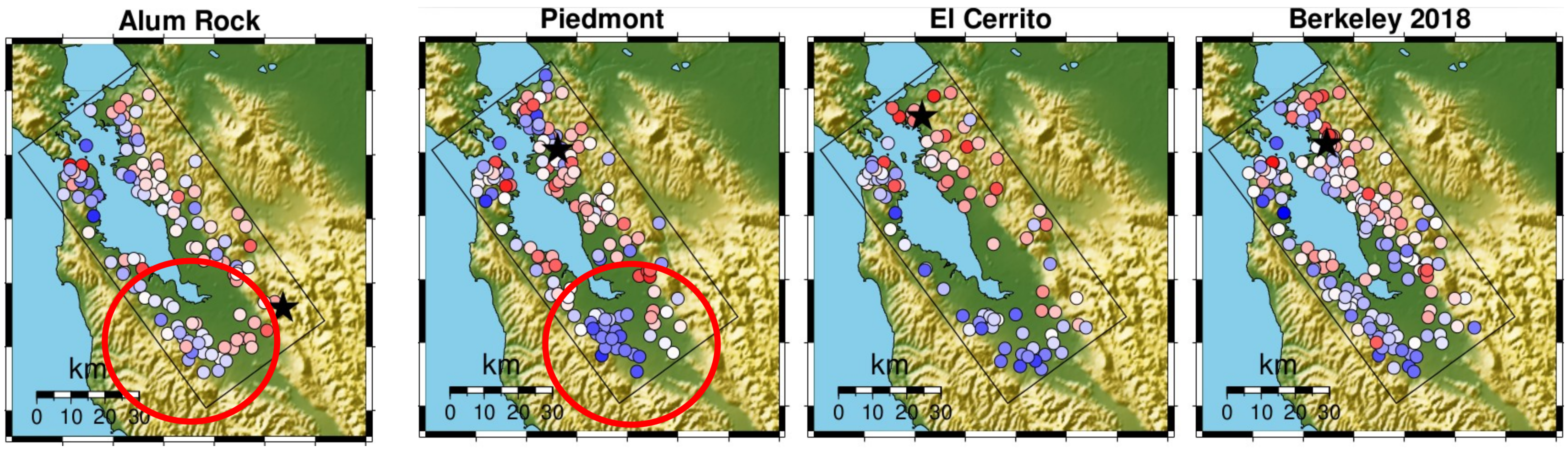


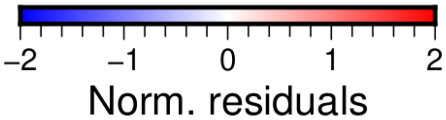
Residuals' comparison per each event

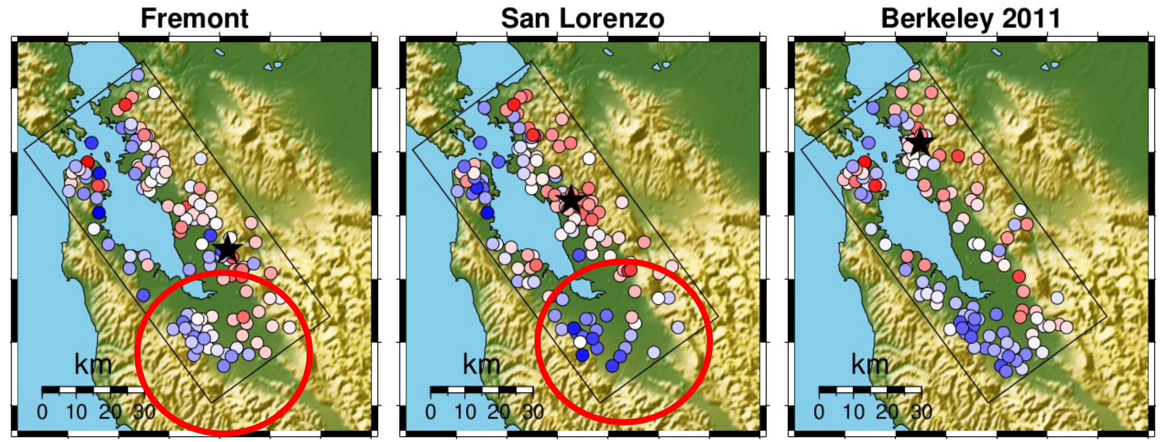


Residuals' spatial distribution – Sa at 1 Hz

- Spatial distribution of residuals allows identification of path effects, especially from basin-induced reverberation and surface waves.
- The southernmost region of our domain manifests unconstrained path effects.







- Red: Under-prediction
- Blue: Over-prediction

Modeling spatially-varying patterns of systematic sub-optimal predictions

- Residuals' decomposition:

$$Y_{es} - \hat{Y}_{es} = \delta B_e + \delta S(\vec{X}) + \delta E_{es}$$

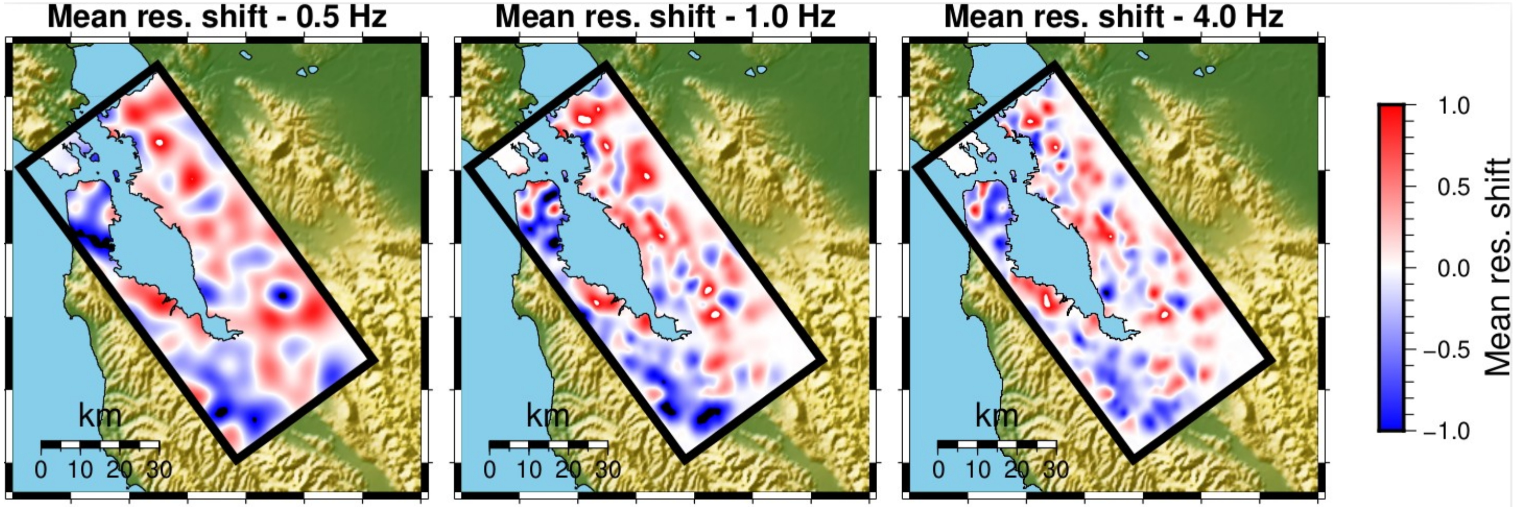
- We solved this equation using Bayesian modeling.
- δS(X) was modeled through a Gaussian process regression.

Hyper-parameters

Spectral frequency (Hz)	ρ (km)	τ (LN units)	σ_0 (LN units)
0.25	4.361	0.095	0.431
0.5	3.701	0.118	0.408
0.75	1.941	0.174	0.383
1	1.687	0.207	0.408
2	1.268	0.27	0.504
4	0.762	0.329	0.625

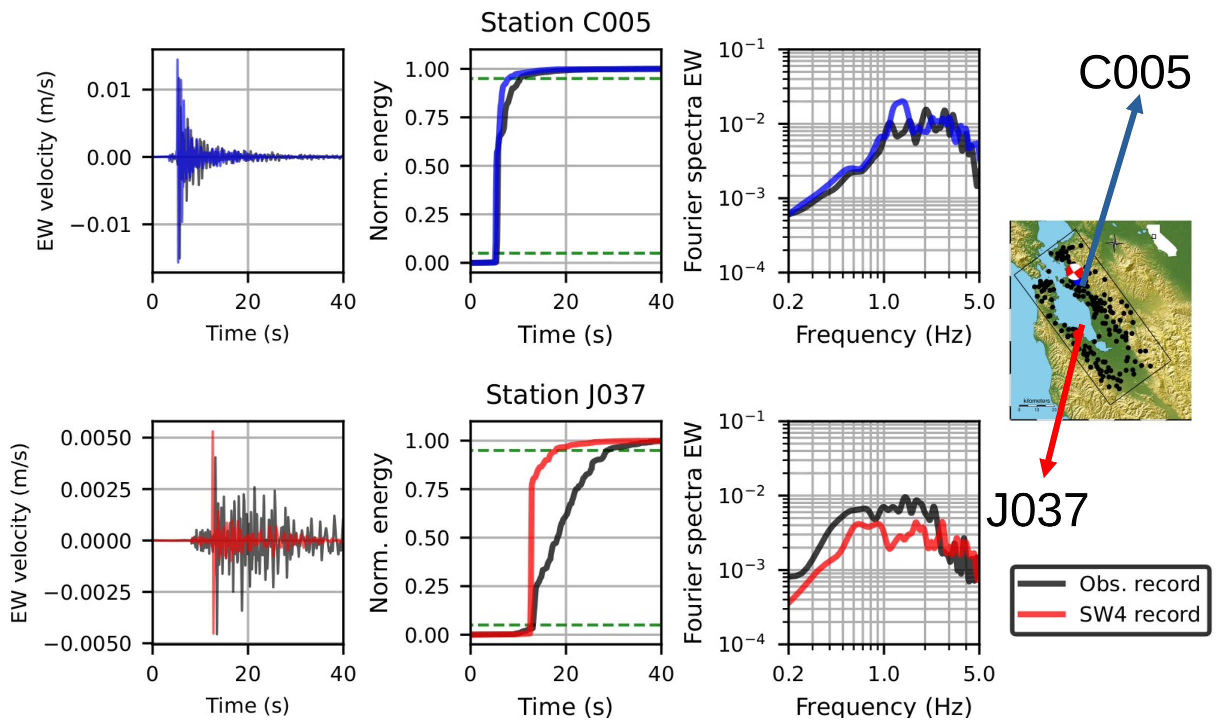
τ: Standard deviation of δB_e.
ρ: Correlation length of δS(X).
σ_0: Remaining unexplained variability.

Median δS(X)



Analysis of the waveforms duration – Example cases

- D_{5-95} captures most of the wave scattering and surface waves manifestations on the wave field (Pinilla-Ramos at al., 2023).
- This parameter illustrates where the velocity structure may be too simple compared to the actual geologic structure.
- Station C005 shows a similar duration between observed and synthetic (2.8 vs 4.9 s). The Husid plot (center plot) shows a similar distribution of the energy in time for both records.
- In contrast, station J037 shows durations of 15.6s and 5.4s for the obs. and synth. signal, respectively.

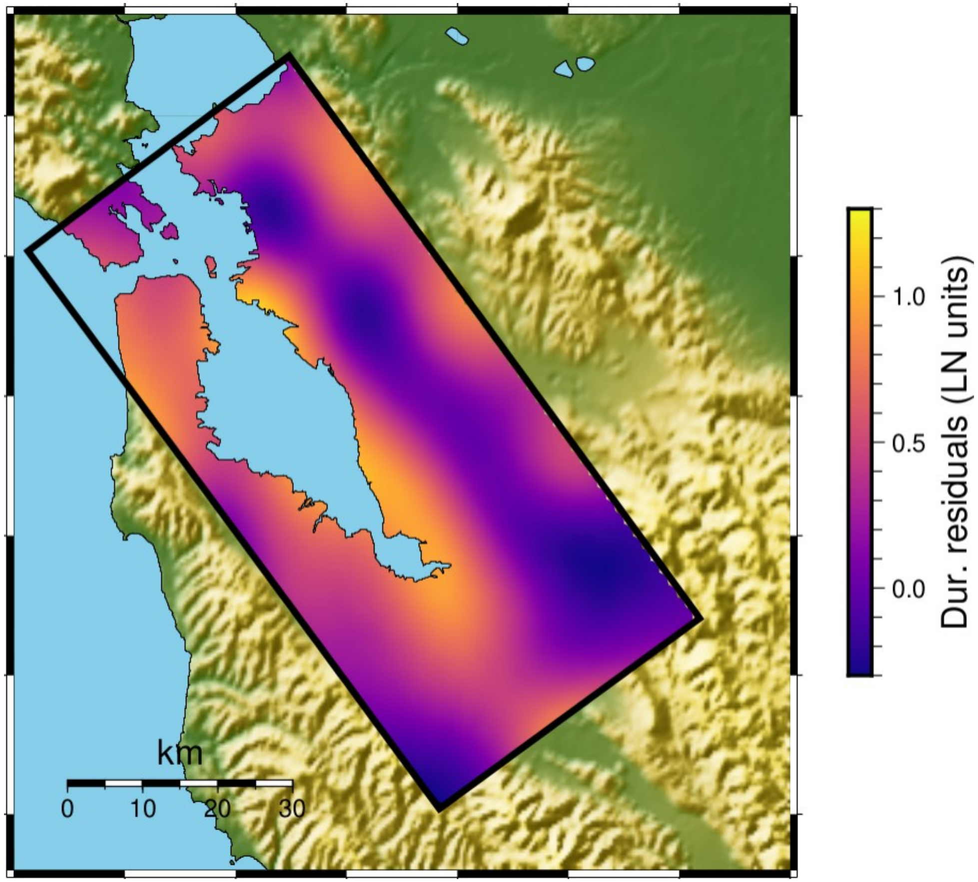


Analysis of waveforms duration – Modeling the residuals' spatial distribution

- Residuals formulated in LN scale, as:

$$e_{es} = \ln\left(\frac{D_{5-95-obs}}{D_{5-95-synth}}\right)_{es}$$

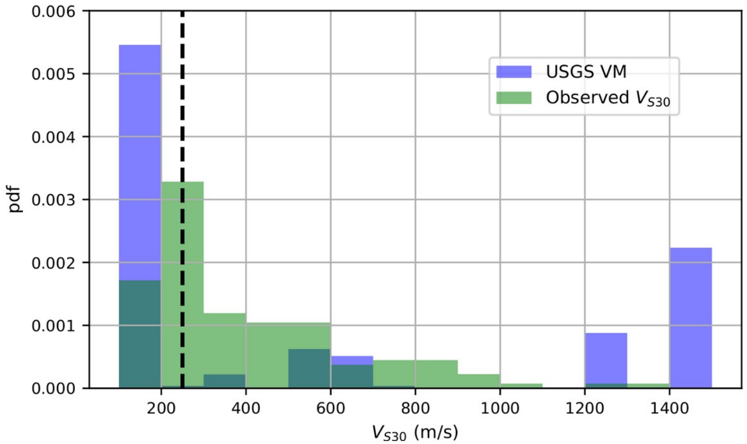
- Modeled using a Gaussian process regression.
- The hills tend to have a ratio of 0 in LN units.
- The sediments of the bay tend to have systematic longer duration.



Median of our model, showing areas with systematic longer or shorter duration.

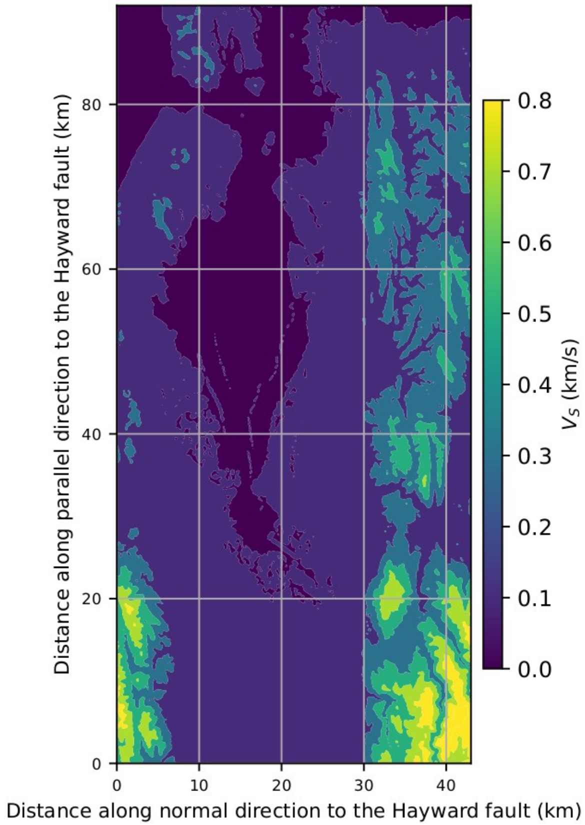
Discussion: Geotechnical layer representation

- V_{S30} is used as a proxy to understand how the geotechnical layer is constrained in the velocity model.
- The velocity model has an over-representation of stiff sites, concentrated mostly on outcrops.
- The sediments of the SFB basin are represented with softer conditions than their actual ones.
- Our $V_{S\text{-}MIN}$ threshold aids in reconciling some differences.



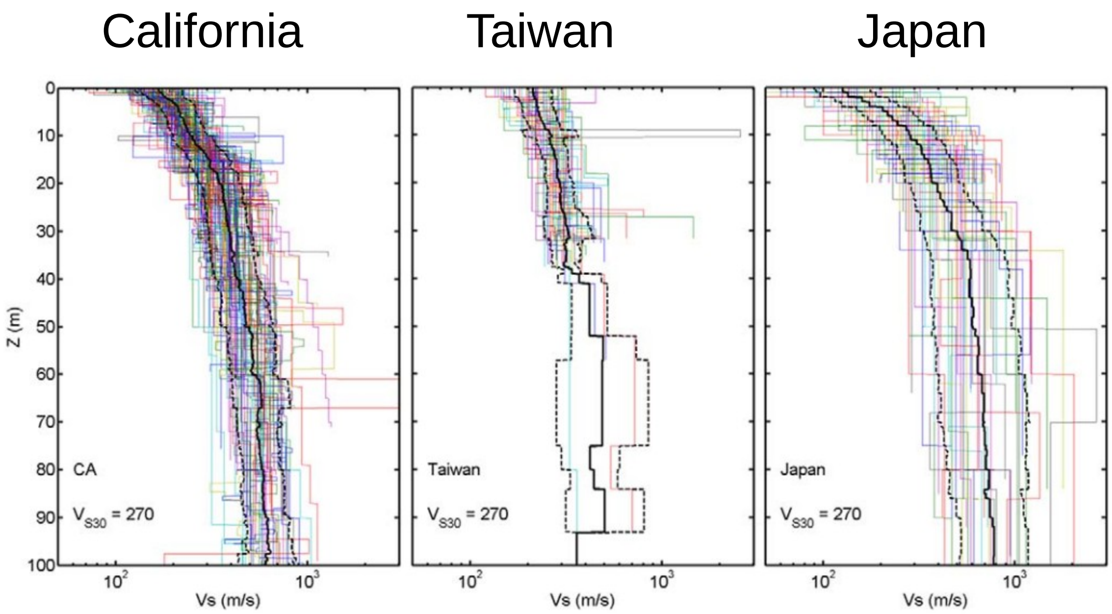
Comparison of the V_{S30} distribution between the USGS VM and an actual dataset (Tehrani et al. 2022).

Map view of the free surface V_S.



Discussion: Geotechnical layer representation

- California has the advantage that the V_S profile in depth is strongly correlated with V_{S30} (Boore et al. 2011, Kamai et al. 2016), making the V_{S30} correlated with site amplification (Pinilla-Ramos et al., 2022). This provides an opportunity to developing rules for updating the velocity model in depth, conditioned by actual velocity values close to the surface.



Velocity profile model in depth for three regions (Kamai et al. 2016).

Discussion: Benefits of this methodology

- The **epistemic uncertainty** between synth. and obs. waveforms comes mainly from:

Epistemic uncertainty decomposition

$$\sigma^2 = \sigma^2_{VM} + \sigma^2_{SC} + \sigma^2_{NL}$$

Total variance

Velocity model

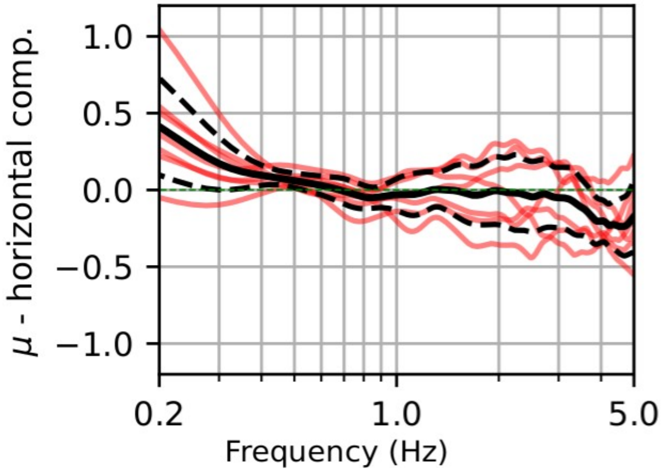
Source characterization

Constitutive models

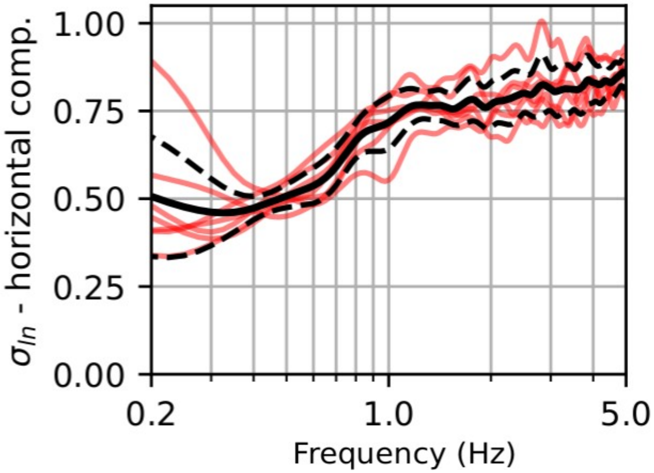
- σ_{VM}, σ_{SC}, and σ_{NL} represents **mismodeled effects.**
- The between-event variability, τ , was significantly smaller than the variability induced by the velocity model: **most of the differences between obs. and synt. waveforms can be attributable to wave propagation effects**.
- This methodology provides a **direct and quantitative estimation of the epistemic uncertainty** induced in wave fields propagated in this velocity model, which can be then **propagated in any seismic hazard analysis applications**.

Conclusions

- The USGS velocity model for the SFBA shows an overall good performance, with residuals zero-centered up to 3.5 Hz, and then at 5 Hz a slight systematic over-prediction of 0.2 in LN units (25%).
- Slight over-prediction at high-frequencies may be improved by including stochastic variability in the velocity model (Savran and Olsen, (2019), Graves and Pitarka, (2016), Abrahamson et al. (2022)).
- The within-event standard deviation ranges from 0.48 LN-units to 0.8 LN-units in the Fourier domain.



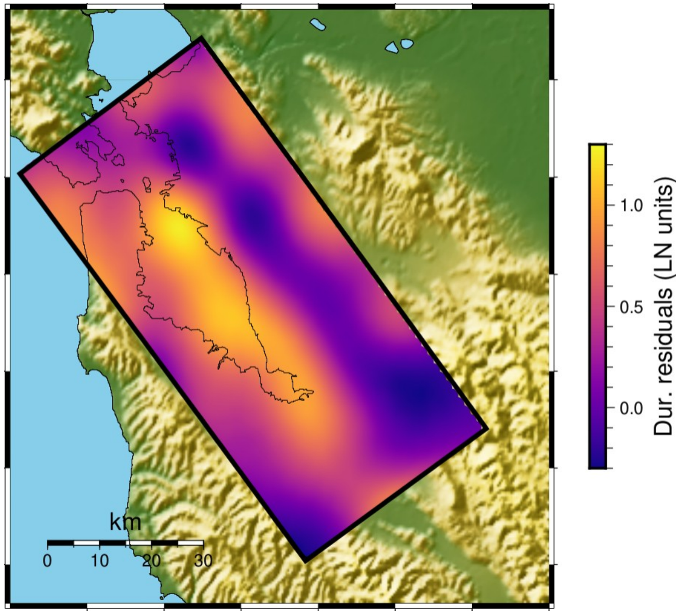
Horizontal mean bias in the Fourier domain.

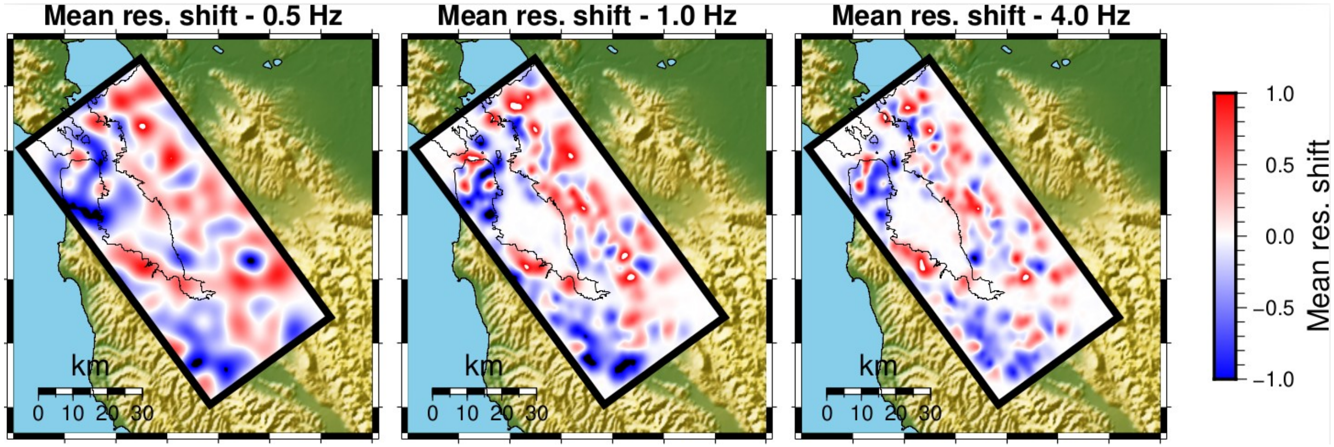


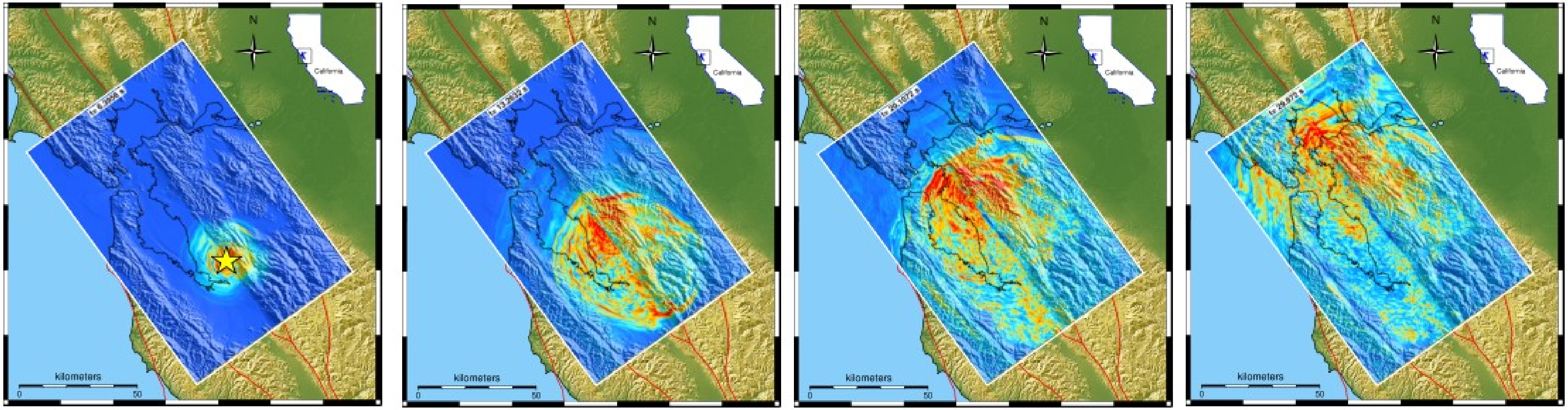
Mean within-event standard deviation.

Conclusions

- Our analysis and maps shows areas with systematic over- or under-prediction, providing an opportunity to improve the velocity model.
- The duration residuals conceptually show areas where the VM has the proper geometry of their buried geologic structures and where refinements would be useful.







Performance evaluation of the USGS velocity model for the San Francisco Bay Area

A PEER – LBNL workshop

January 18-19, 2024

Camilo Ignacio Pinilla Ramos
Critical Infrastructure Program
Energy GeoSciences Division
Lawrence Berkeley National Laboratory

BERKELEY LAB

U.S. DEPARTMENT OF ENERGY

Office of
Cybersecurity, Energy Security, and Emergency Response

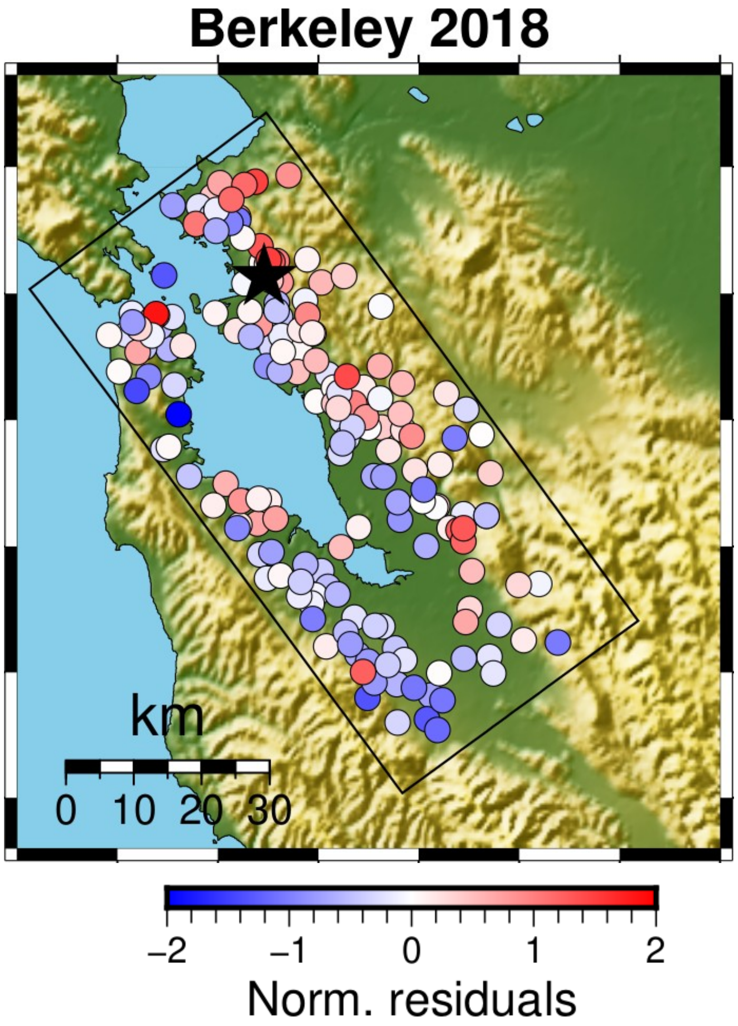
ECP
EXASCALE COMPUTING PROJECT

Residuals' spatial distribution – Sa at 1 Hz

- Spectral acceleration residuals at 1 Hz for the Berkeley 2018 earthquake:

$$e_{es} = \ln\left(\frac{y_{obs-es}}{y_{synth-es}}\right)$$

- We can identify spatial correlation patterns, with spots of over- or under-prediction.



- Red: Under-prediction
- Blue: Over-prediction