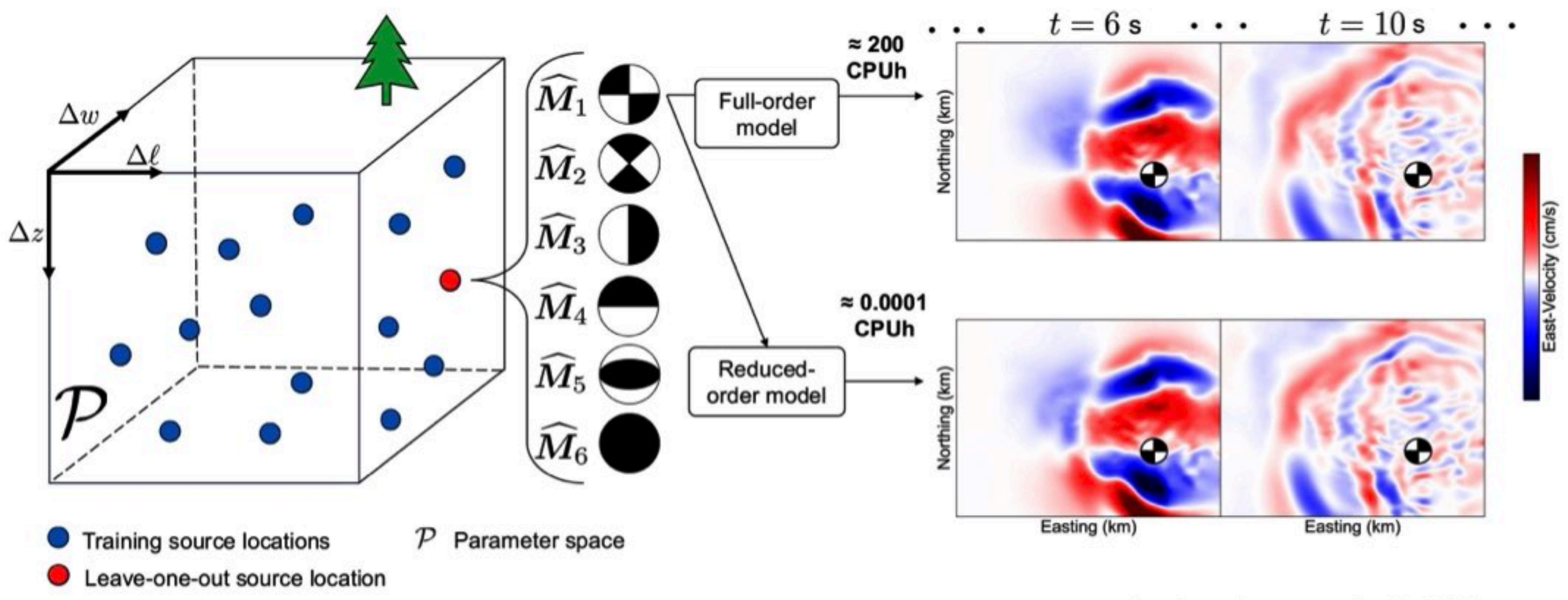
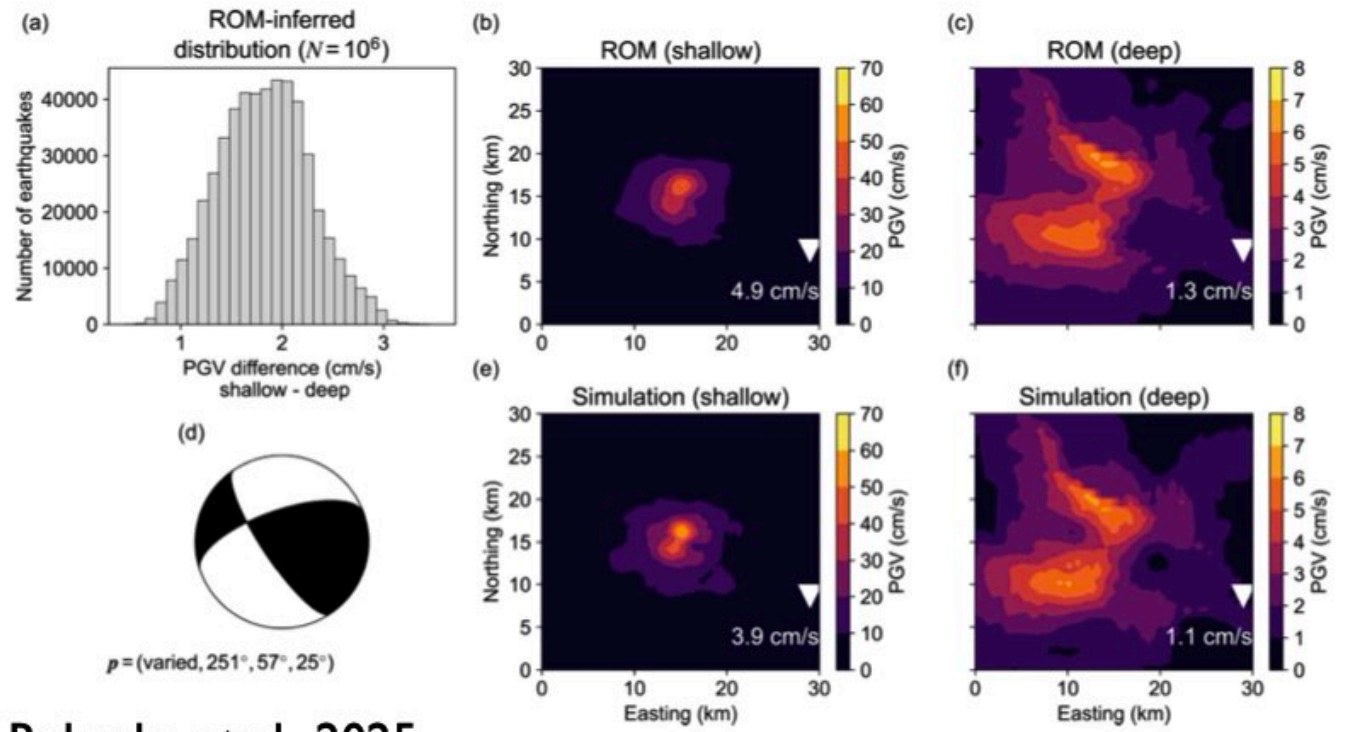
Near-Instantaneous physics-based reduced-order modeling for shake maps, seismograms, dynamic rupture and the seismic cycle

Alice-Agnes Gabriel, Dave May, John Rekoske, Yohai Magen, Gabrielle Hobson

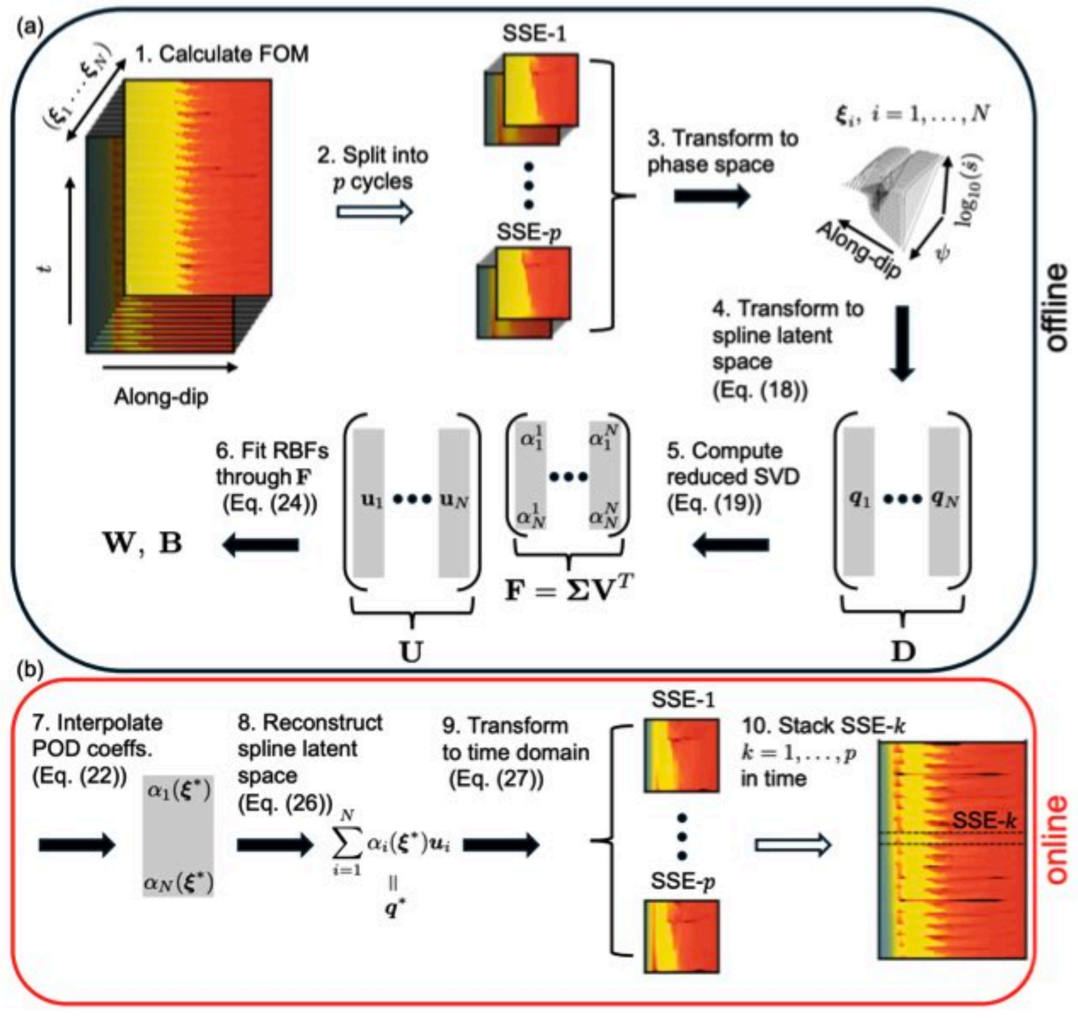
SCRIPPS INSTITUTION OF OCEANOGRAPHY
UC San Diego



Rekoske et al., 2023

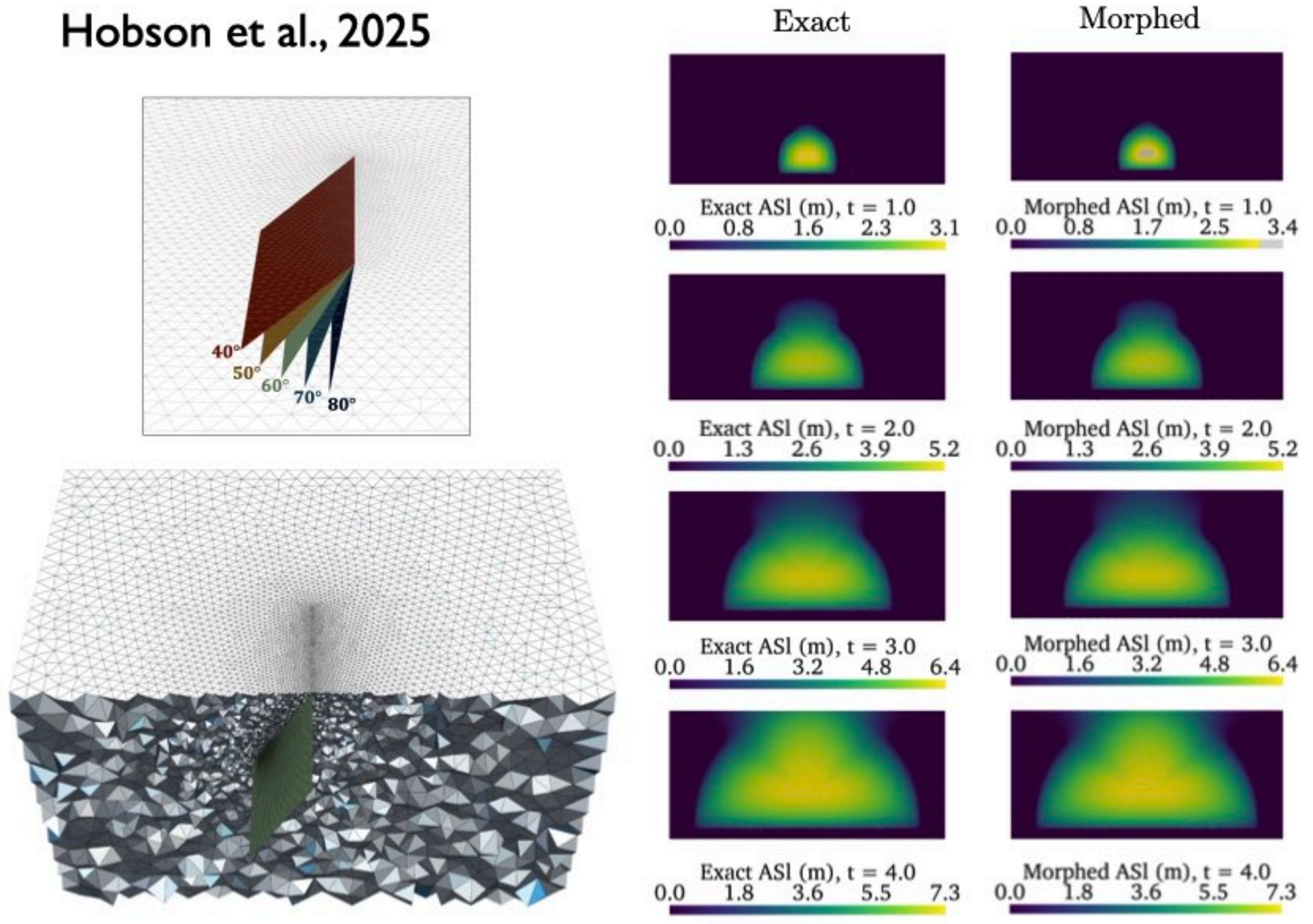


Rekoske et al., 2025



Magen et al., 2025

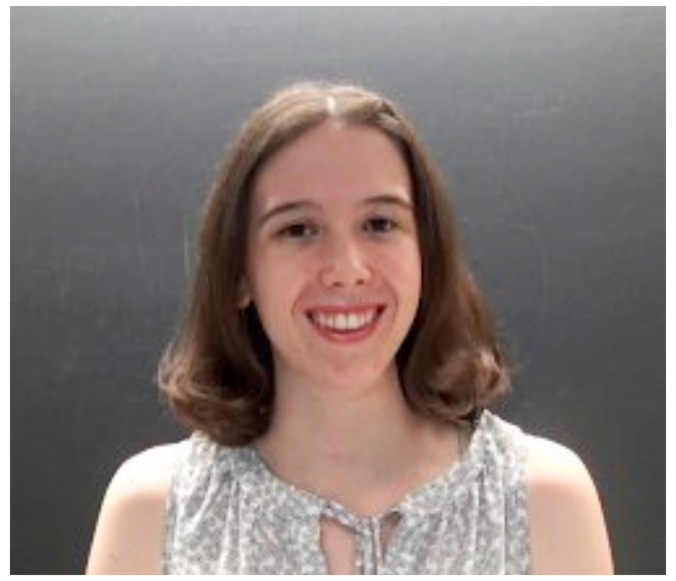
Hobson et al., 2025



Dave May

John Rekoske

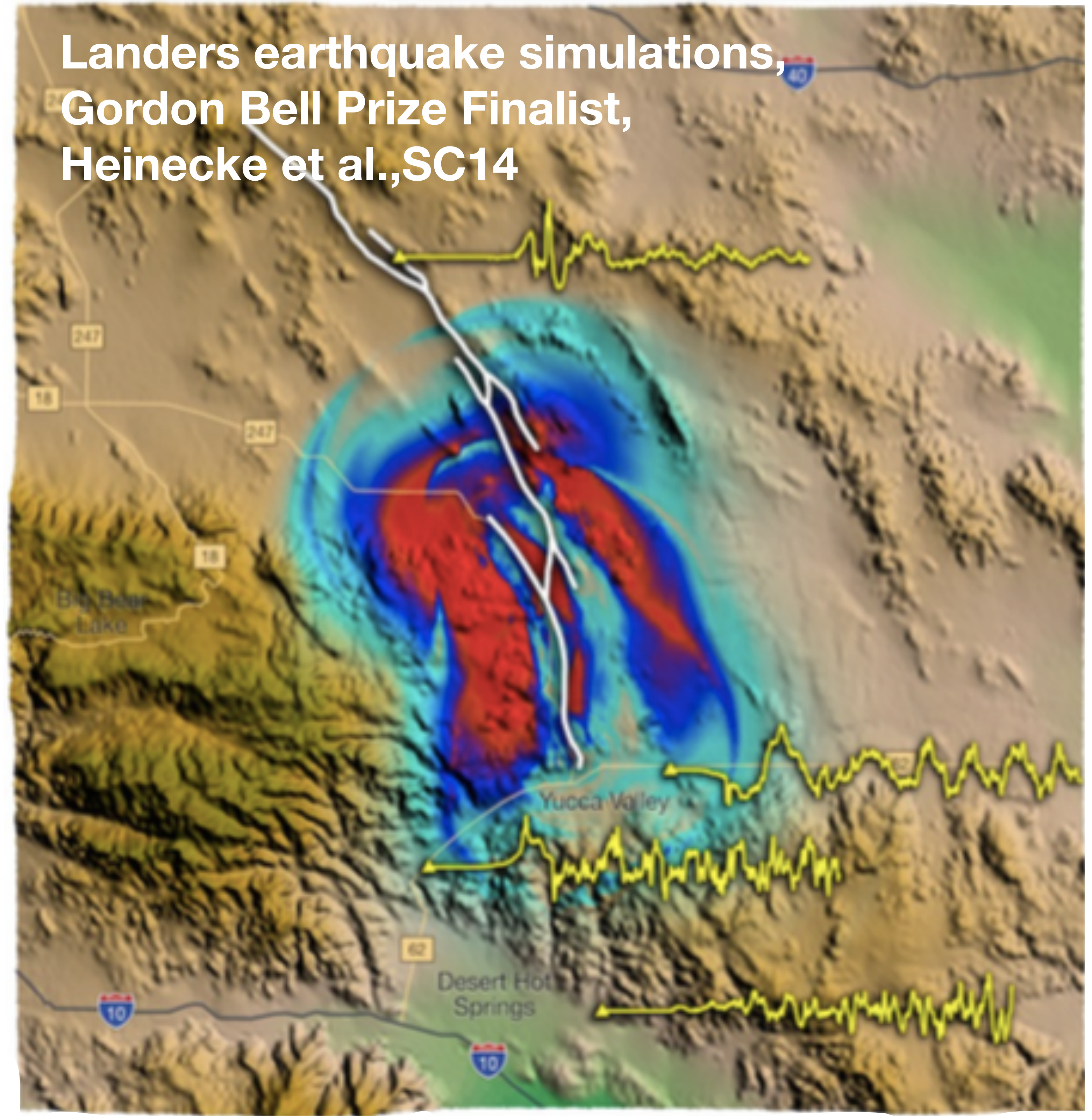
Yohai Magen

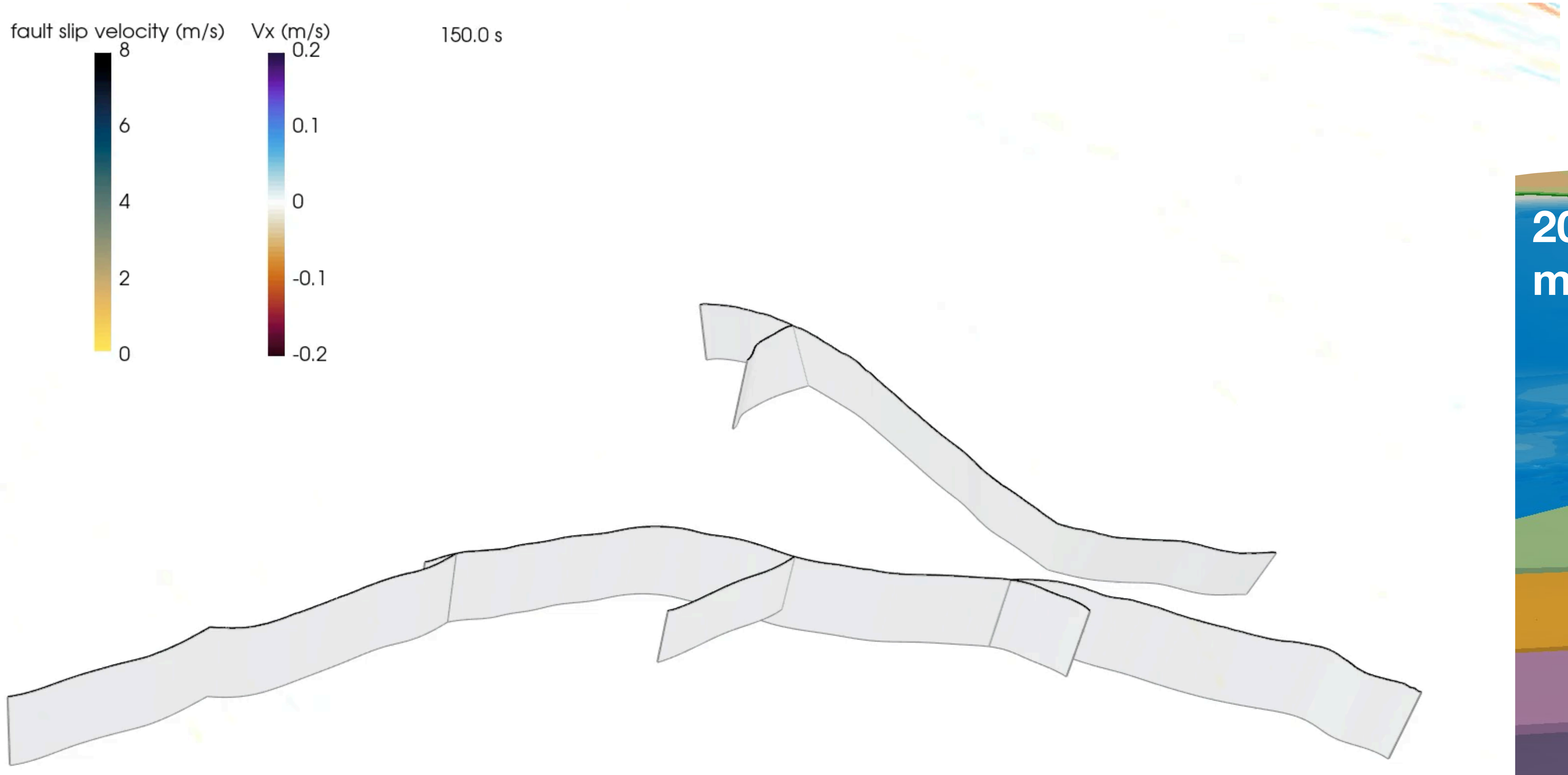


Gabrielle Hobson

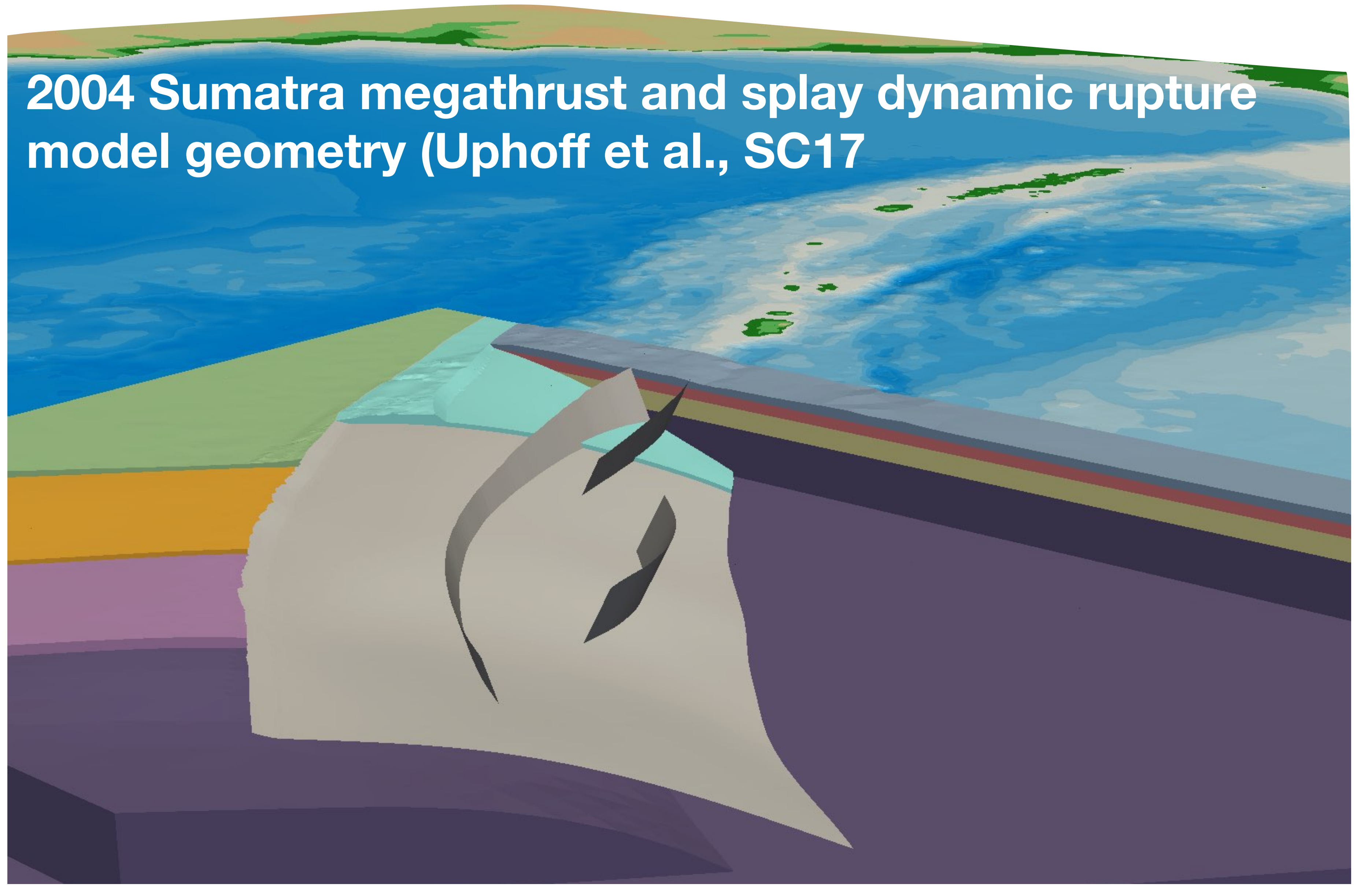
Earthquake simulations become ever more expensive

- 10 Hz Landers earthquake simulation reaching 8.6 PFLOPS on Tianhe-2: **96 billion DoF, 200,000 time steps** (Heinecke et al.,SC14, **Gordon Bell Prize Finalist**)
- **Sumatra tsunami earthquake:** 220 million finite elements (**~111 billion degrees of freedom**) and 3.3M time steps (Uphoff et al., SC17, Best Paper)
- **Rapid response: 5 Hz simulations of the 2023** Turkey earthquakes: 9h on full (~8200nodes) Frontera, and 685 million element mesh, with polynomial order 5 basis functions, **10e12 DoF**



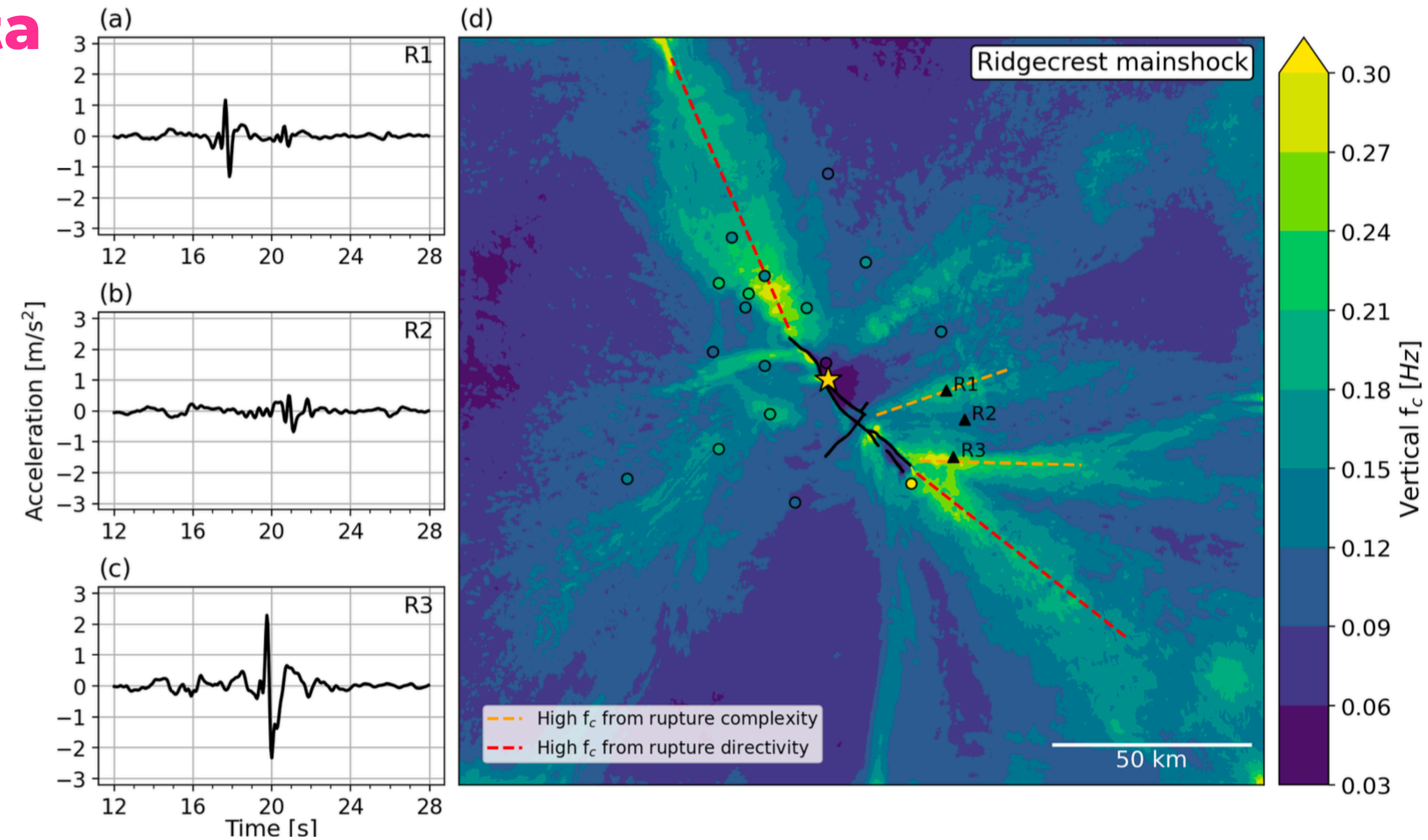


Jia et al., *Science*, 2023
Gabriel et al., *The Seismic Record*, 2023



... and produce ever bigger synthetic data

- 3D simulation output even of small simulations and 2D output of large simulations is **>10s of TeraByte**
- **Reduction is possible but limited:** using modern data formats (**hdf5**), **single precision** (50%), **file-system aware sequential output** (10-20%), or only storying 2D output (at Earth's surface / on faults, ~ hundred GB)
- **FAIR data sharing standards** achieved by archiving simulation input & parameters (~ hundred GB, **no output)**



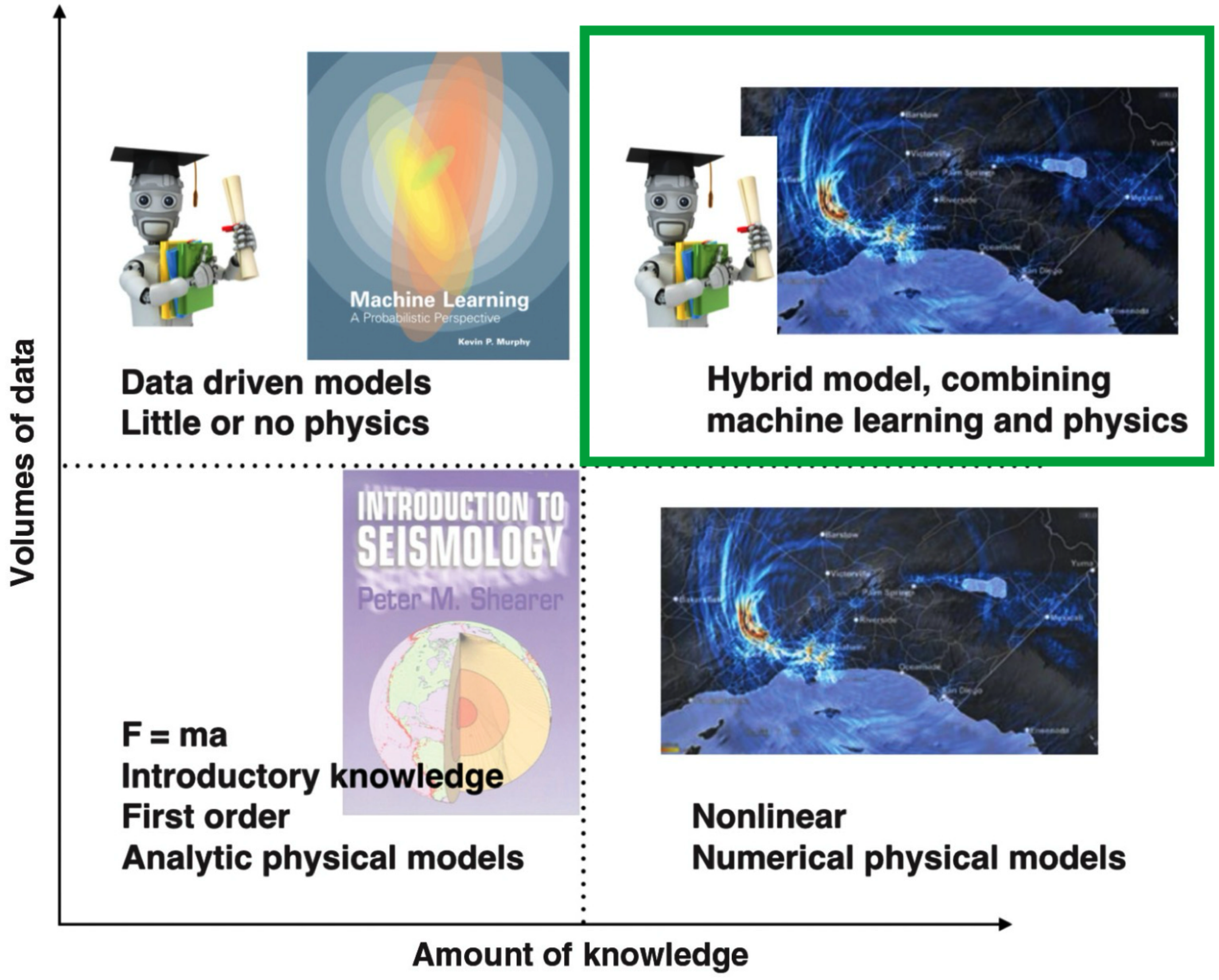
Schliwa & Gabriel, SRL, 2023: Ridgecrest mainshock's synthetic vertical ground accelerations at three selected stations. (d) Map view of the equivalent near-field corner frequency (fc) distribution of the vertical components of synthetic seismograms recorded at ~1,800,000 virtual seismic stations. The synthetic seismograms are generated from a complex dynamic rupture model of the 2019 Ridgecrest mainshock. Black lines indicate the fault traces, the star marks the epicenter, colored dots show fc values of recorded ground motion spectra, and triangles show the virtual station locations of the analyzed accelerograms. Orange and red lines mark different high-fc features. Bottom: Peak dip-slip isochrones of stations R1 and R3.

Visualization of 15 TB of 3D volumetric data on unstructured tetrahedral meshes on Frontera.
Abrams et al., SC'23, Visualization showcase

Reduced-order modeling for on-demand and physics-informed earthquake model surrogates

- None of the existing physics-based simulation methods are efficient enough for real-time (**early warning**) or routine full physics-based **probabilistic seismic hazard assessment** (evaluating 10,000s of complex models)

- **The results of data-intensive computations can be used to construct surrogates, e.g., reduced-order models using low dimensional information which enable the evaluation of new earthquake & seismic cycle scenarios near-instantaneously**



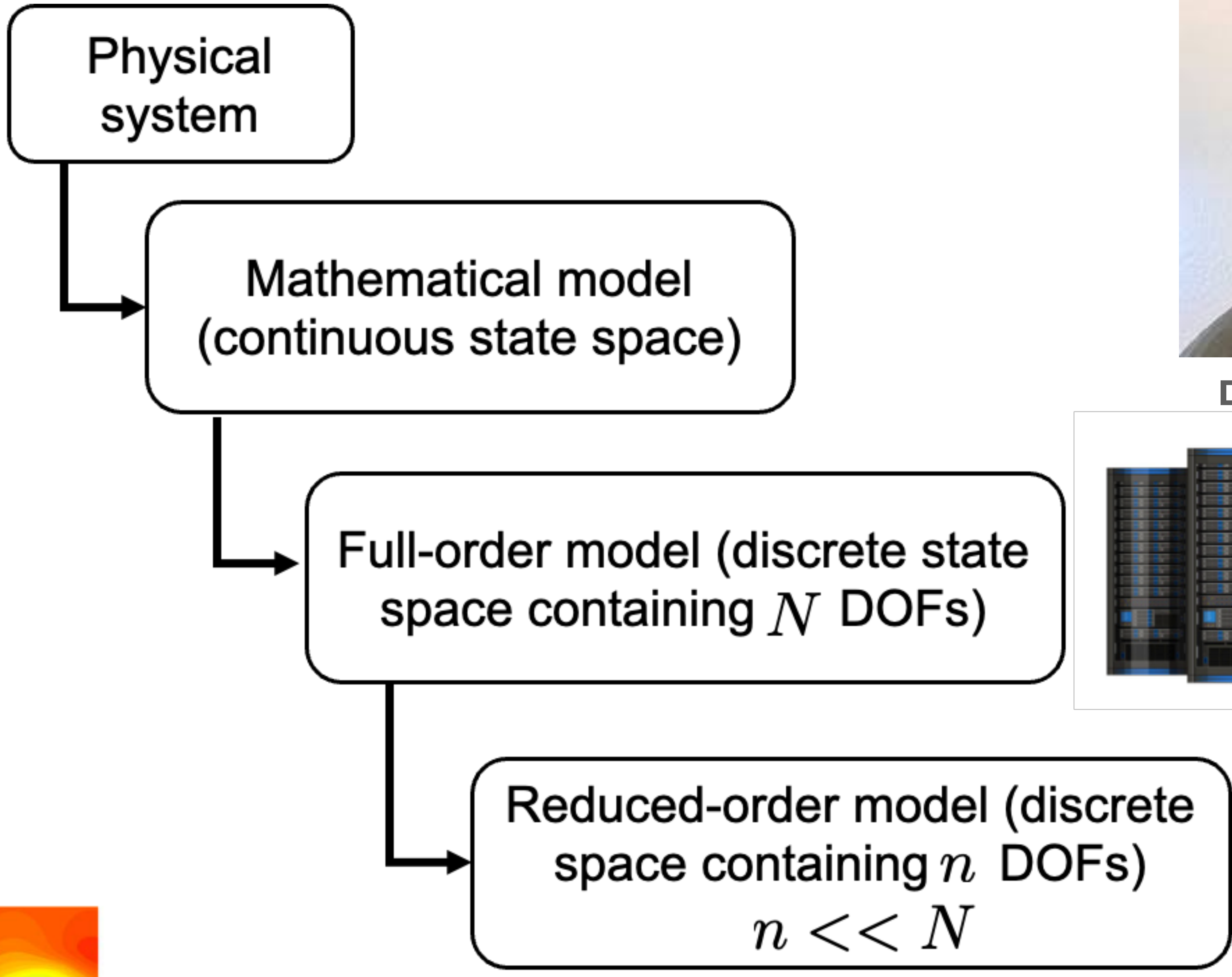
Kong et al. (2019)

Reduced-Order Models (ROMs)

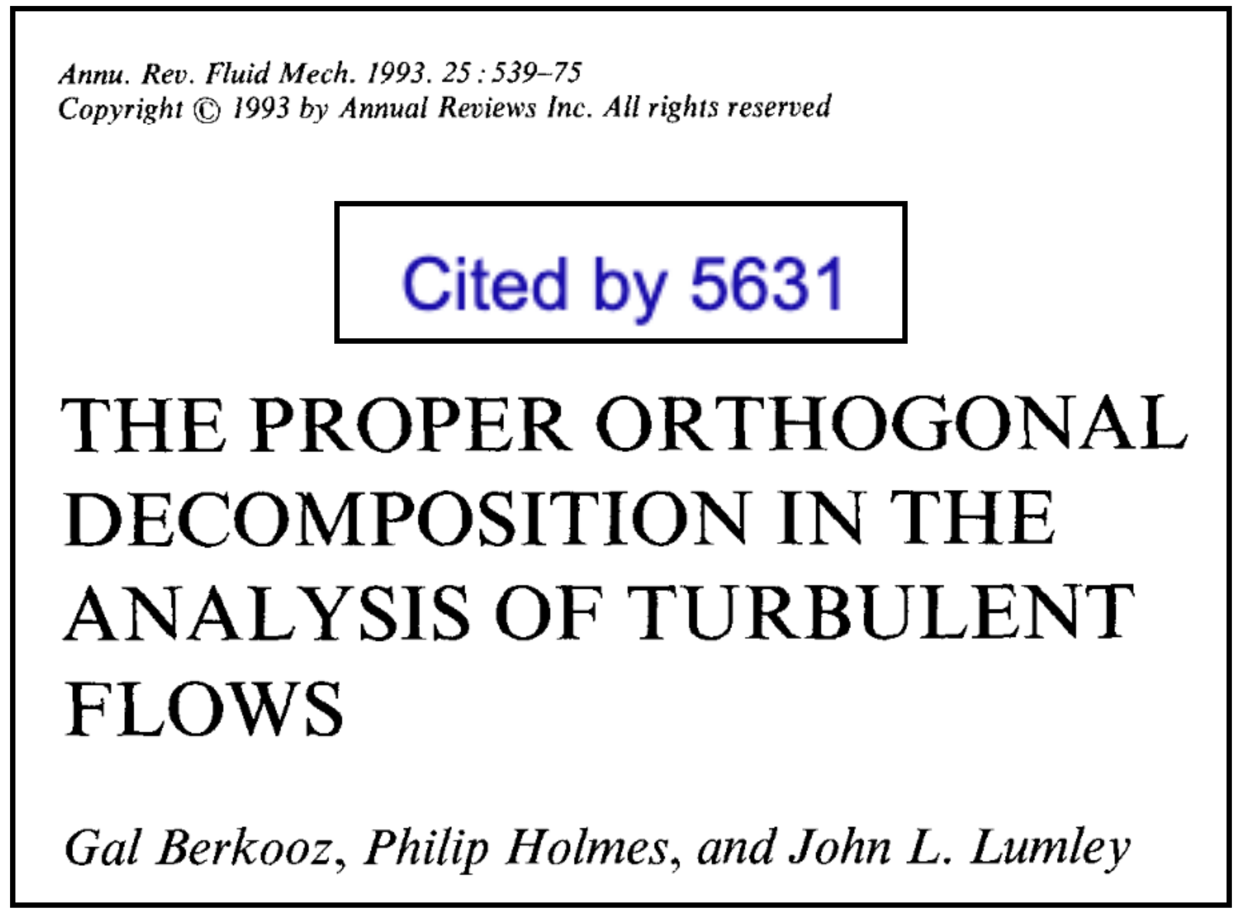
- Identify and exploit a **basis** to represent families of PDE solutions

The choice of basis is key!

- **Project** high-fidelity simulations (large discretizations) onto a low-dimensional subspace
- Classic origins: **fluid mechanics / turbulence studies** (Lumley, **1967**)
- Goal: **MUCH** faster evaluation, **while retaining essential physics**



Dave May (UCSD)

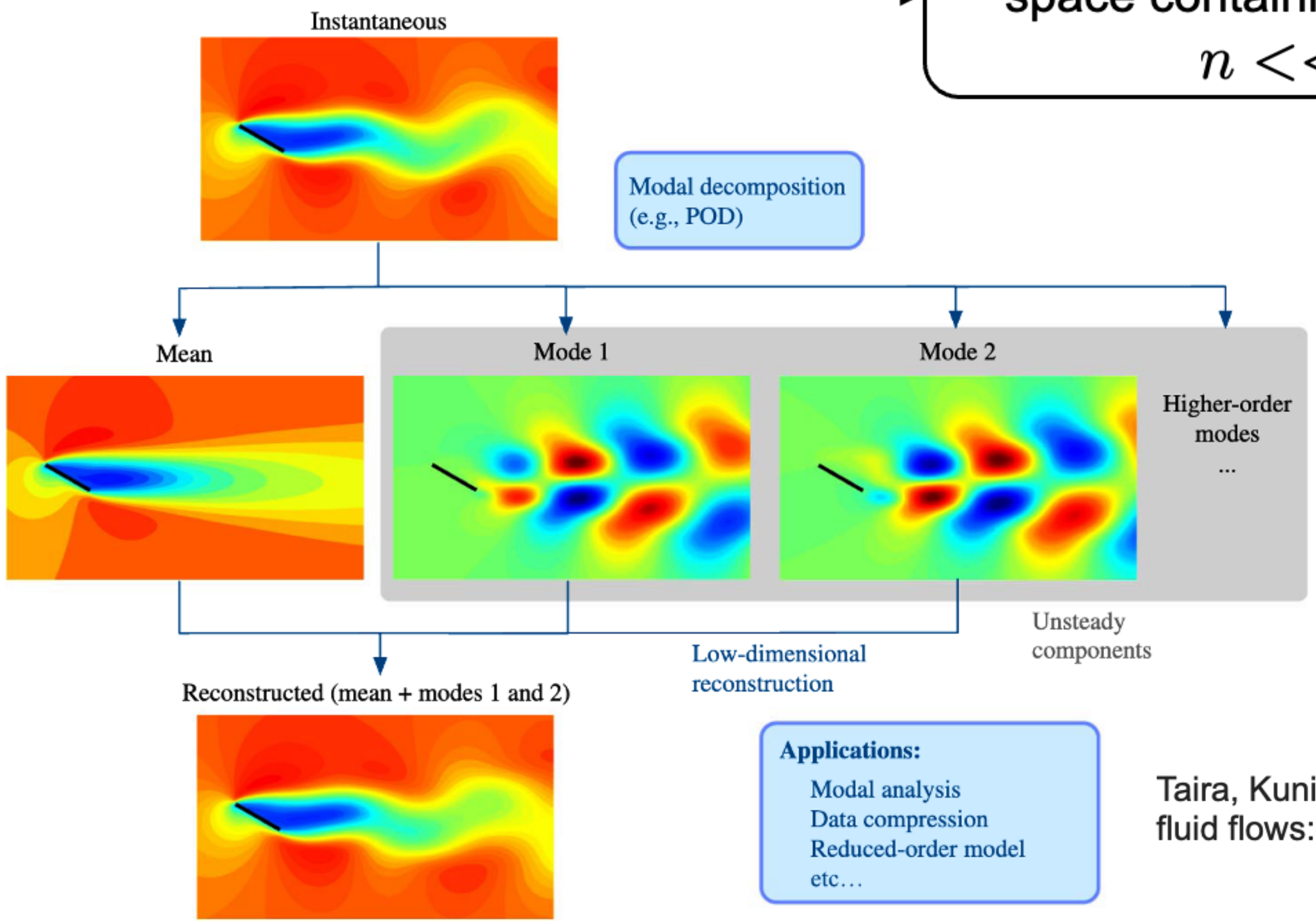


Annu. Rev. Fluid Mech. 1993. 25 : 539–75

THE PROPER ORTHOGONAL DECOMPOSITION IN THE ANALYSIS OF TURBULENT FLOWS

Gal Berkooz, Philip Holmes, and John L. Lumley

(Also: Sirovich, 1987; Holmes et al., 2012)



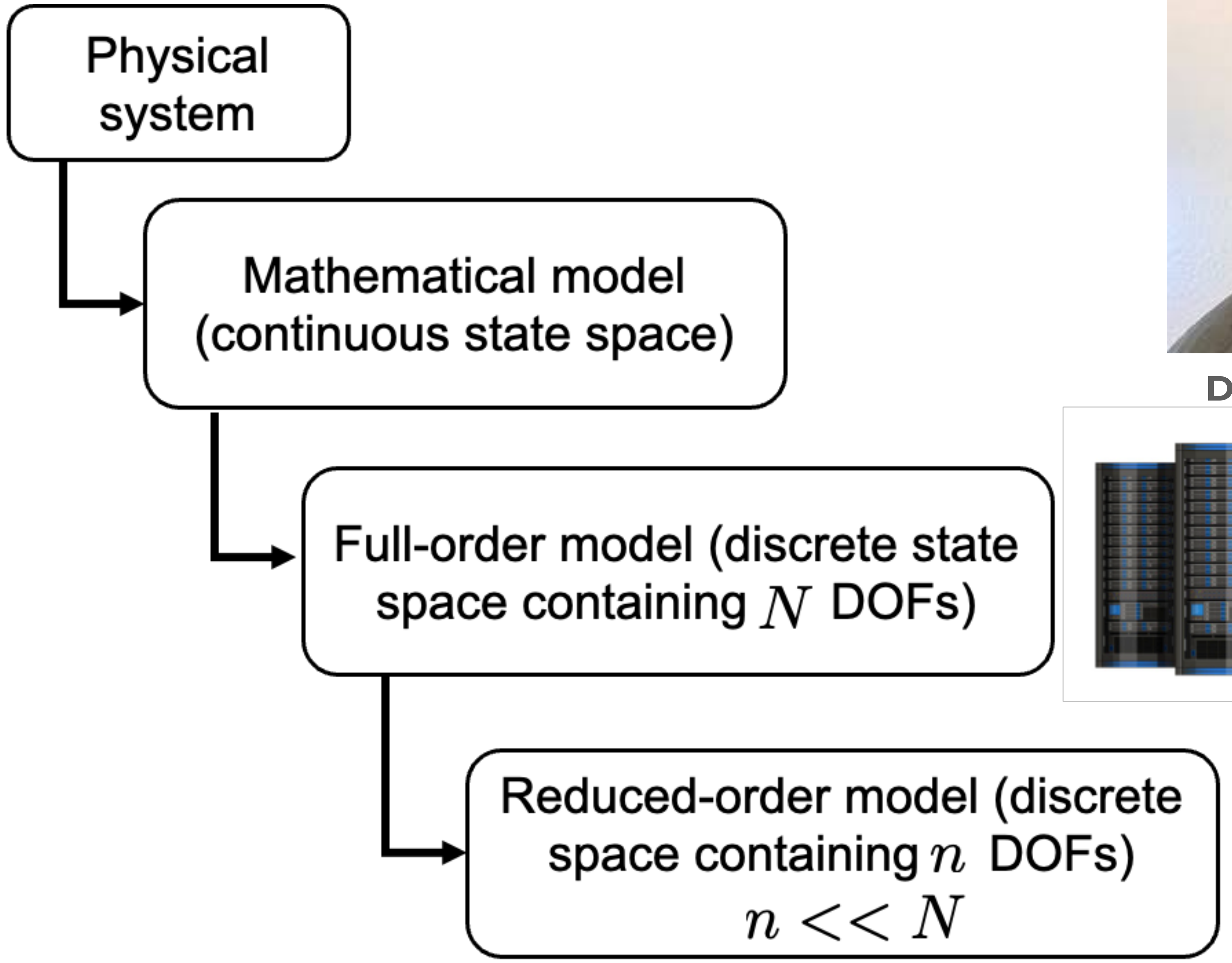
Taira, Kunihiko, et al. "Modal analysis of fluid flows: An overview." *AIAA journal*

Reduced-Order Models (ROMs)

- Identify and exploit a **basis** to represent families of PDE solutions

The choice of basis is key!

- **Project** high-fidelity simulations (large discretizations) onto a low-dimensional subspace
- Classic origins: **fluid mechanics / turbulence studies** (Lumley, **1967**)
- Goal: **MUCH** faster evaluation, **while retaining essential physics**




Dave May (UCSD)

- **Advantage**: **Not a black-box, verifiable SciML! iPOD inherits** error analysis **from interpolation theory!**
- **Parallel, scalable software (C, PETSc) implementation capable to work with > 1 billion DoF snapshots + accessible Python implementation**
- **Limitations**: extrapolation in parameter space challenging; design of parameter space is problem-dependent (our current form is suitable for <10 dimensions)

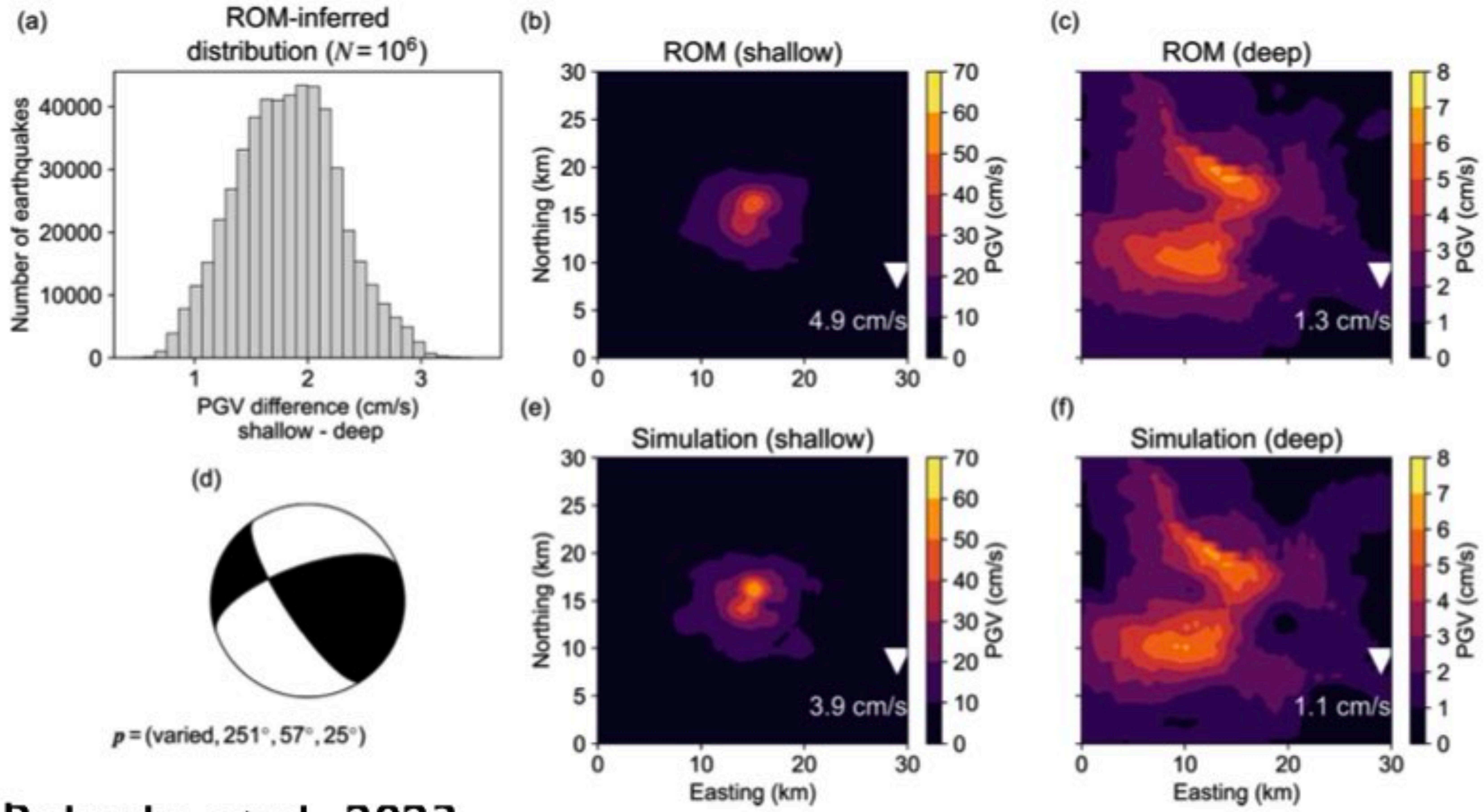
JGR Solid Earth

Research Article | Free Access

Instantaneous Physics-Based Ground Motion Maps Using Reduced-Order Modeling

John M. Rekoske, Alice-Agnes Gabriel, Dave A. May

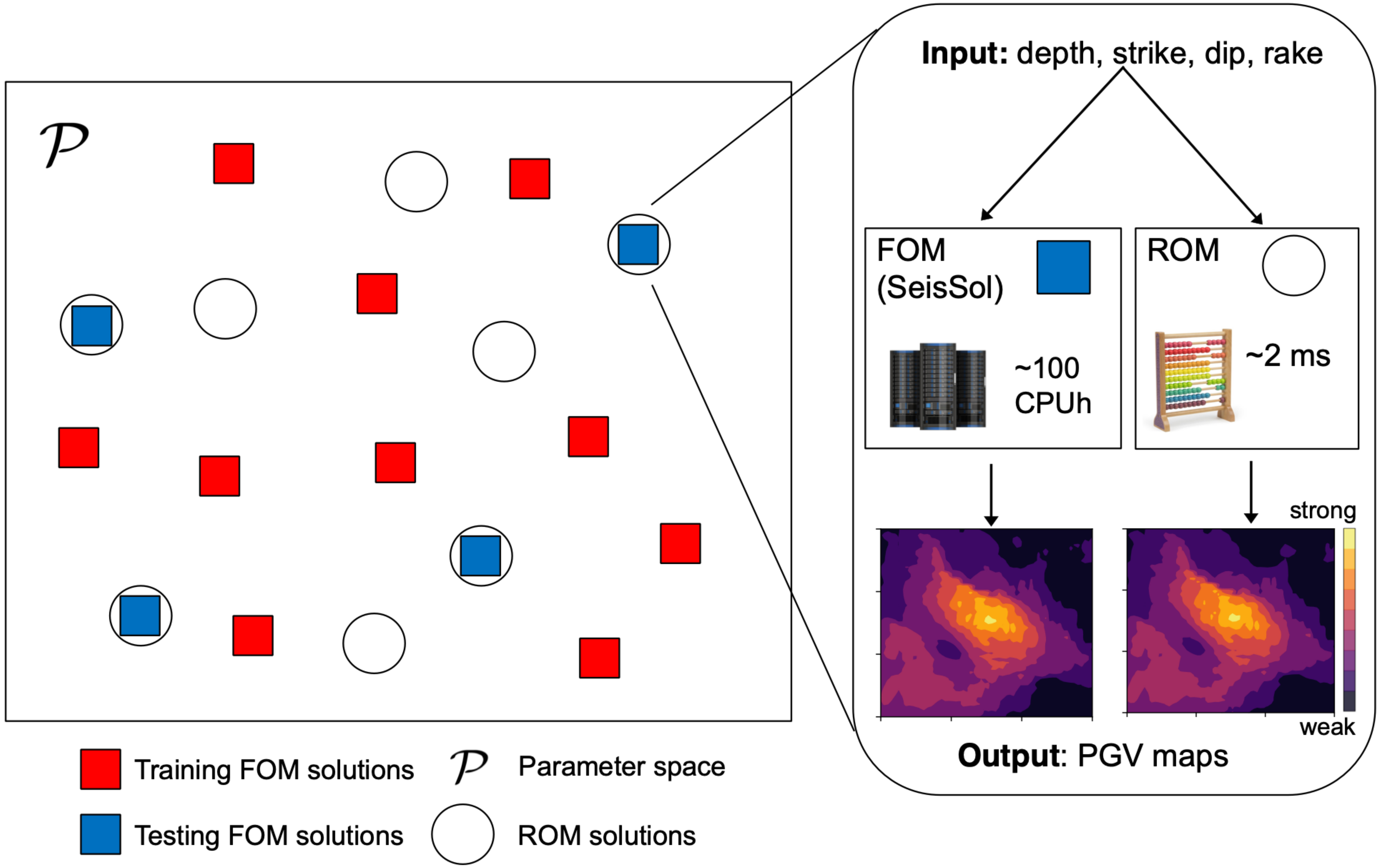
John Rekoske



Rekoske et al., 2023

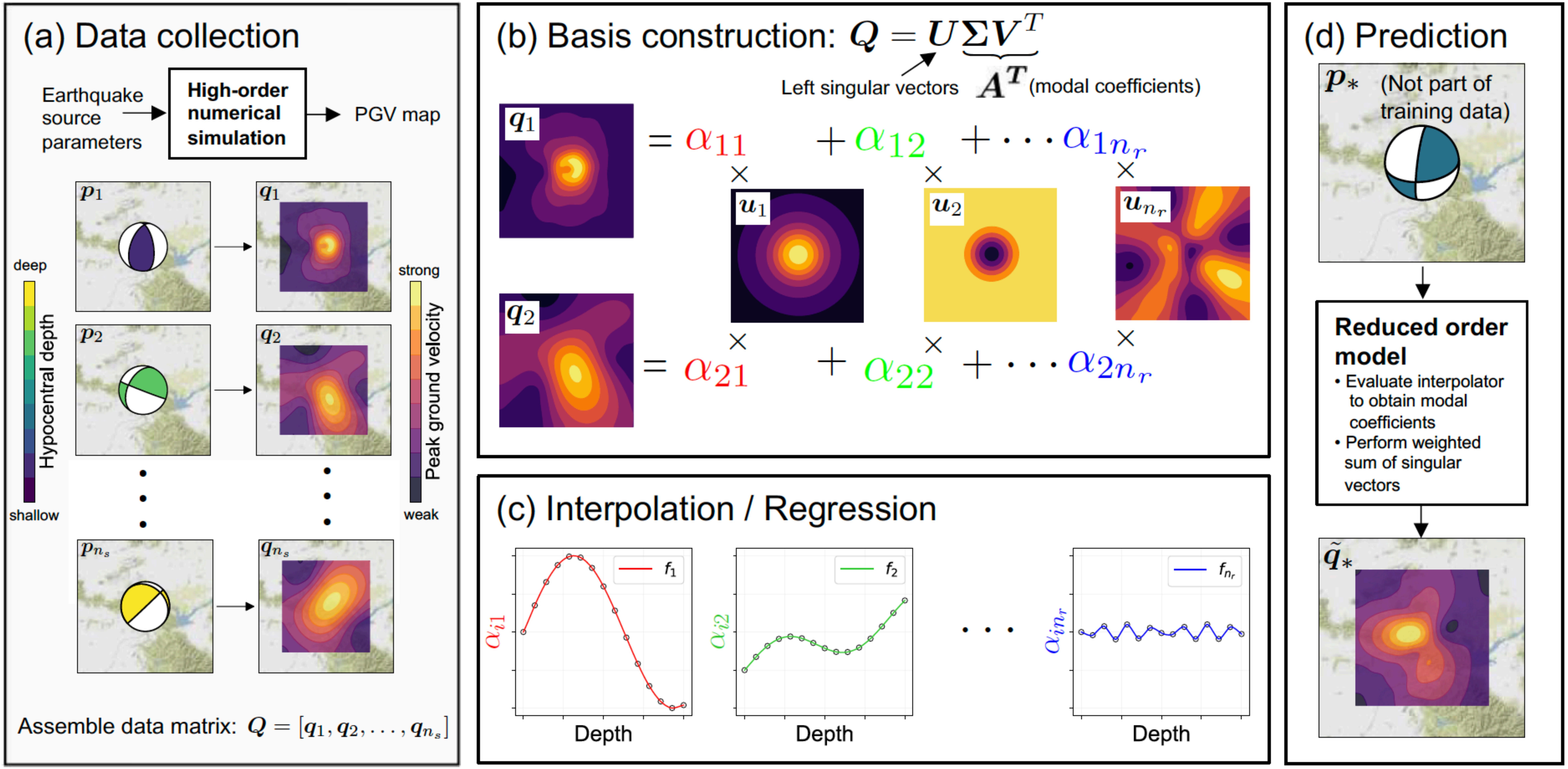
Rapid ground motion maps using data-driven ROMs

- 1 Hz **3D wave propagation simulations** (SeisSol) with **topography**, viscoelastic **attenuation**, **3D velocity** model (Vs>500 m/s) and varying earthquake point sources
- iPOD ROM for instantaneous predictions of **peak-ground velocity (PGV) maps**
- Comparison of different types of **interpolators** used in iPOD

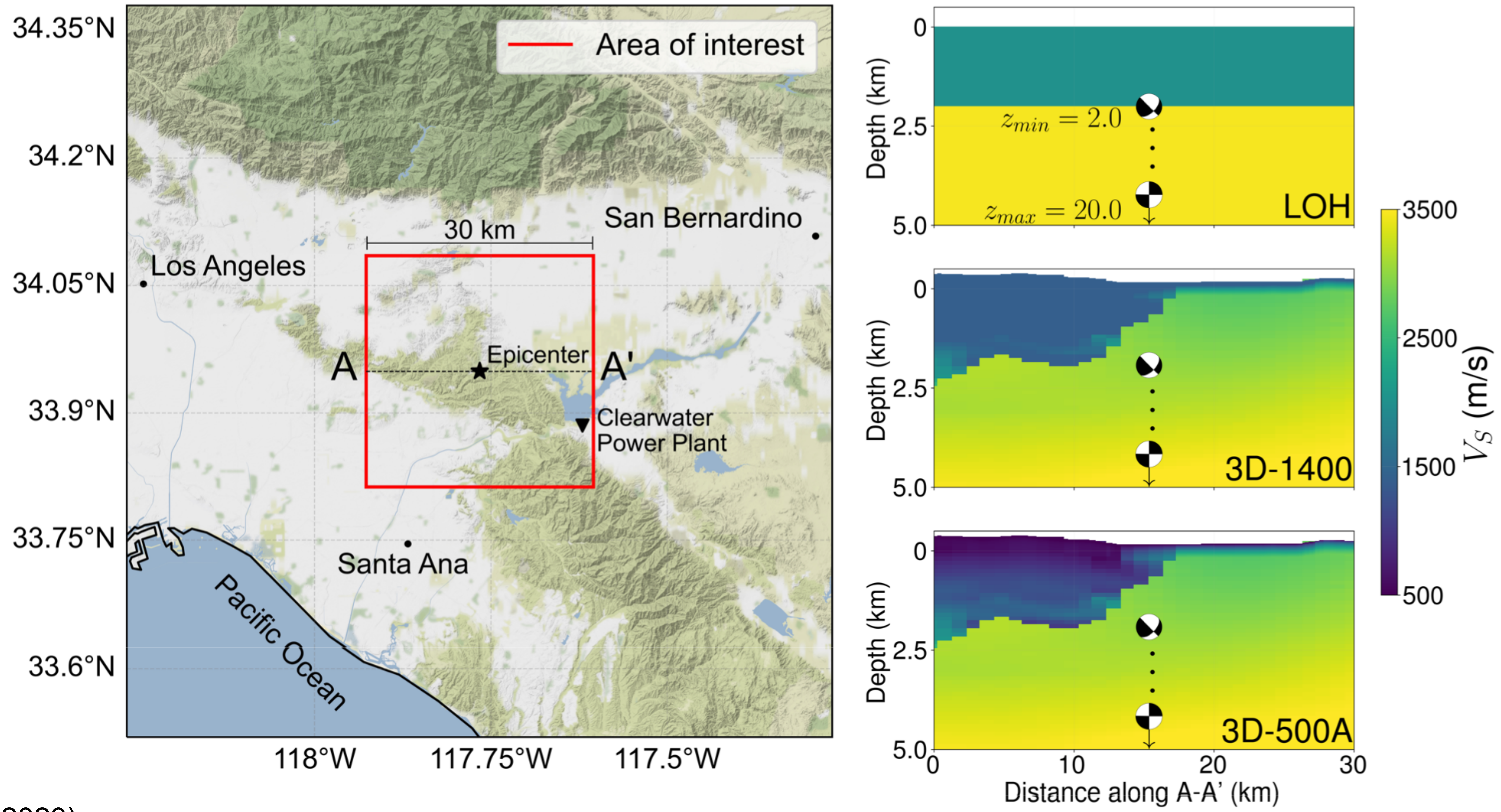


Rekoske, J. M., Gabriel, A.-A., & May, D. A. (2023). Instantaneous physics-based ground motion maps using reduced-order modeling. *Journal of Geophysical Research: Solid Earth*, 128, e2023JB026975. https://doi.org/10.1029/2023JB026975

Interpolated proper orthogonal decomposition (iPOD)



Full-order model (FOM): PGV maps for varying earthquake source depths and focal mechanisms

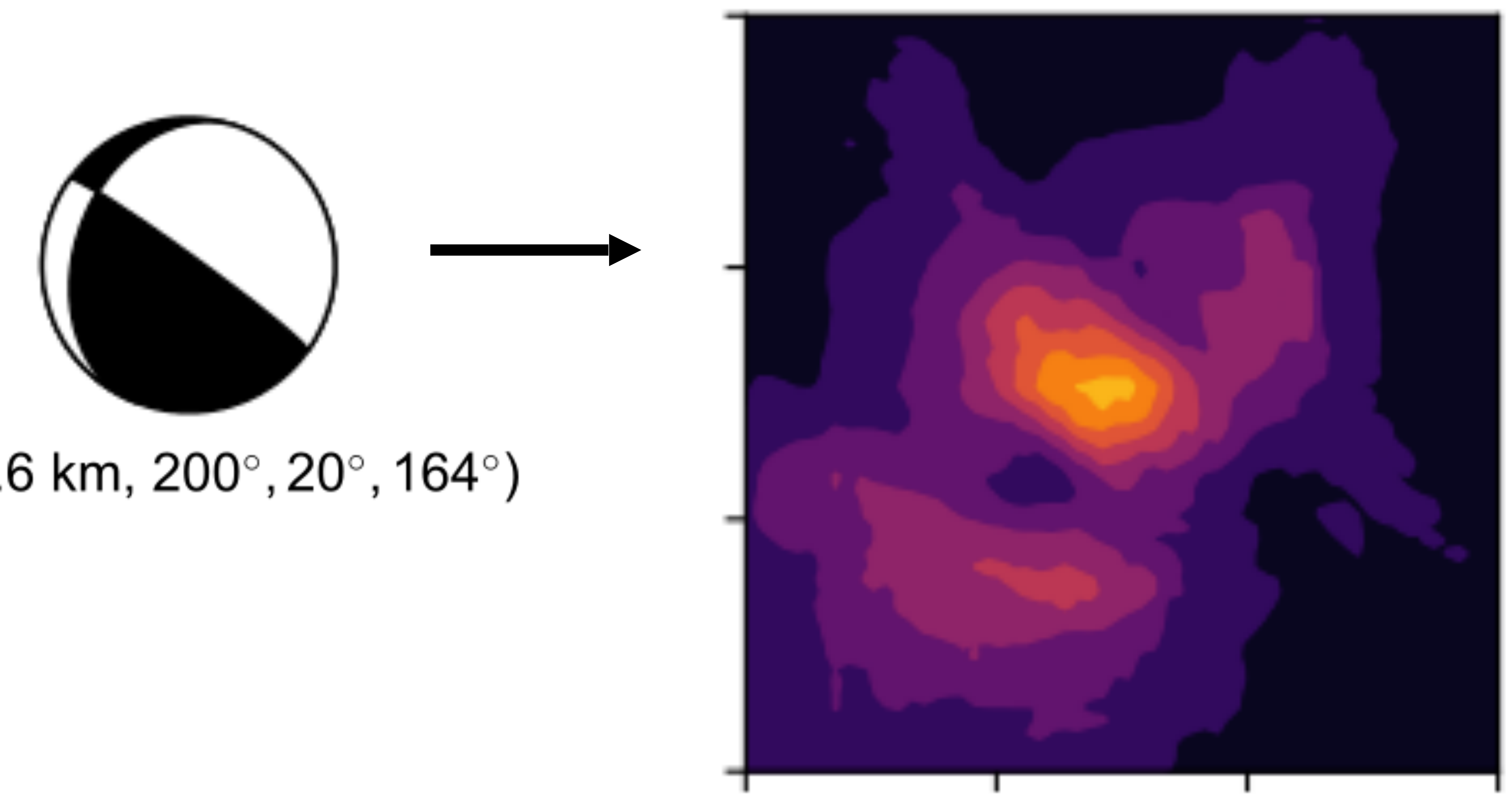


Rekoske et al. (2023)

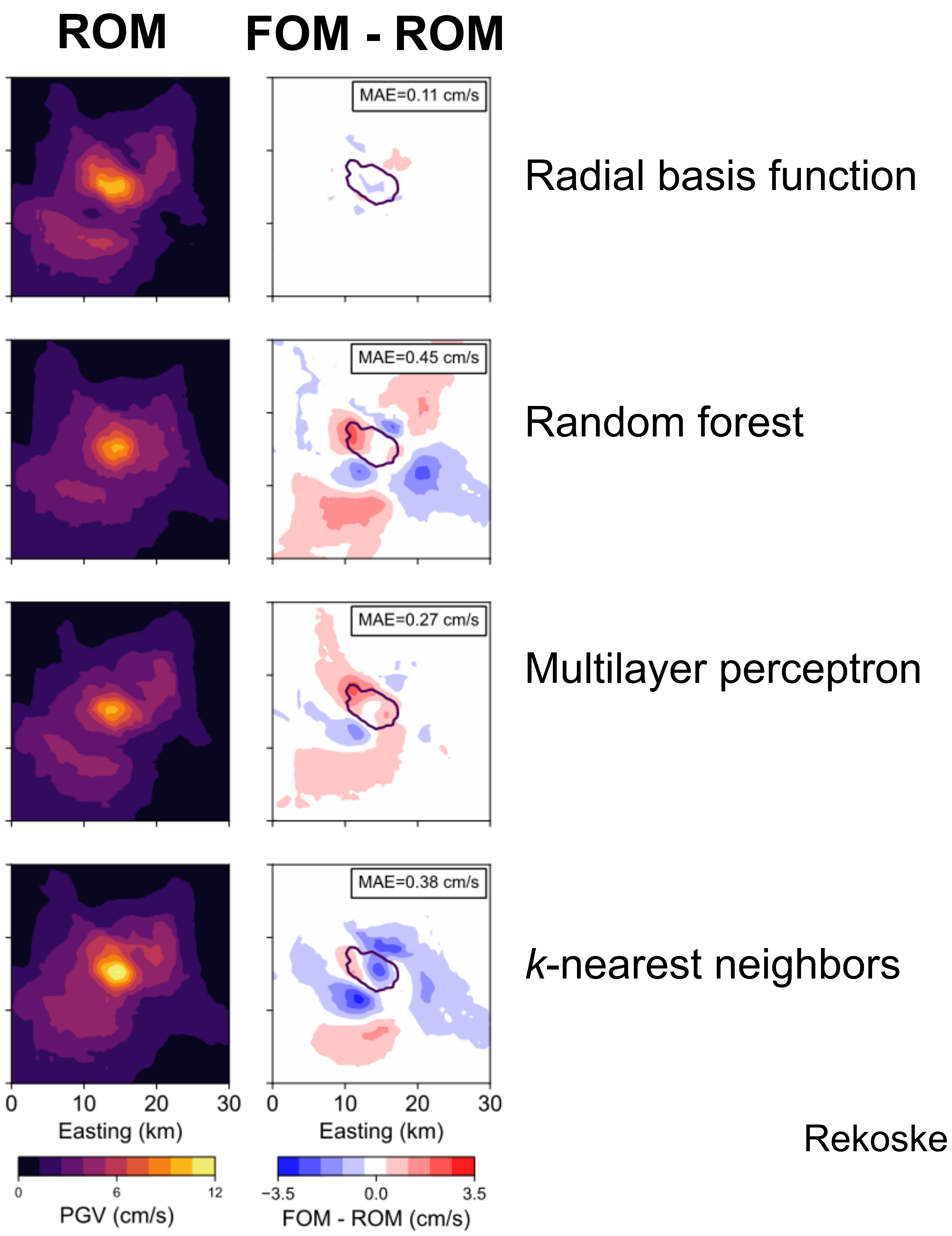
Accuracy comparison of predicted PGV maps for different iPOD interpolators

- Comparing **radial basis function interpolation** (<u>**no hyper parameters!)**</u> with **various ML regression methods** (k-nearest neighbor, multi-layer perceptron, random forest)
- **RBF is uniformly superior in accuracy independent of model complexity**

FOM PGV map (3D-500A)



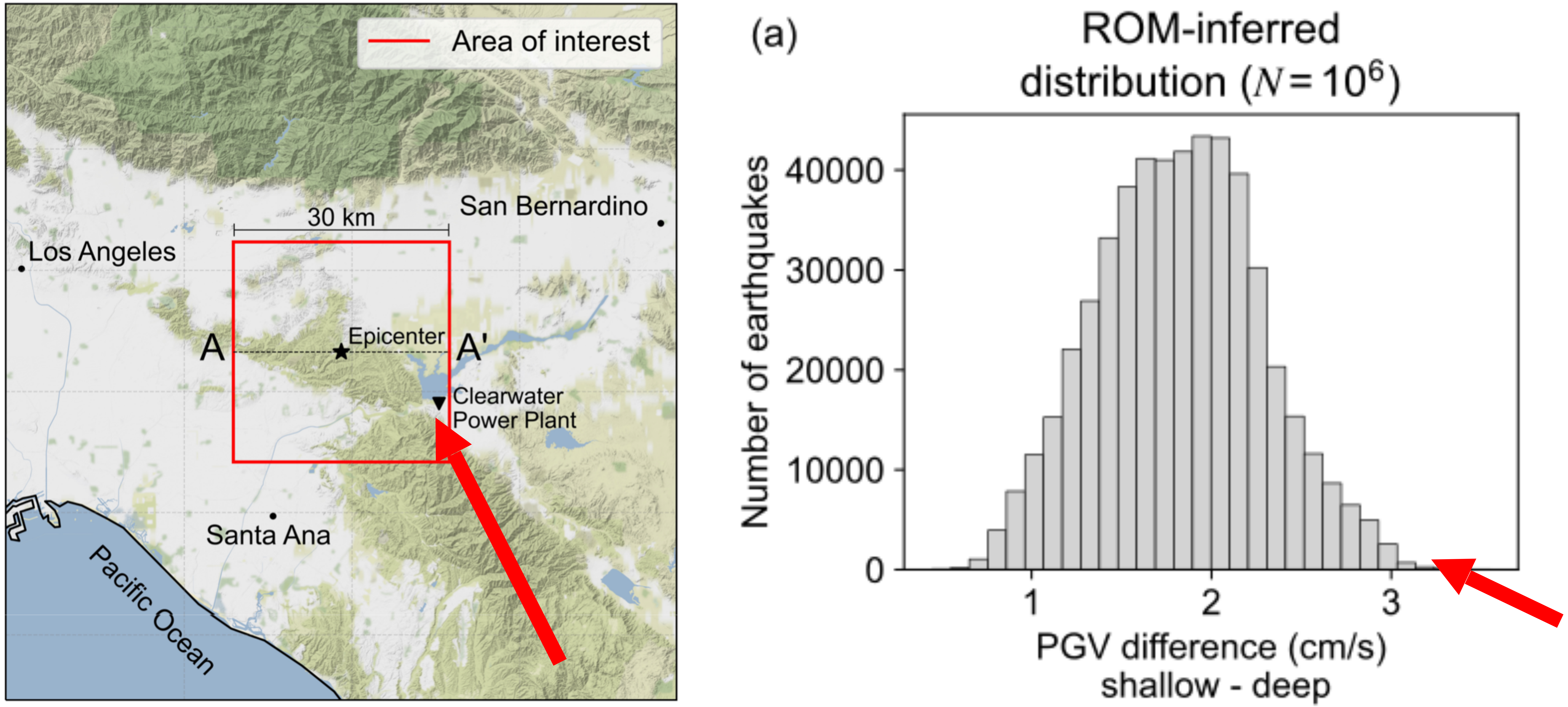
$p = (6.6\text{ km}, 200°, 20°, 164°)$

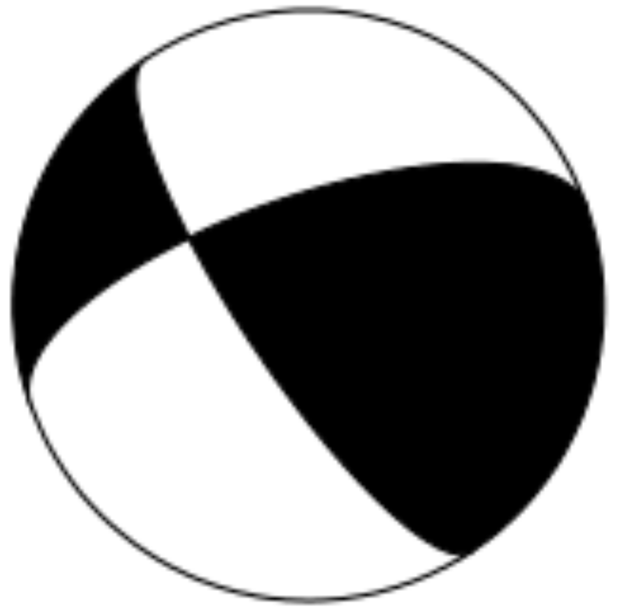


Rekoske et al. (2023)

Solving a maximization problem with our ROM by evaluating millions of scenarios - Ensemble forecasting

- Example application: **A worst-case scenario**
- **ShakeAlert fixes earthquake depths at 8 km**: What is the range of error that could be introduced by this assumption? What is the worst possible case?
- Evaluate difference in ROM for shallow (2 km) and deep (8 km) **for one million focal mechanisms**: in the worst-case scenario, the PGV predictions based on an 8.0 km hypocentral depth could underestimate the true PGV at this site by up to 3.6 cm/s if the true hypocentral depth is 2.0 km





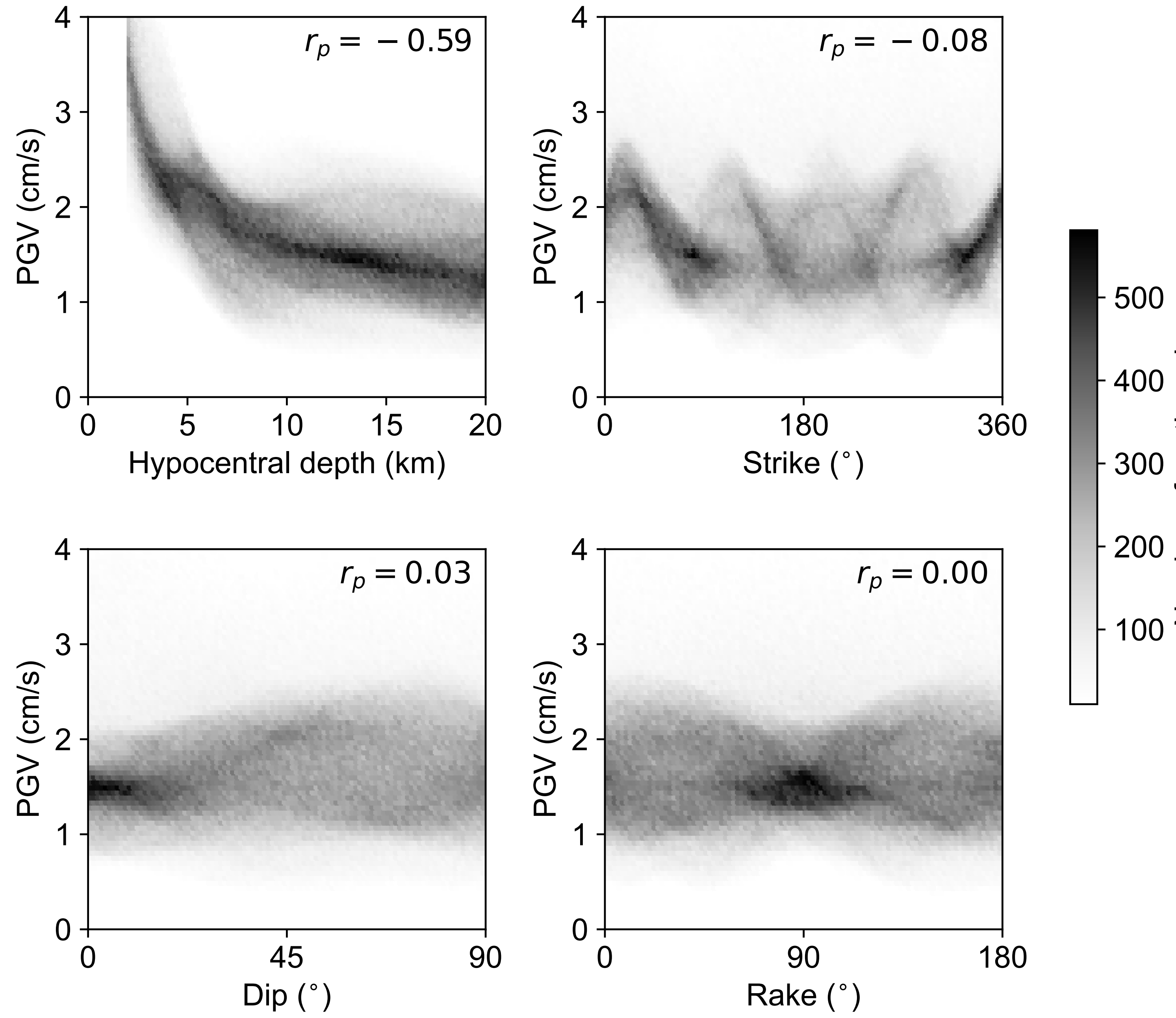
The ROM-identified worst-case scenario

$p = (\text{varied}, 251°, 57°, 25°)$

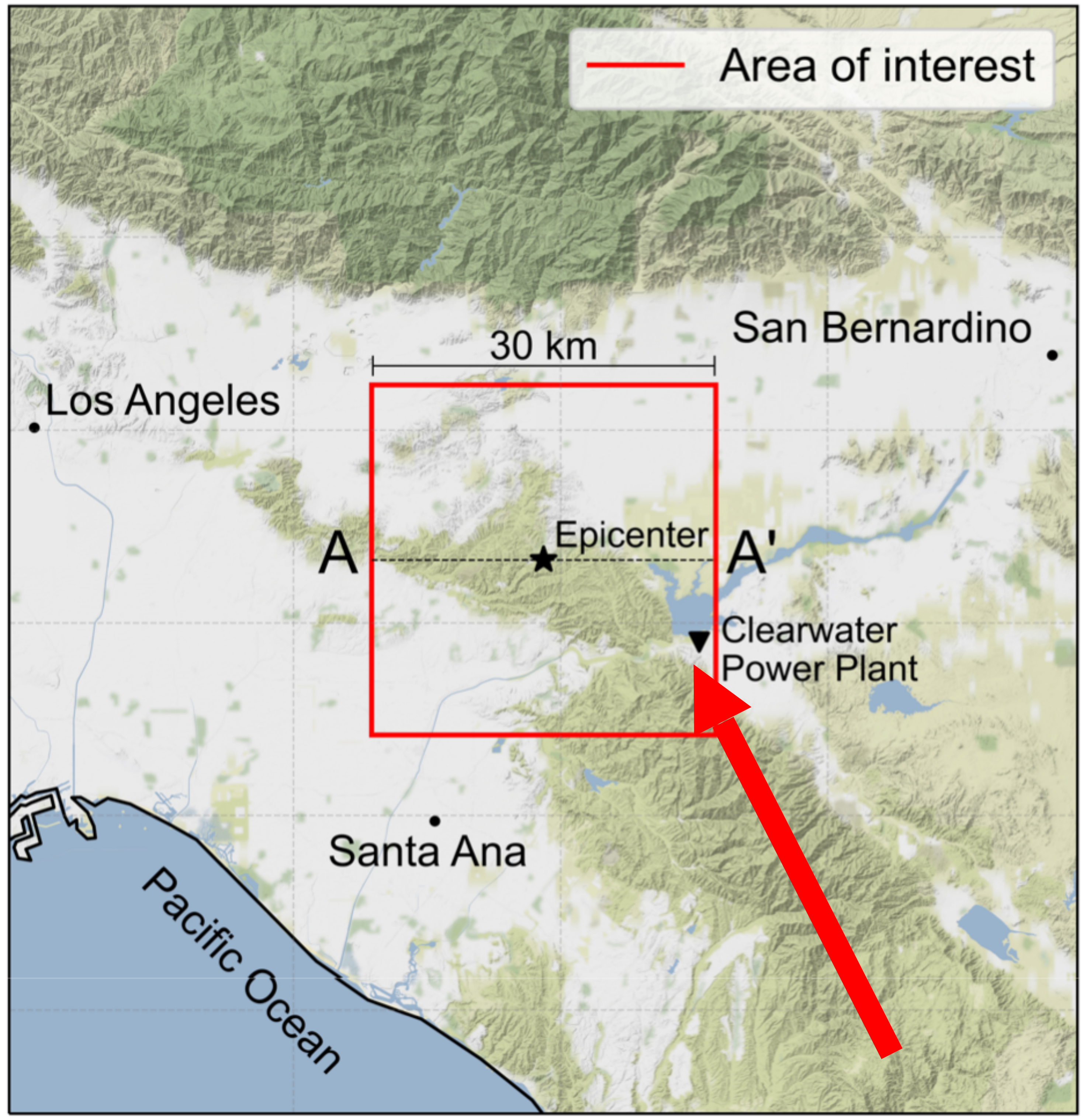
Rekoske et al. (2023)

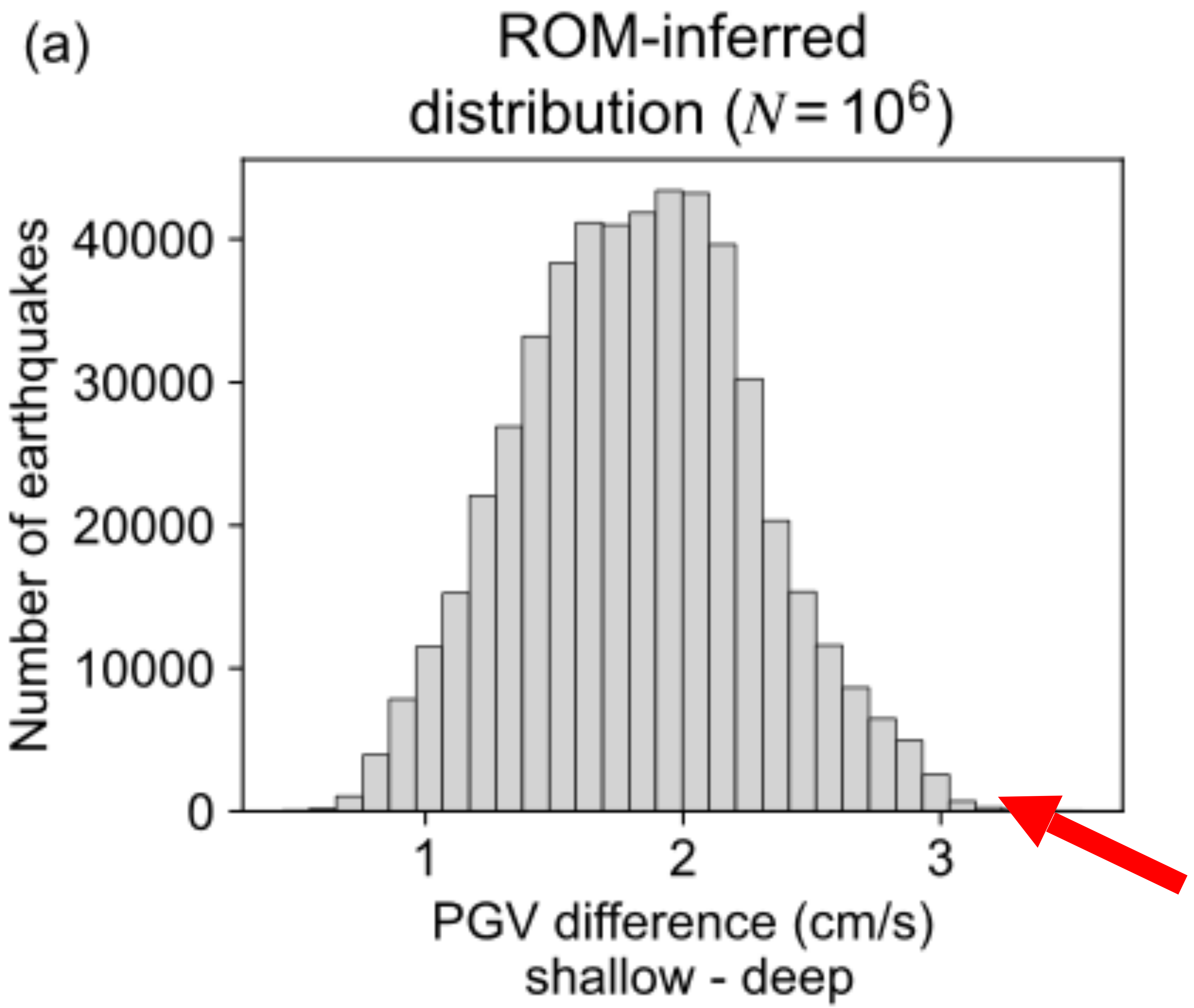
Solving a maximization problem with our ROM by evaluating millions of scenarios - Ensemble forecasting

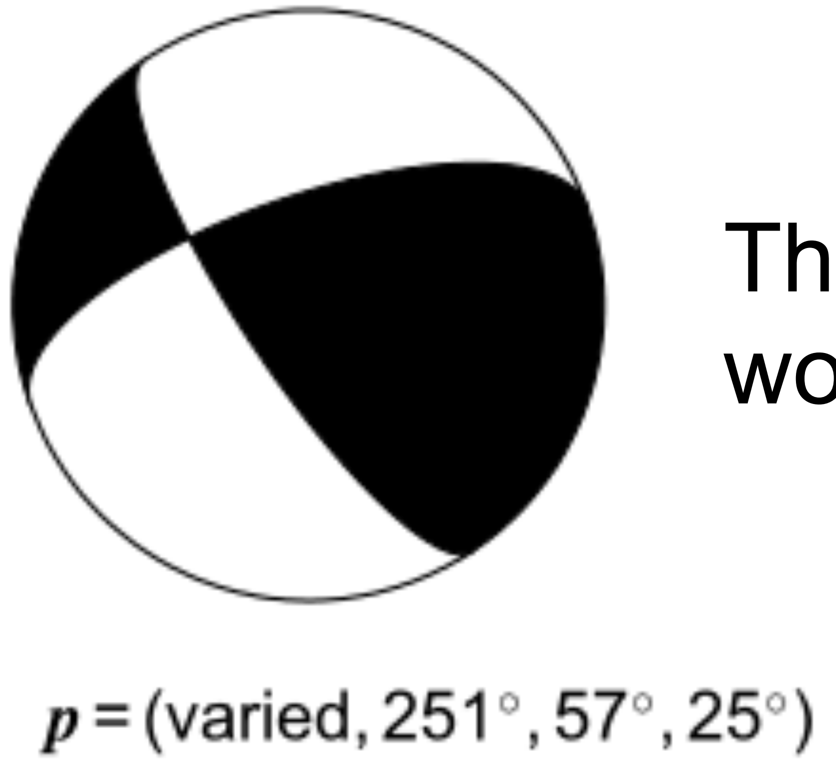
- Example application: **A worst-case scenario**
- **ShakeAlert fixes earthquake depths at 8 km**: What is the range of error that could be introduced by this assumption? What is the worst possible case?
- Evaluate difference in ROM for shallow (2 km) and deep (8 km) **for one million focal mechanisms**: in the worst-case scenario, the PGV predictions based on an 8.0 km hypocentral depth could underestimate the true PGV at this site by up to 3.6 cm/s if the true hypocentral depth is 2.0 km



global sensitivity analysis / forward uncertainty quantification





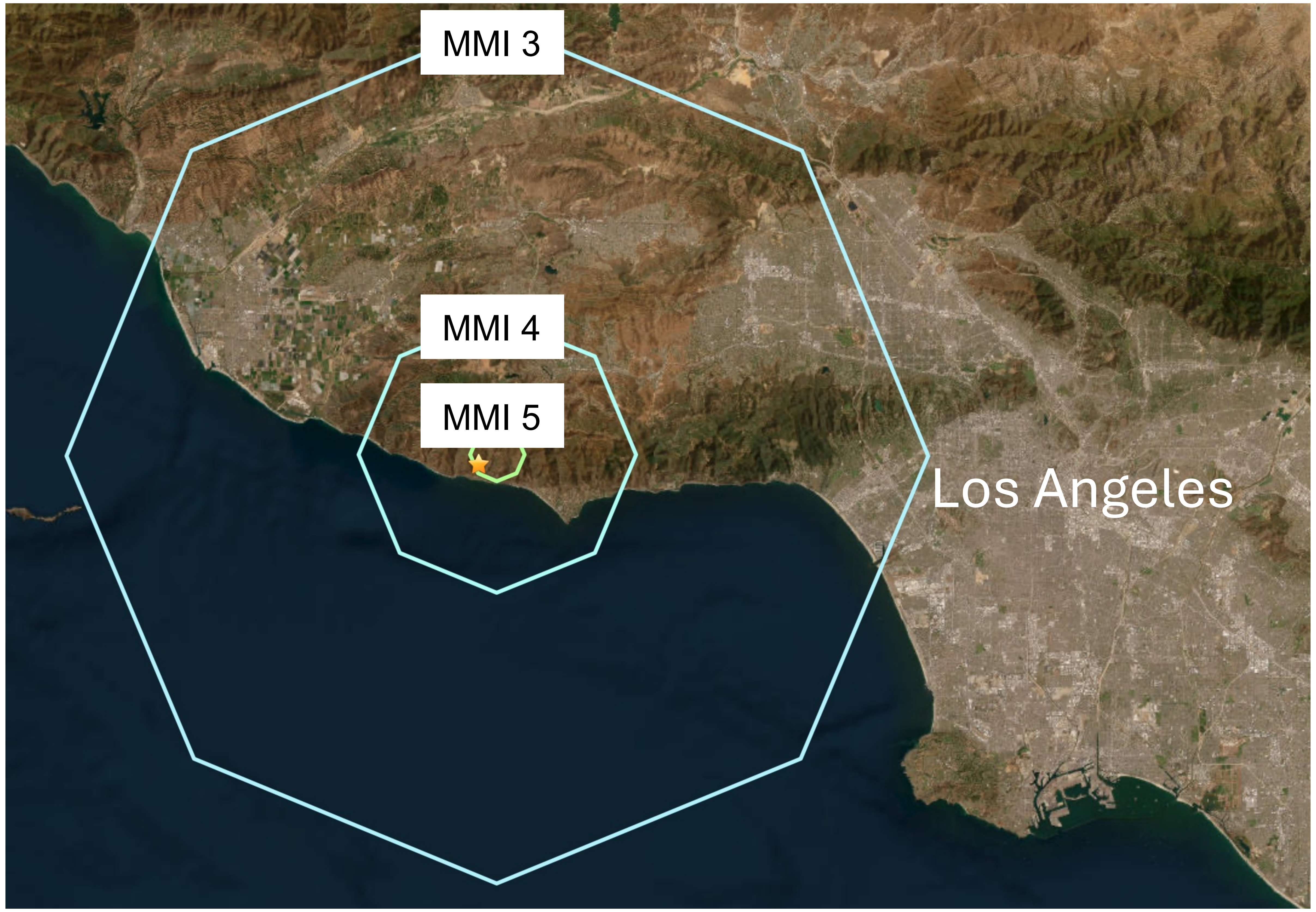


The ROM-identified worst-case scenario

Rekoske et al. (2023)

Rapidly predicted shaking contours may not always match real shaking due to source or site effects.

ShakeAlert MMI contours



ShakeMap MMI contours

Example: **M**4.6 Malibu, CA Earthquake, February 2024
ShakeAlert final estimated magnitude: 4.7

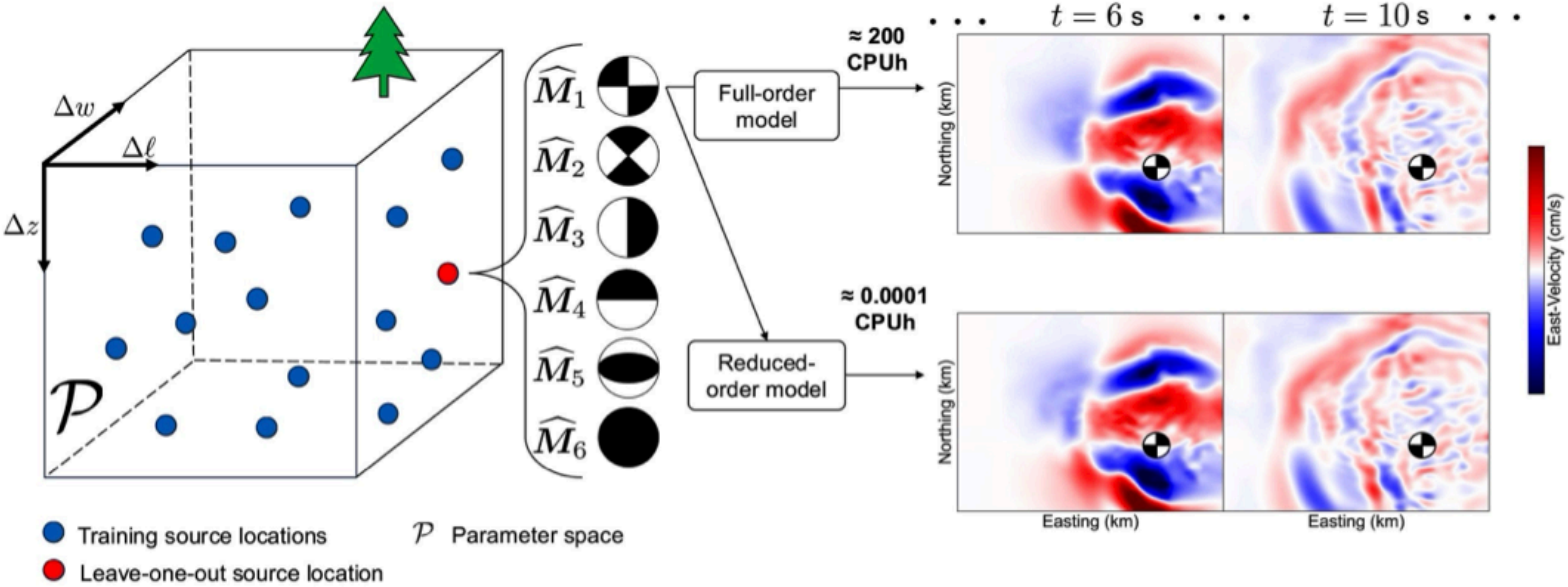
USGS EEW proposal recommended for funding

Reduced-order modelling for complex three-dimensional seismic wave propagation

John M Rekoske, Dave A May, Alice-Agnes Gabriel

Geophysical Journal International, Volume 241, Issue 1, April 2025, Pages 526–548,
https://doi.org/10.1093/gji/ggaf049

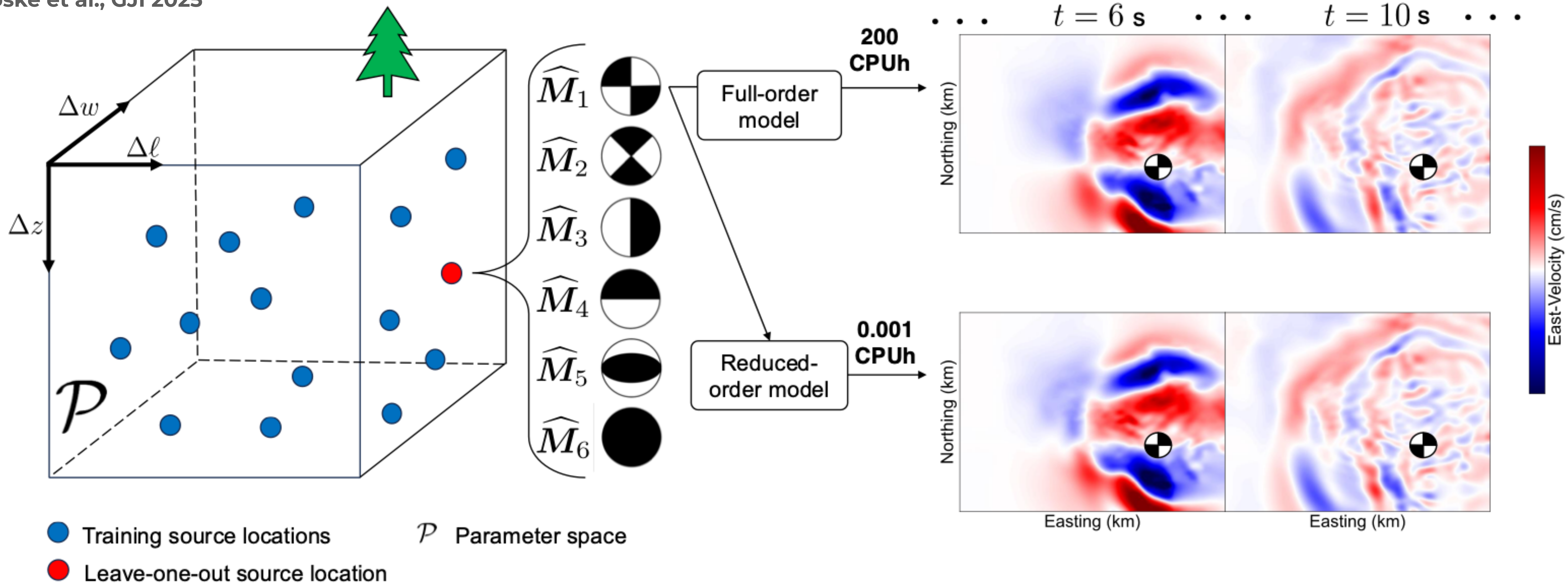
John Rekoske



Rekoske et al., 2025

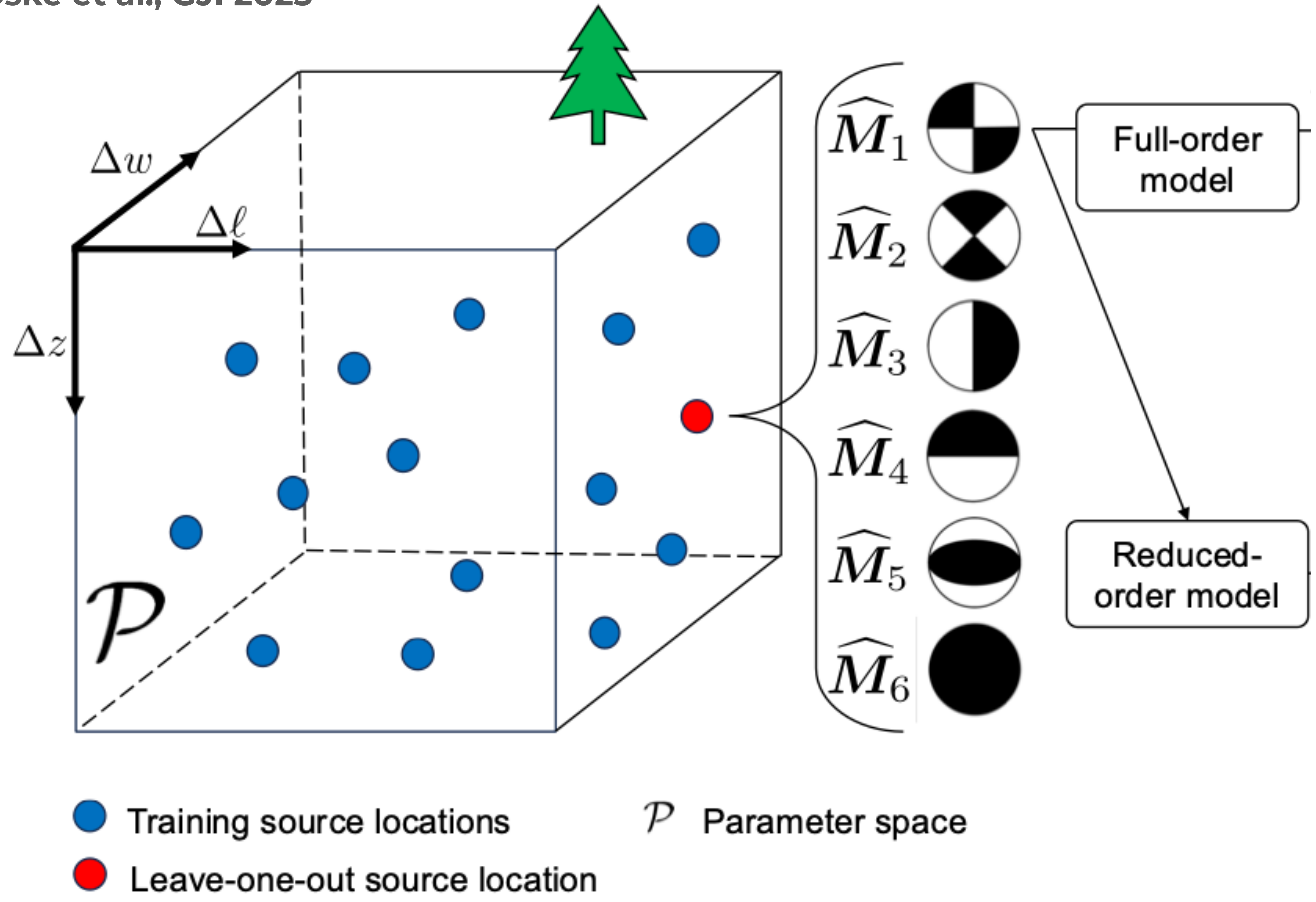
Reduced-order modeling for complex 3D seismic wave propagation

Rekoske et al., GJI 2025



Reduced-order modeling for complex 3D seismic wave propagation

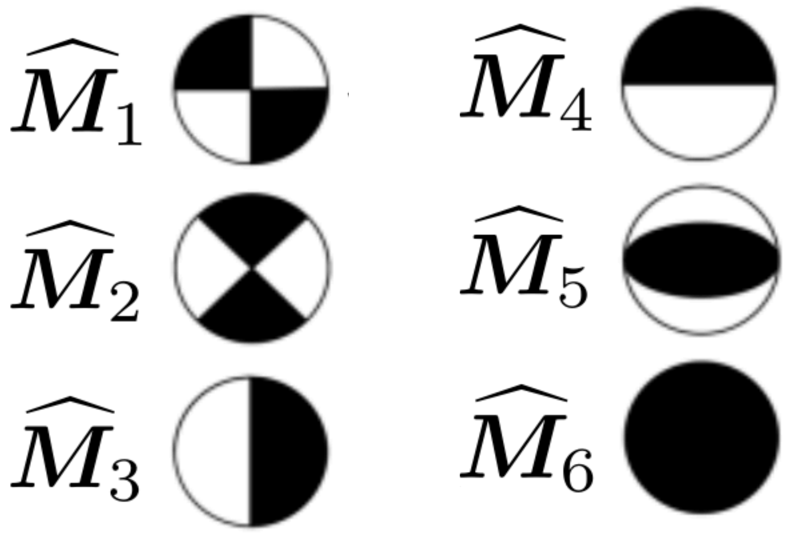
Rekoske et al., GJI 2025



$$\boldsymbol{D} = \underbrace{\begin{bmatrix} \boldsymbol{d}_{11} & \dots & \boldsymbol{d}_{1n_s} \\ \vdots & \ddots & \vdots \\ \boldsymbol{d}_{61} & \dots & \boldsymbol{d}_{6n_s} \end{bmatrix}}_{\mathbb{R}^{(6\cdot n_r \cdot n_t \cdot 3)\times n_s}}, \quad \boldsymbol{d}_{mp} = \underbrace{\begin{bmatrix} \boldsymbol{x}_{mp1} \\ \boldsymbol{x}_{mp2} \\ \vdots \\ \boldsymbol{x}_{mpn_r} \end{bmatrix}}_{\mathbb{R}^{n_r \times 1}}, \quad \boldsymbol{x}_{mpr} = \underbrace{\begin{bmatrix} u_{mpr1} \\ \vdots \\ u_{mprn_t} \\ v_{mpr1} \\ \vdots \\ v_{mprn_t} \\ w_{mpr1} \\ \vdots \\ w_{mprn_t} \end{bmatrix}}_{\mathbb{R}^{3n_t \times 1}}$$

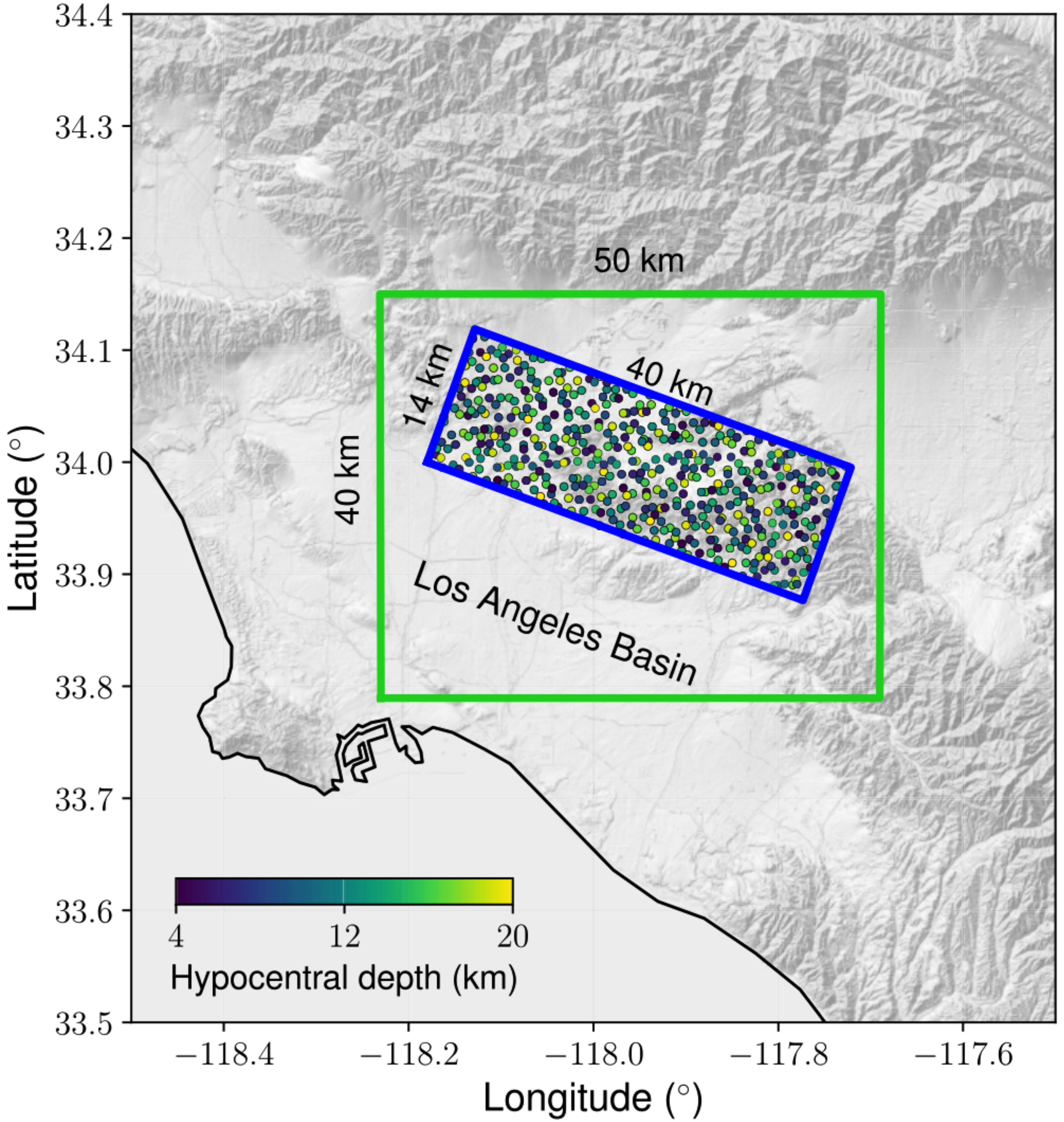
- **Snapshots** in time-dependent case consist of all **(flattened) space-time values**
- For elementary moment tensors, **6 Green's functions** computed

$$\bar{\boldsymbol{M}} = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{12} & M_{22} & M_{23} \\ M_{13} & M_{23} & M_{33} \end{bmatrix} = \sum_{i=1}^{6} c_i \widehat{\boldsymbol{M}}_i$$



Earthquake and station locations for Green's function calculations

- ROM constructed for this region enables **rapid computation** (0.0001 CPU hr) of complete, high-resolution (500 m spacing), 0.5 Hz surface velocity wavefields that are accurate for a shortest wavelength of 1.0 km for a single elementary moment tensor source (& account for geotechnical layer!)



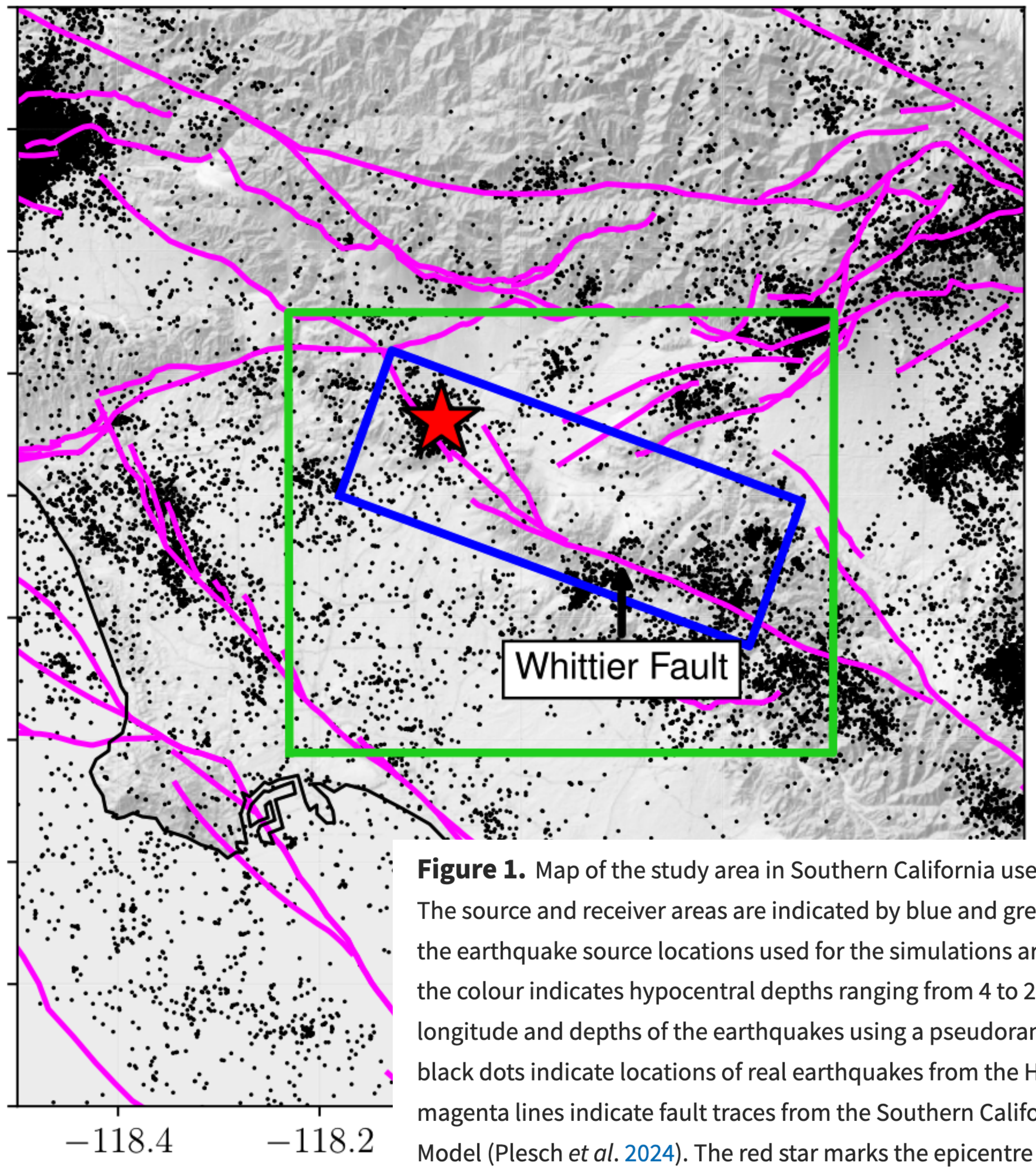
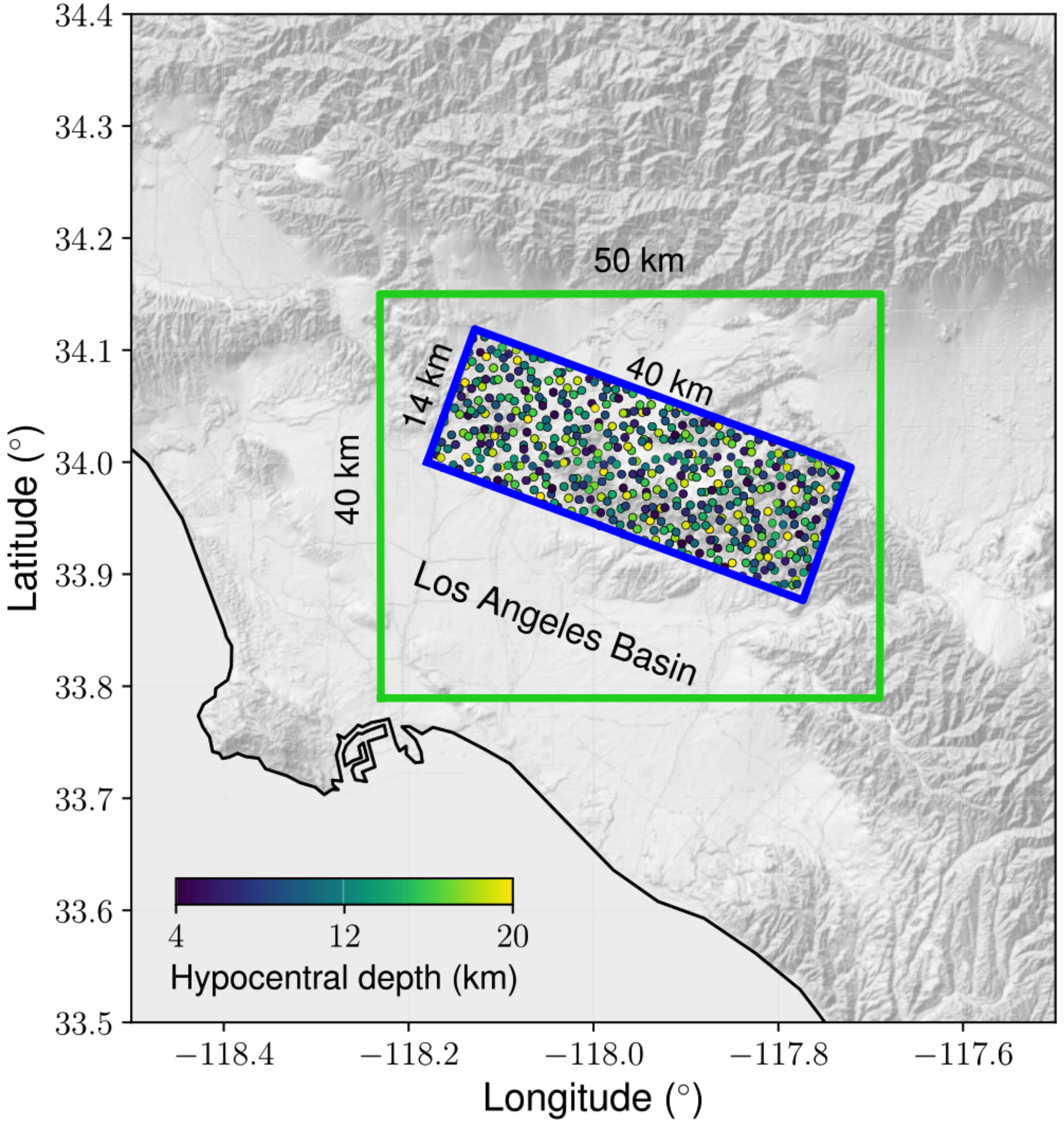


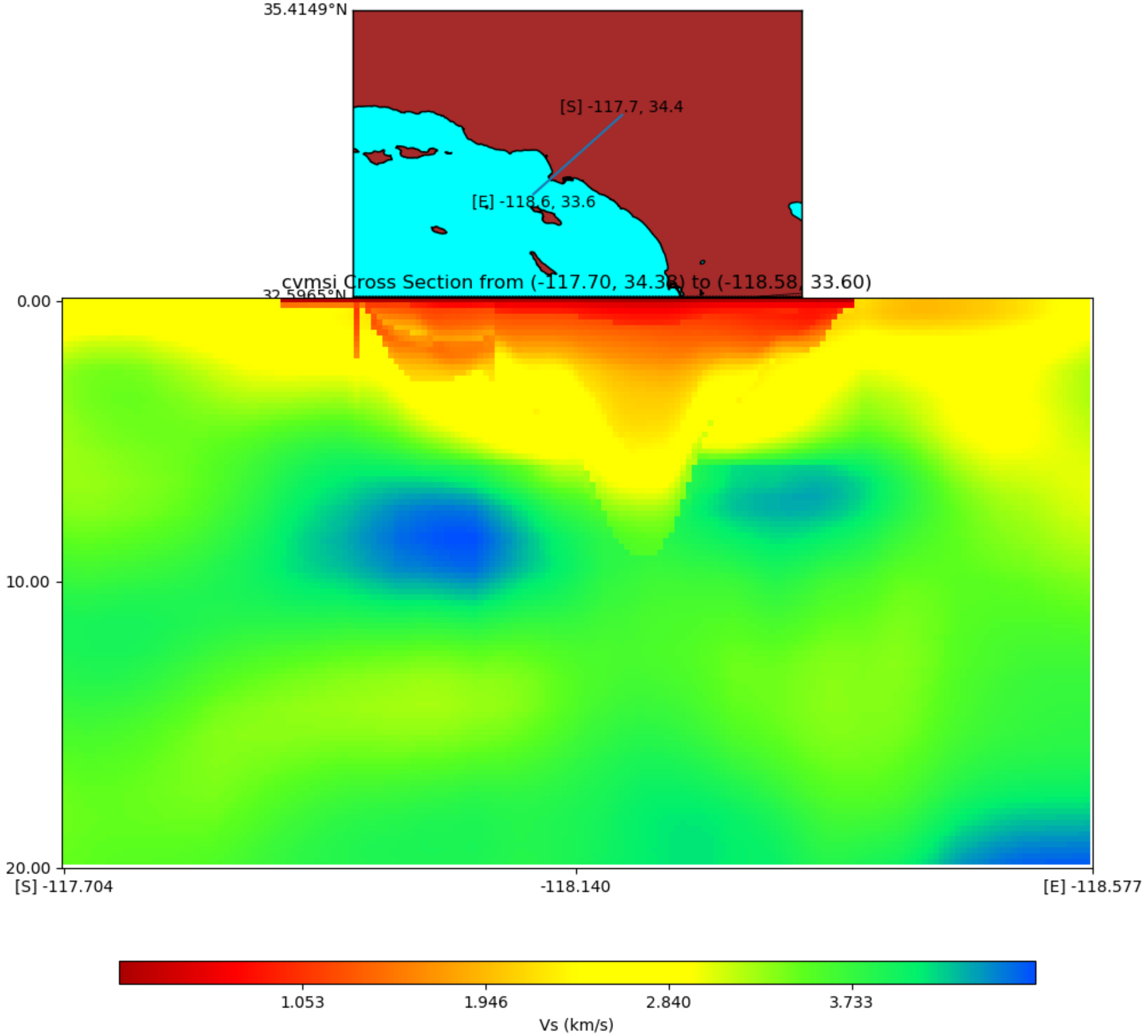
Figure 1. Map of the study area in Southern California used for computing rapid seismic wavefields. The source and receiver areas are indicated by blue and green rectangles, respectively. In panel (a), the earthquake source locations used for the simulations are indicated by the coloured circles, where the colour indicates hypocentral depths ranging from 4 to 20 km. We determine the latitude, longitude and depths of the earthquakes using a pseudorandom Halton sequence. In panel (b), the black dots indicate locations of real earthquakes from the Hauksson *et al*. (2012) catalogue, and the magenta lines indicate fault traces from the Southern California Earthquake Center Community Fault Model (Plesch *et al*. 2024). The red star marks the epicentre of the 1987 M_W 5.9 Whittier Narrows earthquake, which we use as a demonstrator for our finite fault rupture modelling approach.

Earthquake and station locations for Green's function calculations

- ROM constructed for this region enables **rapid computation** (0.0001 CPU hr) of complete, high-resolution (500 m spacing), 0.5 Hz surface velocity wavefields that are accurate for a shortest wavelength of 1.0 km for a single elementary moment tensor source (& account for geotechnical layer!)

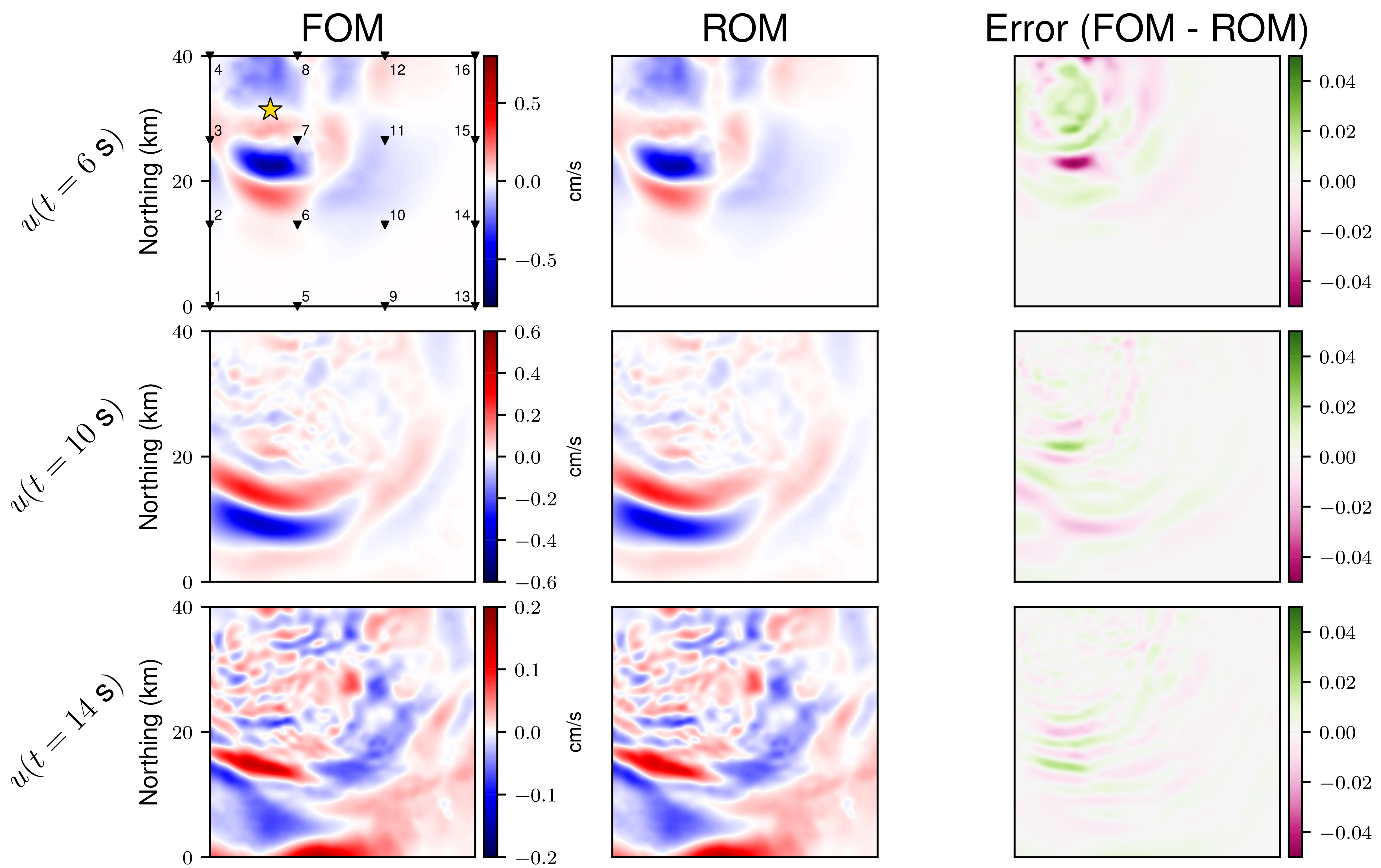


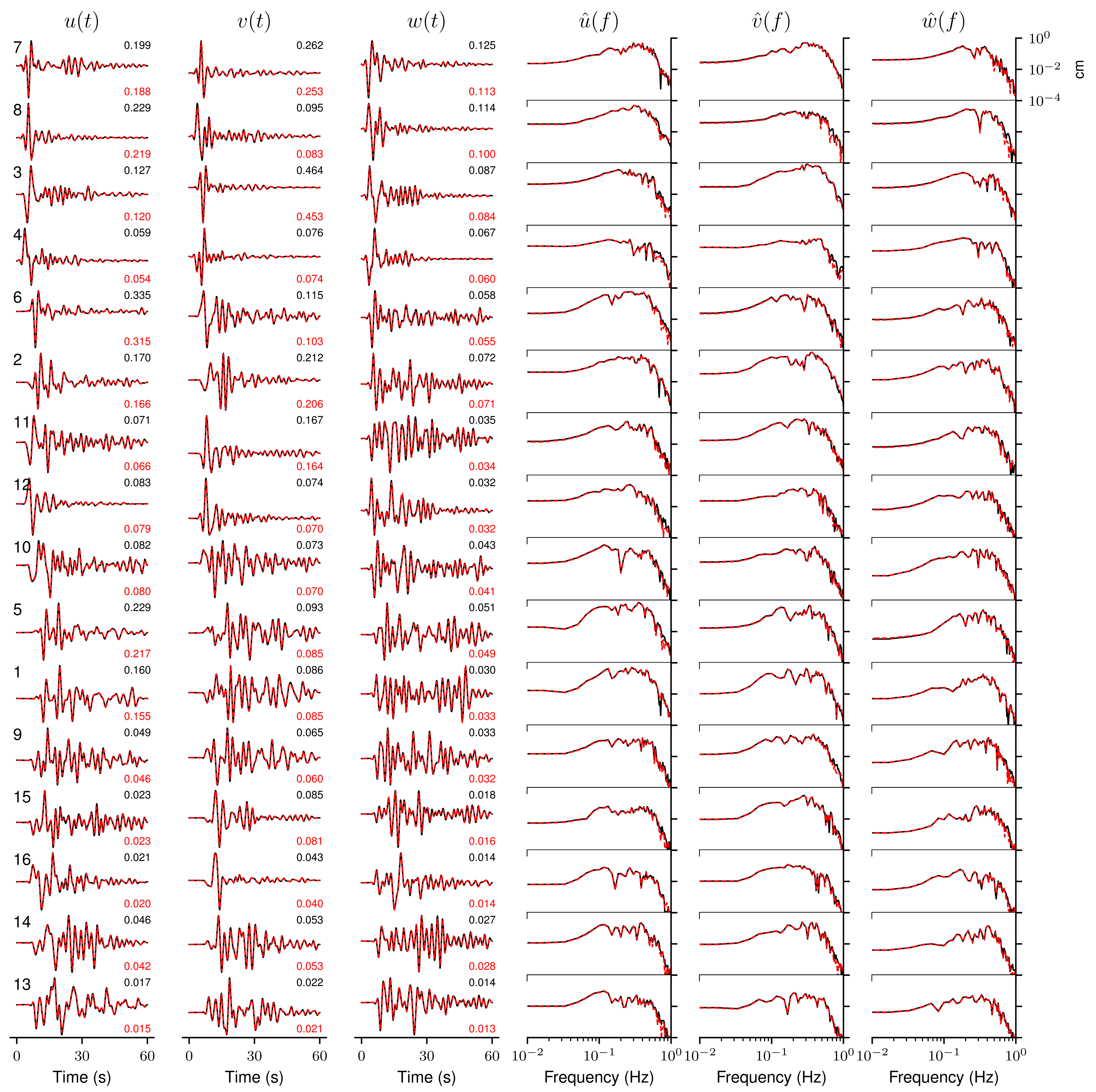
CVM-S4.26.M01 cross-section



Wave Propagation

Approximation Errors

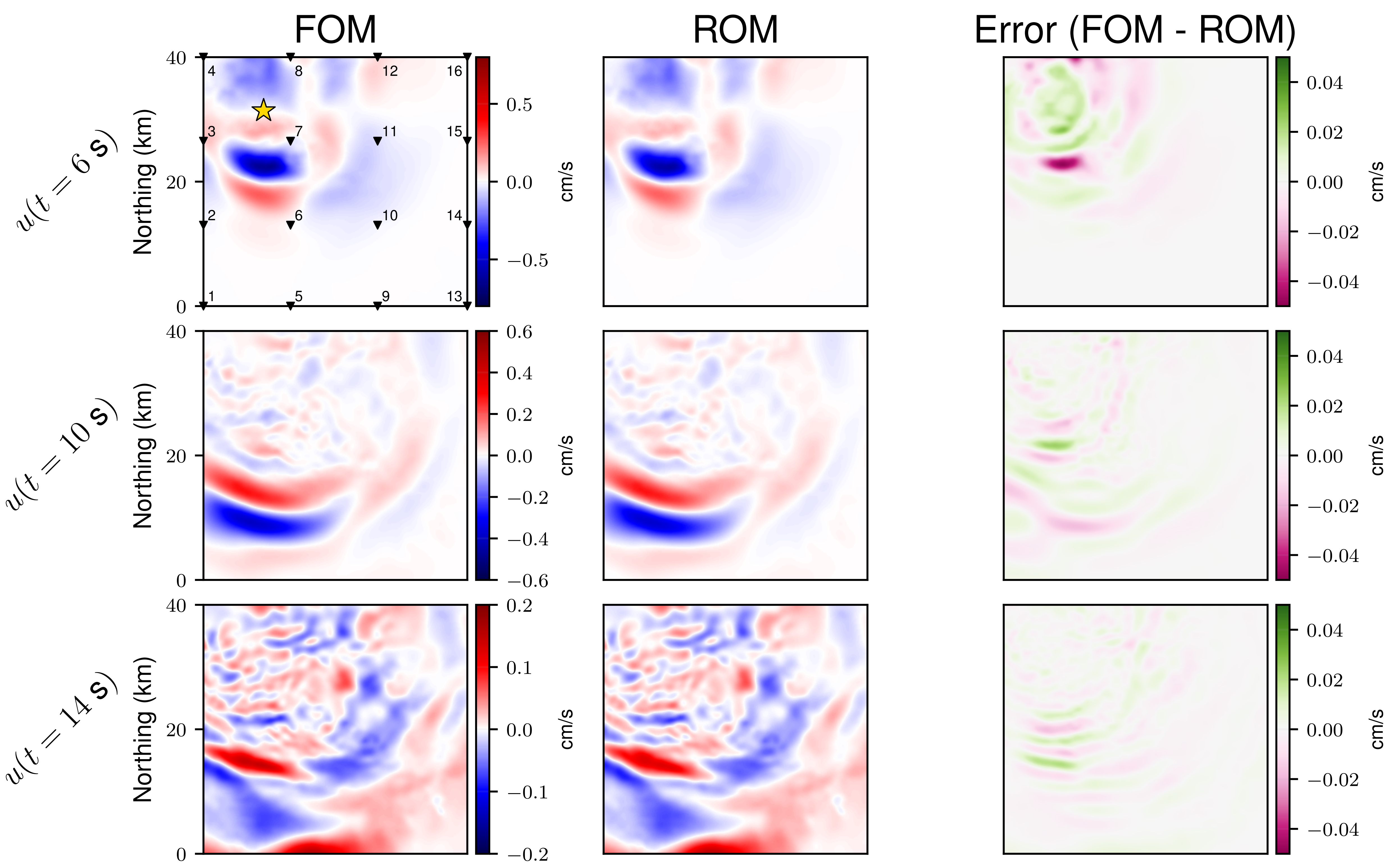


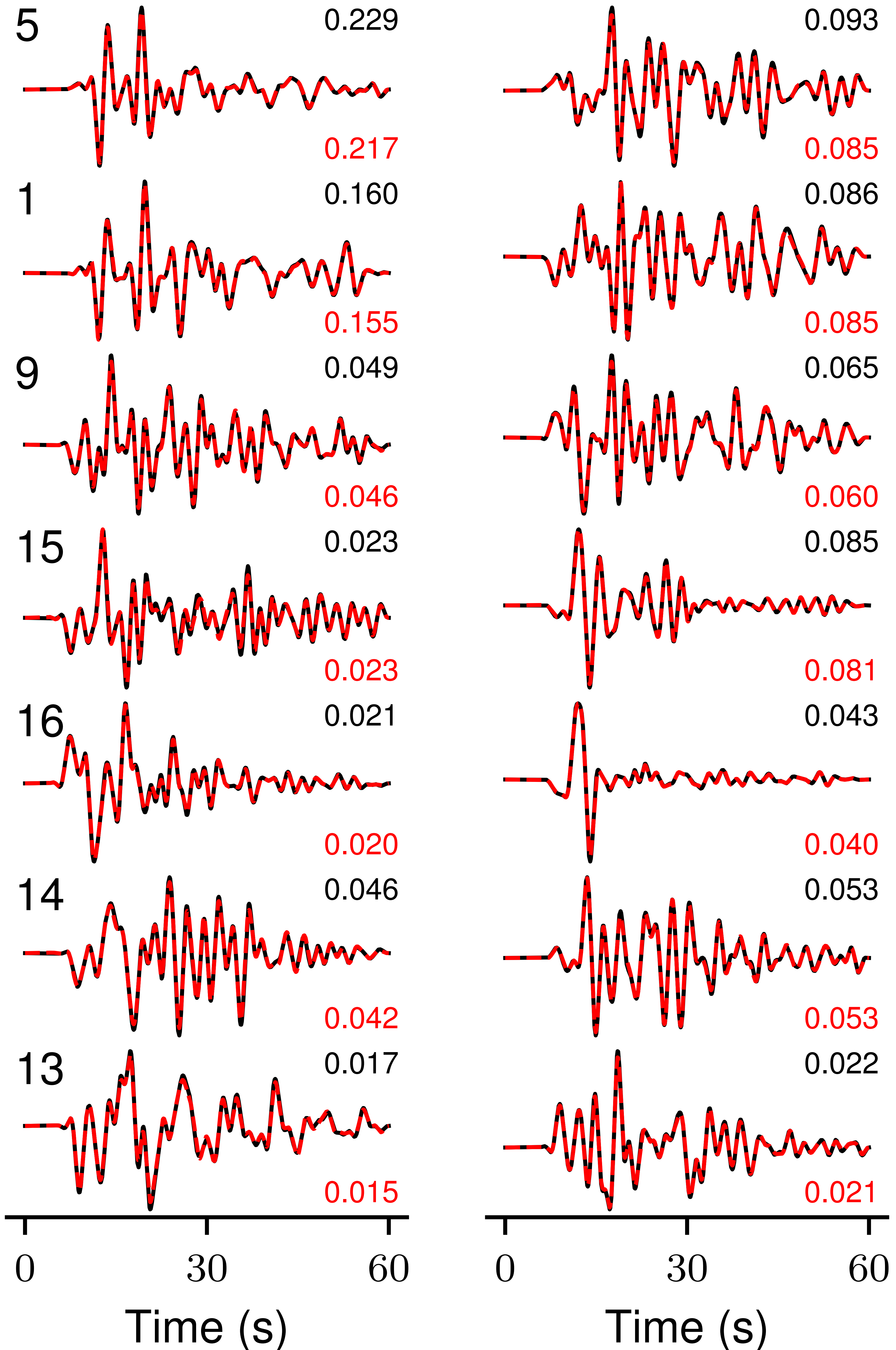


- *Errors are small in space, time and frequency domain*

Wave Propagation

Approximation Errors



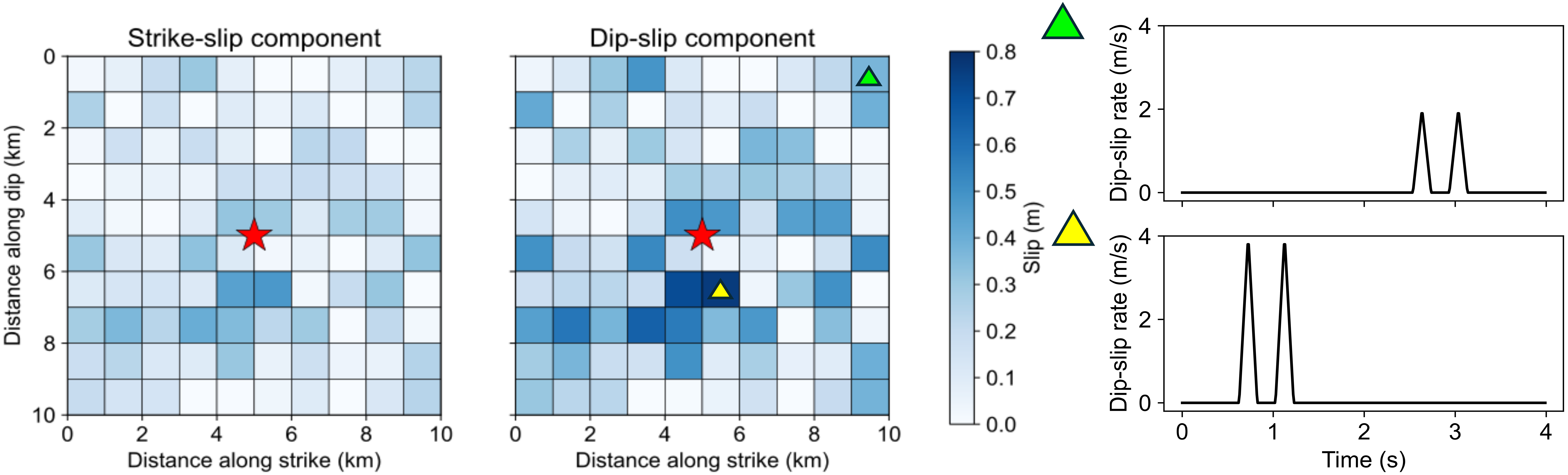


- *Errors are small in space, time and frequency domain*

Our approach can also simulate seismograms for kinematic, finite fault models

Discrete representation theorem:

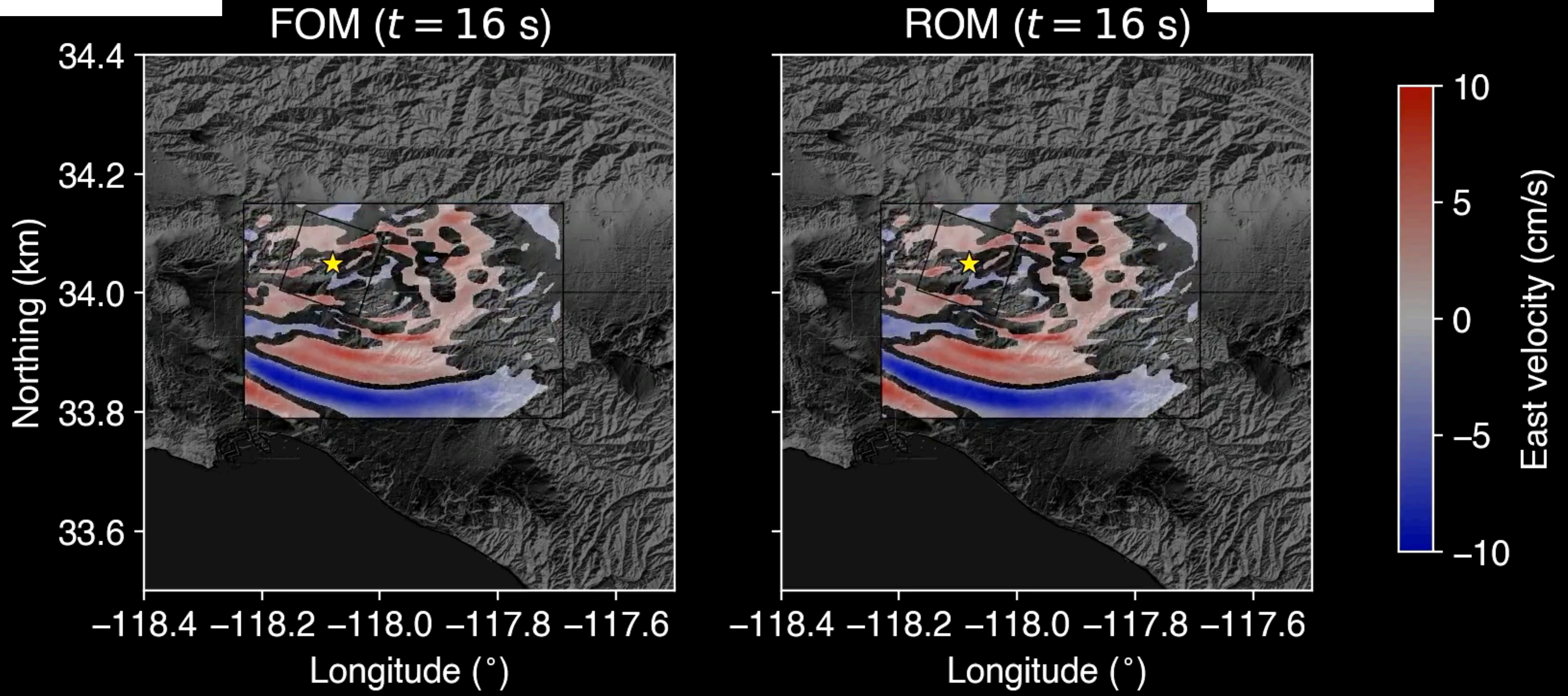
$$u_n(\boldsymbol{x}_r, t) = \sum_{i=1}^{n} M_{pq}(\boldsymbol{x}_i, t) * G_{np,q}(\boldsymbol{x}_i, \boldsymbol{x}_r, t)$$



Hartzell & Iida (1990)

Reduced-order modeling for complex 3D seismic wave propagation

Rekoske et al., GJI 2025



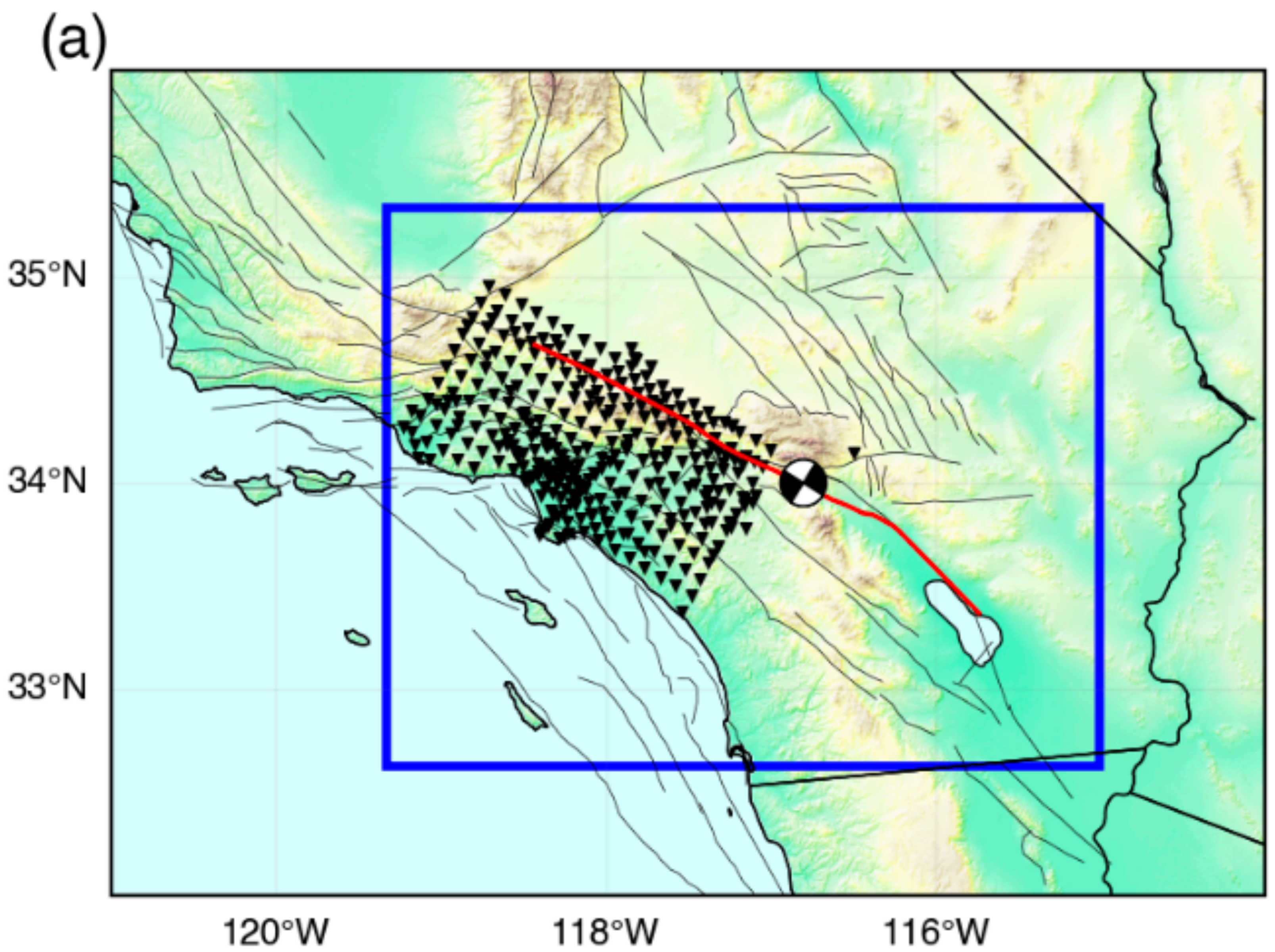
Increasing resolution in Cybershake physics-based PSHA via ROMs

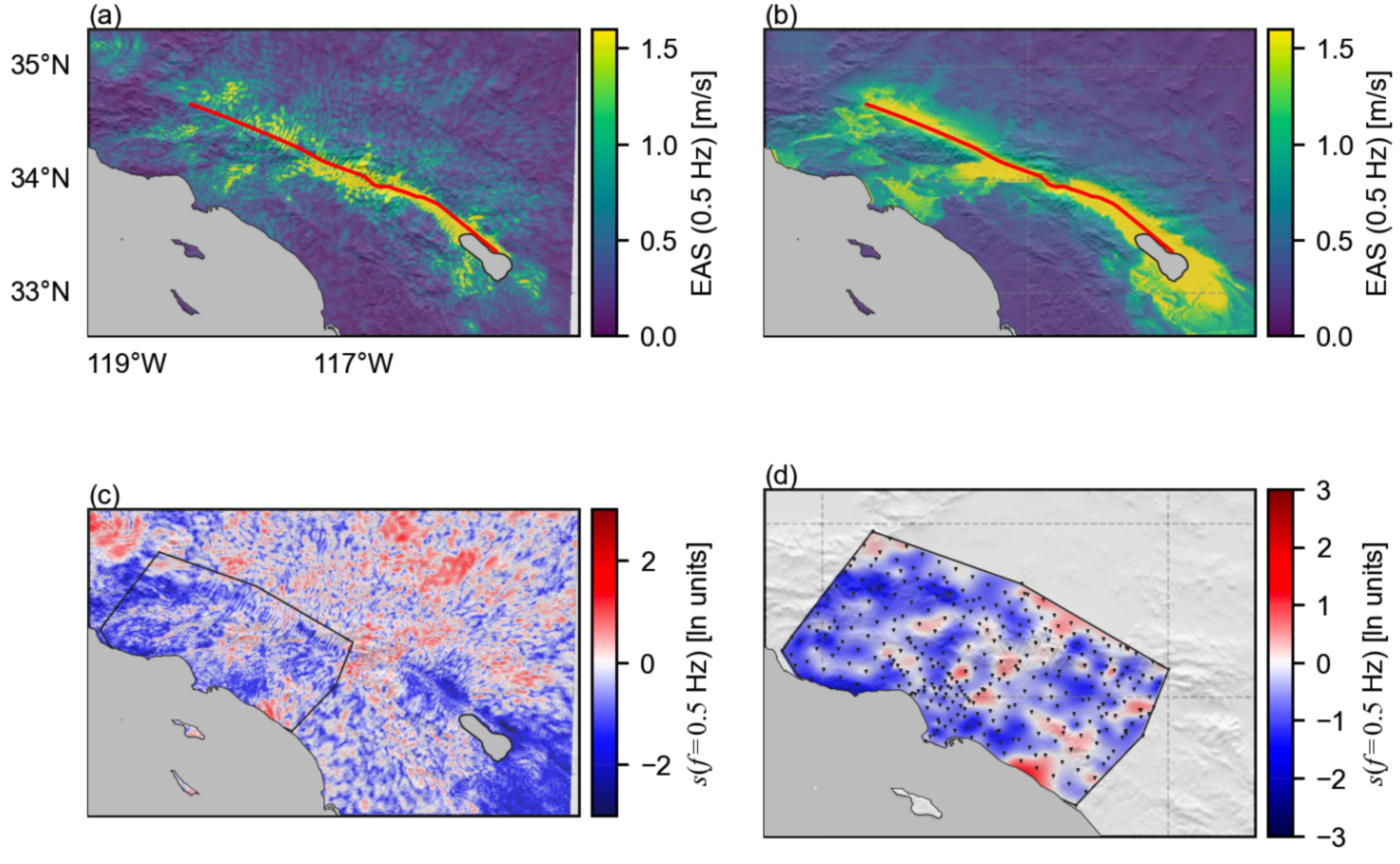
submitted to BSSA

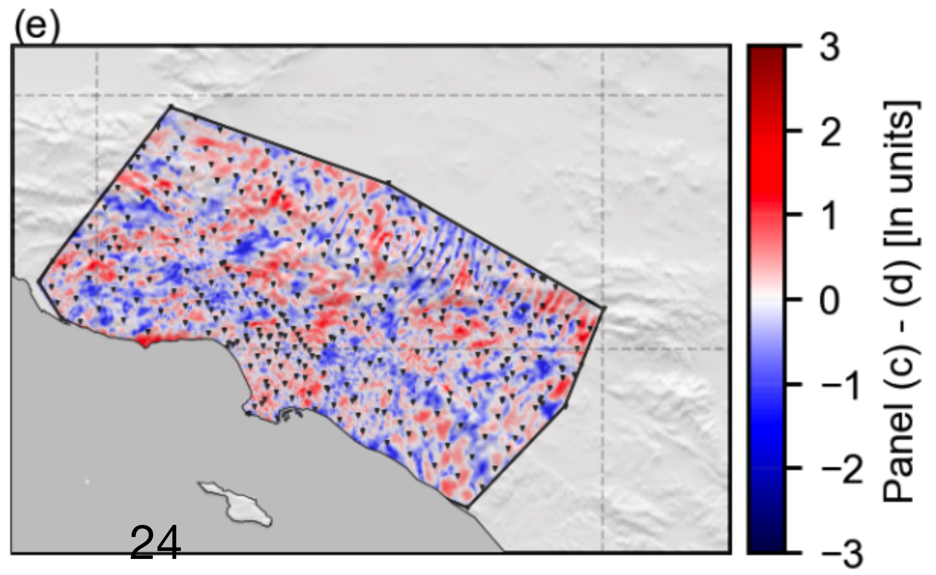
- **CyberShake** hazard maps interpolate ground motion from sparse grids (~**335** sites)
- **SeisSol** simulations for **480,000 sites** (500m spacing) along the Southern San Andreas Fault

CyberShake physics-based seismic hazard maps using reduced-order models: results from Southern San Andreas earthquake scenarios

John M. Rekoske[1,*], Scott Callaghan[2], Kevin Milner[3], Dave A. May[1], and Alice-Agnes Gabriel[1,4]







Alt-text: Five-panel map showing 0.5 Hz ground motion results for a kinematic rupture model. Layout identical to point source figure with panels for simulated EAS, GMM prediction, residuals, interpolated residuals, and interpolation errors across Southern California.

Increasing resolution in Cybershake physics-based PSHA via ROMs

- Using the BA18 GMM as a baseline: significant non-ergodic effects (e.g., radiation patterns); GMMs generally underpredict simulated spectra in **deep basins** like Fillmore
- Interpolation misses localized amplification near basin edges, causing errors **up to a factor of ~3**
- **Errors increase with frequency** (0.1–1.0 Hz) as wavelengths shorten relative to station spacing
- Our ROM model enables high-resolution mapping with a **336x speedup** compared to scaling the standard reciprocity approach

CyberShake physics-based seismic hazard maps using reduced-order models: results from Southern San Andreas earthquake scenarios

John M. Rekoske[1,*], Scott Callaghan[2], Kevin Milner[3], Dave A. May[1], and Alice-Agnes Gabriel[1,4]

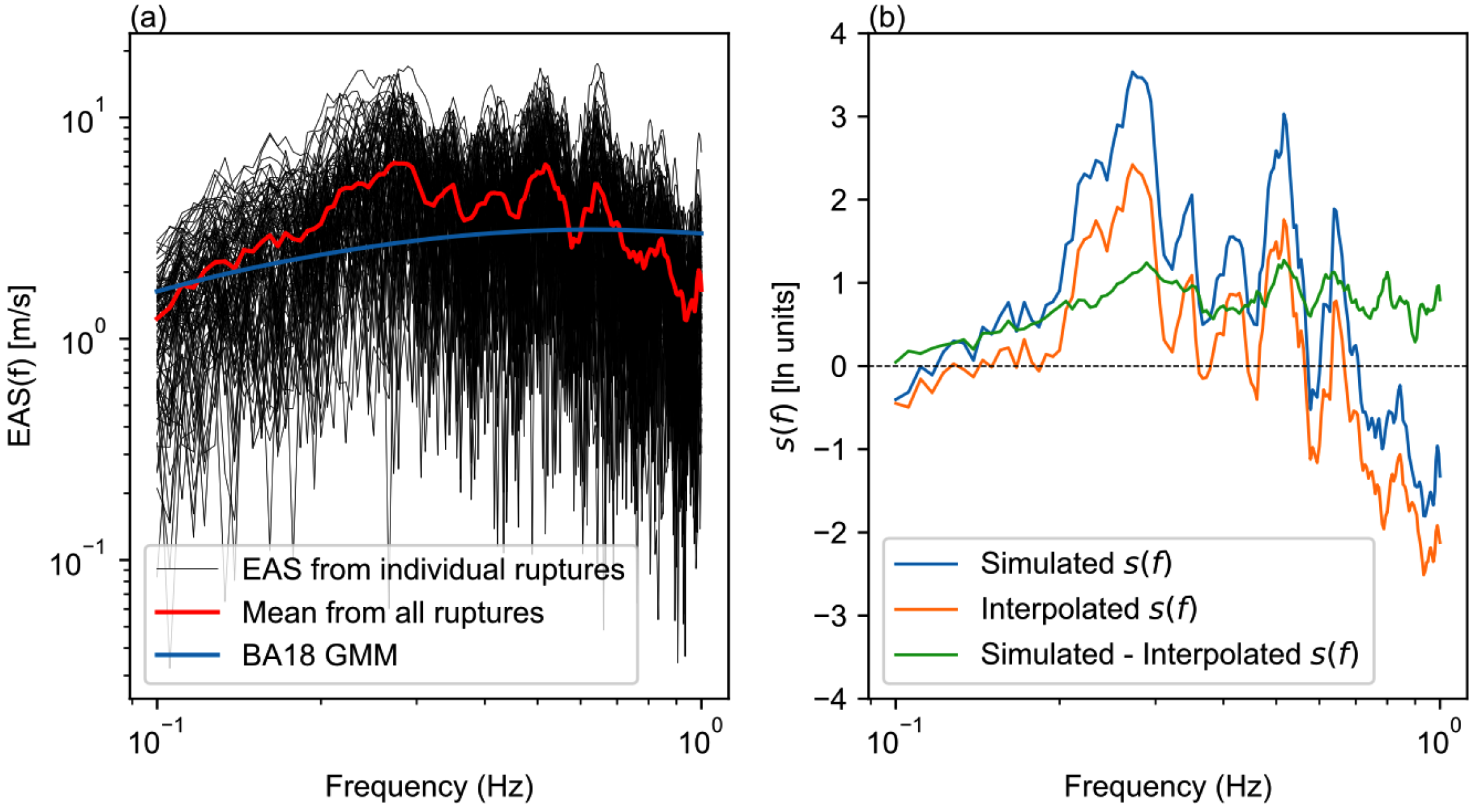
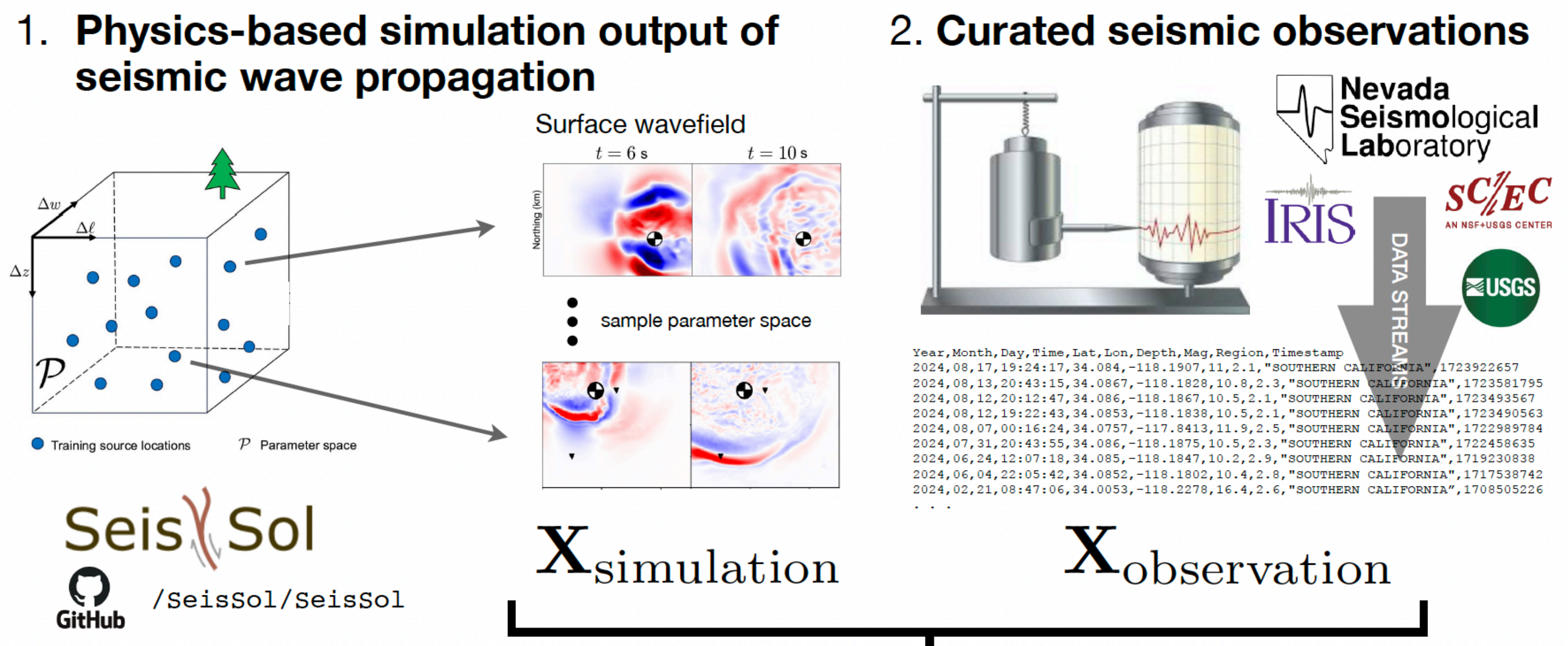


Figure 10: EAS for one location in the Fillmore Basin indicated by the star in Figure 2b. The simulated minus interpolated $s(f)$ is just over one natural log unit for some frequencies.

Next: "Space-time completeness of seismic ground motions via non-intrusive model order reduction"

- NSF Collaborations in Artificial Intelligence and Geoscience (CAIG)
  - D May, A Gabriel, D Trugman, B Kramer
- **What we will do**
  - Build time-dependent parametric surrogate models of seismic ground motions **fusing simulated ground motion wavefields with curated observational data from regional earthquakes in Southern California and Nevada**
  - Examine the fundamental differences in source and path affecting seismic ground motions from earthquakes in different geologic settings and with distinct mechanisms, depths, and source properties
  - Deploy models for seismic hazard assessment and provide physics-based earthquake early-warning and rapid response
- **Methods**
  - **Gappy POD (data fusion)**
  - ROM —> Operator Inference (OpInf)
  - Open-source HPC toolkit for large scale POD (SVD) and OpInf



3. Data Fusion

Fuse **simulation** and **observation** data matrices into a self consistent **basis** describing space, time, parameter variations

Gappy POD [68—71]

Left singular vectors $\mathbf{\Phi}_v$

4. Operator Inference [65 — 67]

I. Exploit **projection** to define a **structure-preserving low-dimensional model** $\quad \widehat{\boldsymbol{v}} = \mathbf{\Phi}_v^\top \boldsymbol{v}$ — Low-dimensional snapshot from **Gappy POD** left singular vectors

II. Define the **structure** of the reduced-order model $\quad \dot{\widehat{\boldsymbol{v}}} = \widehat{\boldsymbol{A}}\,\widehat{\boldsymbol{v}} + \widehat{\boldsymbol{H}}\,(\widehat{\boldsymbol{v}} \otimes \widehat{\boldsymbol{v}}) + \widehat{\boldsymbol{B}}\,\boldsymbol{f}$

Non-intrusive learning through **inferring** the low-dimensional operators from **observation and simulation** data

$$\widehat{\boldsymbol{V}} = (\widehat{\boldsymbol{v}}(t_1), \widehat{\boldsymbol{v}}(t_2), \ldots, \widehat{\boldsymbol{v}}(t_K))$$

$$\min_{\widehat{\boldsymbol{A}},\widehat{\boldsymbol{H}},\widehat{\boldsymbol{B}}} \left\| \dot{\widehat{\boldsymbol{V}}}^\top - \widehat{\boldsymbol{V}}^\top \widehat{\boldsymbol{A}}^\top - (\widehat{\boldsymbol{V}} \odot \widehat{\boldsymbol{V}})^\top \widehat{\boldsymbol{H}}^\top - \boldsymbol{F}^\top \widehat{\boldsymbol{B}}^\top \right\|_F^2 + \mathcal{R}$$

Low-dimensional operators define the reduced model as a dynamical system

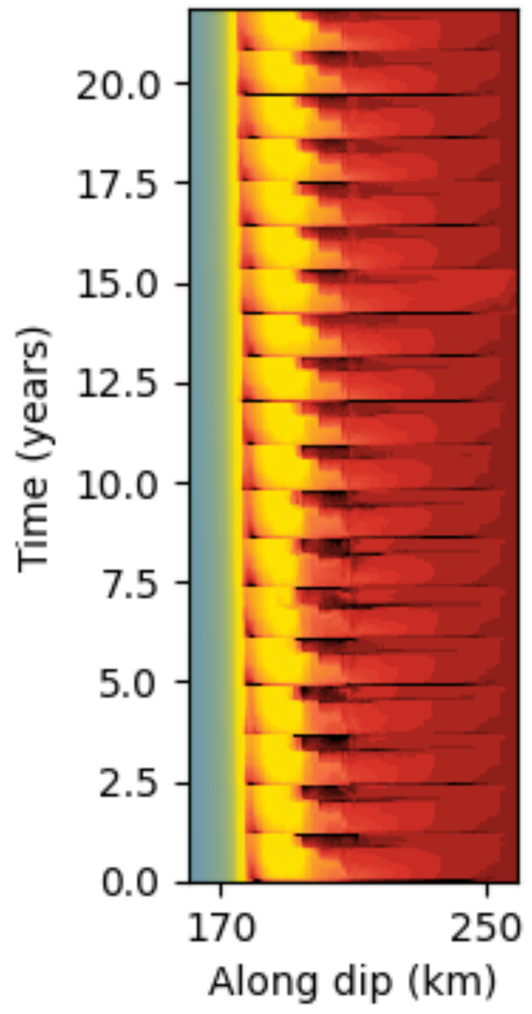
Minimum residual formulation leads to a regularized least squares problem

Regularization terms

Summary

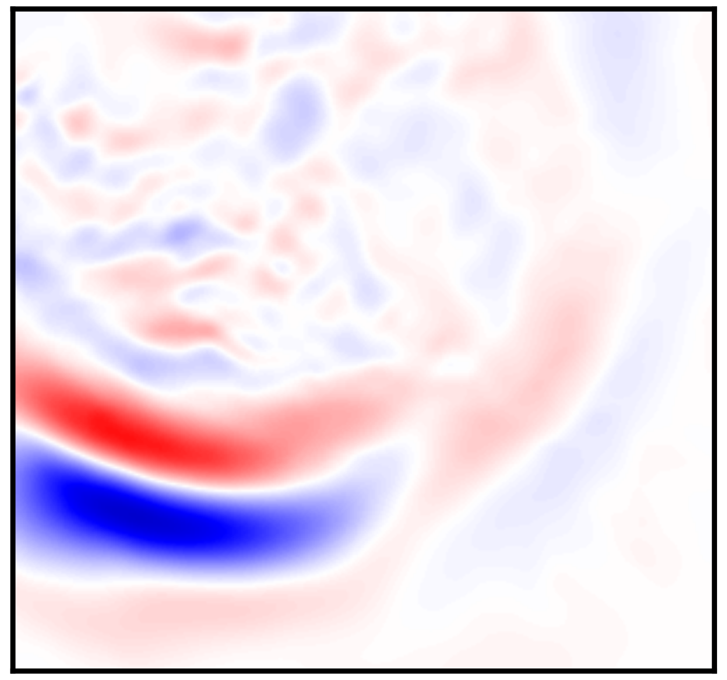
- **Non-intrusive interpolated POD reduced-order models are:**
  - **Easy to use;**
  - **As accurate as required, without needing vast volumes of training data;**
  - **Easy to interpret;**
  - **Remarkably accurate**
  - **Verifiable (no black box, we have error estimators);**
  - **Remarkably fast to evaluate**
  - **Applicable to shake maps, seismic wave propagation, dynamic rupture, SSE cycles, …**
- **We still need high-quality HPC simulations, but we can "do more" with fewer FOMs by exploiting ROMs**

Speedup summary



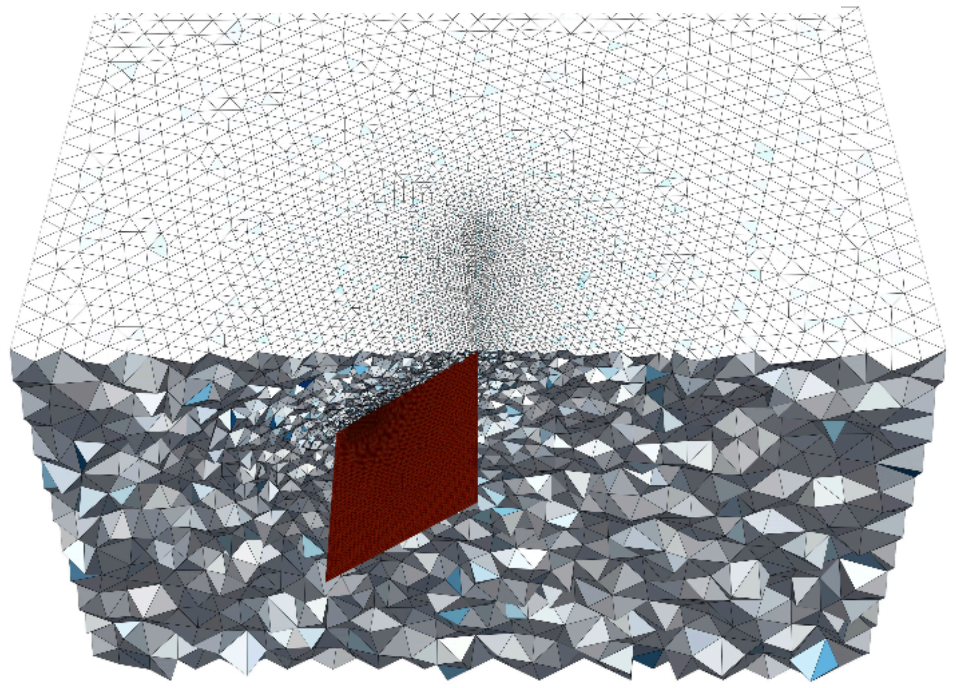
Seismic Cycles

10^5 x



Wave Propagation

10^5 x



Dynamic Rupture

10^8 x

References:

Rekoske, Gabriel, May (2023), "Instantaneous physics-based ground motion maps using reduced-order modeling", J. Geophys. Res., 128, e2023JB026975, doi:10.1029/2023JB026975

Rekoske, May, Gabriel (2025). "Reduced-order modeling for complex 3D seismic wave propagation", Geophysical Journal International, 241, 526–548, doi:10.1093/gji/ggaf049

Hobson, May, Gabriel (2025), "Quantifying the influence of fault geometry via mesh morphing with applications to earthquake dynamic rupture and thermal models of subduction", accepted in G3, arXiv preprint doi:arXiv:2506.15892

Magen, May & Gabriel (2025), "Reduced-order modelling of Cascadia's slow slip cycles", in press GJI, EarthArXiv preprint doi:10.31223/X5QT7V