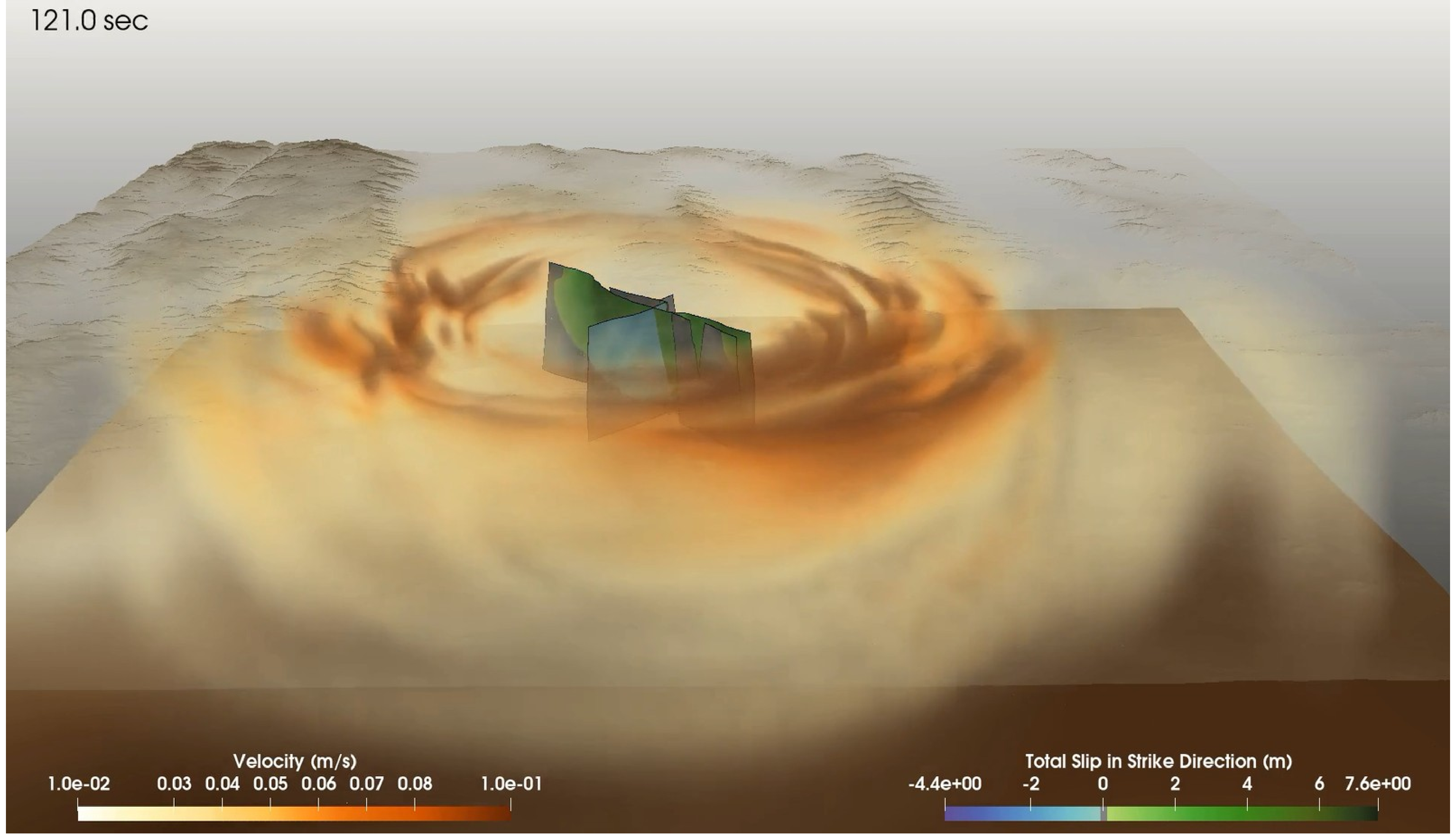
SCRIPPS INSTITUTION OF OCEANOGRAPHY
UC San Diego

Alice-Agnes Gabriel

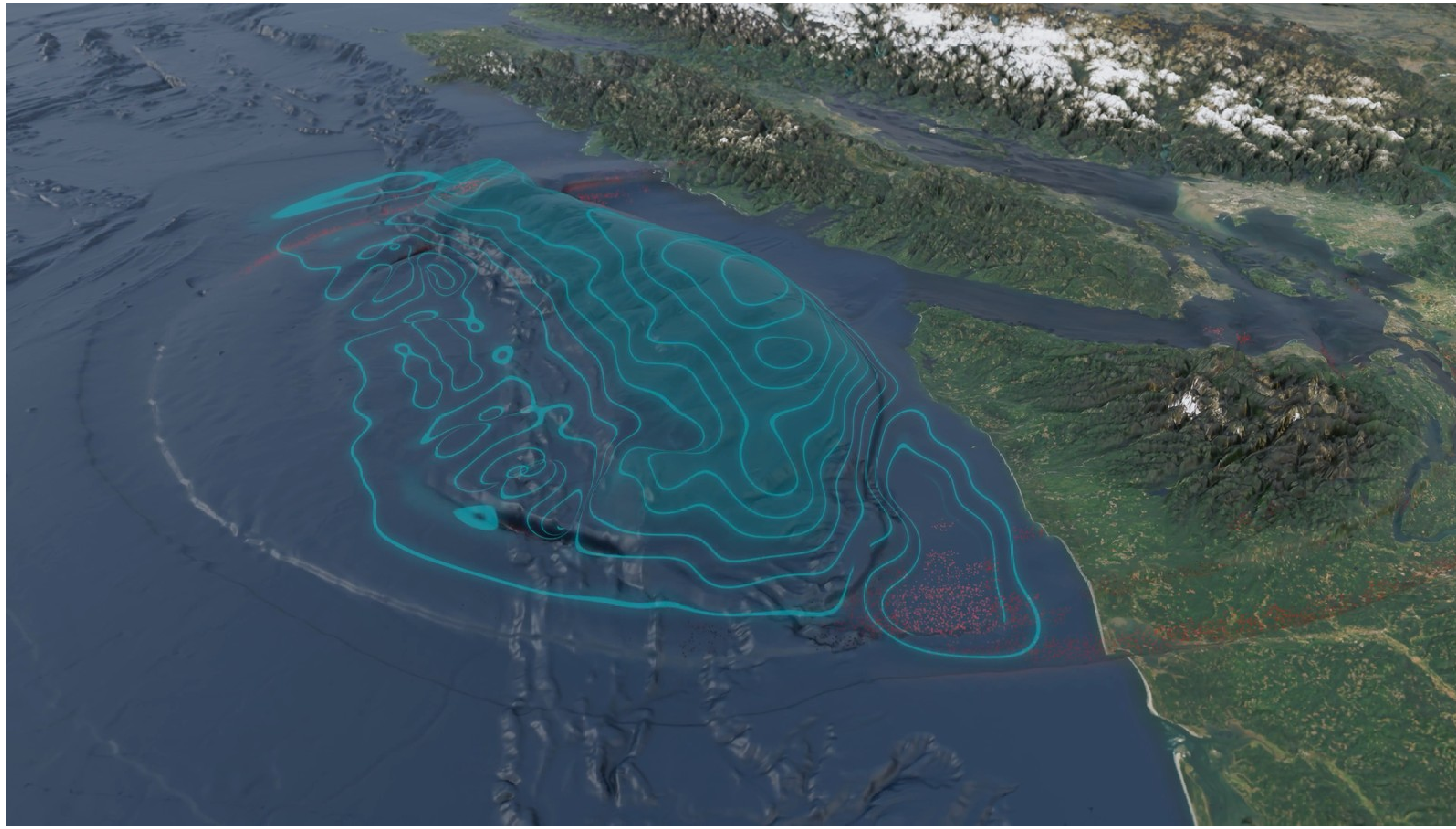
(and many outstanding students, postdocs, & collaborators, as referenced throughout this talk)

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Earthquakes, Tsunamis, and Supercomputers:
From Hero Runs to Digital Twins



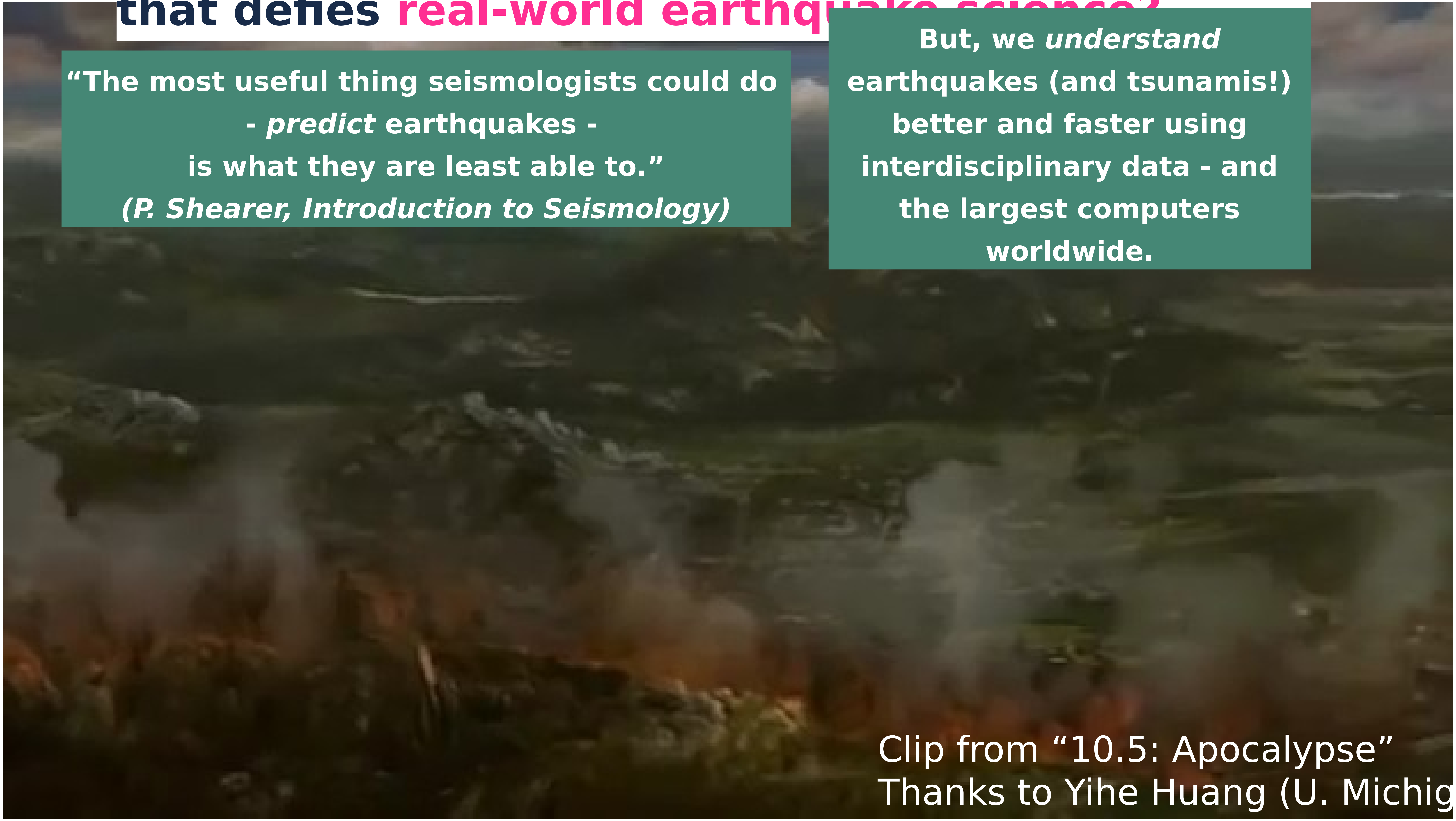
Propagation of seismic waves and "unzipping" of faults during dynamic rupture simulations of the 2019 Ridgecrest earthquakes. Visualization of 15 TB of 3D volumetric data on unstructured tetrahedral meshes on Frontera.
Taufiqurrahman et al., Nature, 2023
Abrams et al., SC23, Visualization showcase



High-resolution acoustic–gravity simulation for real-time Bayesian inference at extreme scale, a digital twin for tsunami early warning applied to the Cascadia subduction zone
Henneking et al., SC25, Art in HPC

IXPUG Workshop, ISC, June 26, 2026

Do you notice something about this movie scene that defies real-world earthquake science?



“The most useful thing seismologists could do - *predict* earthquakes - is what they are least able to.”
(P. Shearer, Introduction to Seismology)

But, we *understand* earthquakes (and tsunamis!) better and faster using interdisciplinary data - and the largest computers worldwide.

Clip from “10.5: Apocalypse”
Thanks to Yihe Huang (U. Michig

Earthquake observations are increasingly interdisciplinary

- And even include CCTV cameras!

e.g., Kearse & Kaneko, TSR. 2

Earthquakes may cause destructive tsunamis

- **Since 1900, 58 tsunamis have claimed more than 260,000 lives,** averaging ~4,600 deaths per event, more than any other natural hazard
- 2004 Sumatra / Indian Ocean (M9.1): **~230,000 deaths across 14 countries**
- 2011 Tōhoku-Oki, Japan (M9.1): **~18,000 deaths, ~$240B in damage**

This iconic photograph shows the massive 2011 Tohoku tsunami wave breaching a seawall and engulfing a street in Miyako City, Iwate Prefecture, Japan. (Mainichi Shimbun / Reuters file)

Kamchatka, 29 July 2025 (M8.8 earthquake)

- Tied for **sixth-largest earthquake instrumentally recorded**, no casualties from shaking or tsunami waves

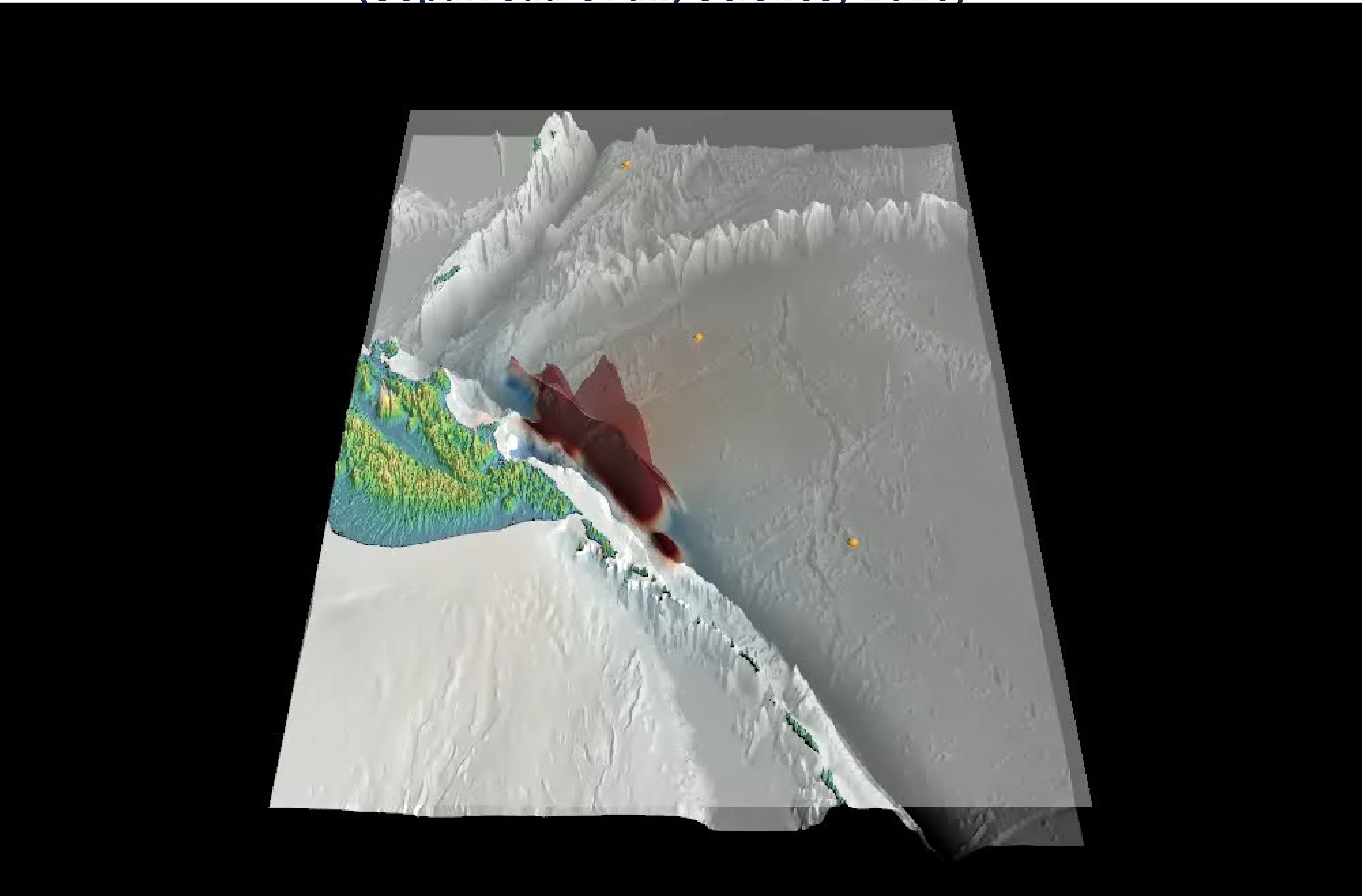
- Warnings: 1.9 million evacuated in Japan, 2700 in the Sakhalin (Russia) region

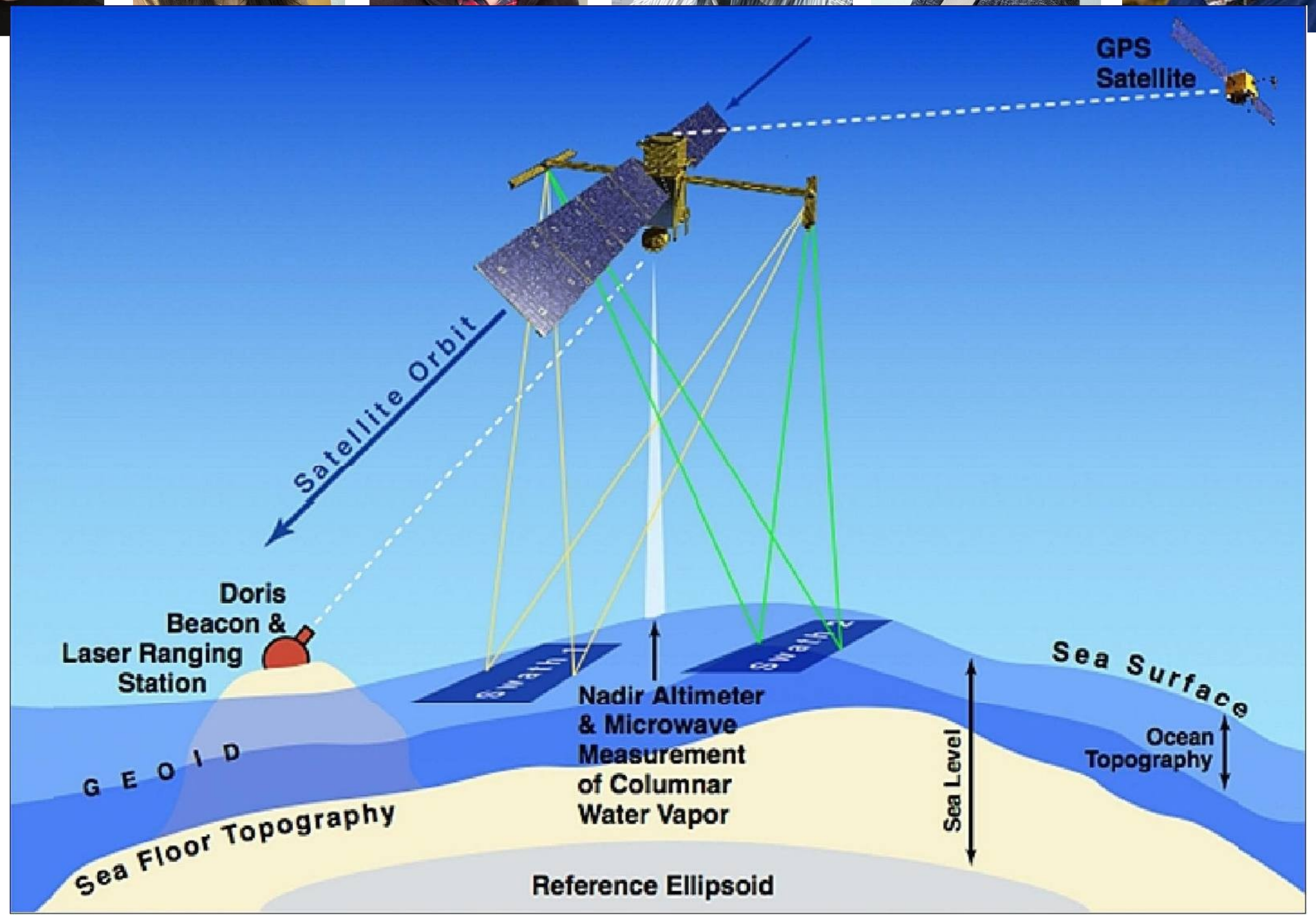
- Tsunami weaker than the earthquake magnitude implied (~1 m at most far-field gauges), but ~5-6 m local waves at Severo-Kurilsk, with run-ups reaching ~33 m in a remote bay

"A man walking his dog captures footage as the tsunami strike Kamchatka's coast" (verified)
https://www.youtube.com/watch?v=ZH0CKMKrQ4Q

A lucky overflight

SWOT Satellite Spots Tsunami Wave After Kamchatka Quake
(Sepulveda et al., *Science*, 2026)



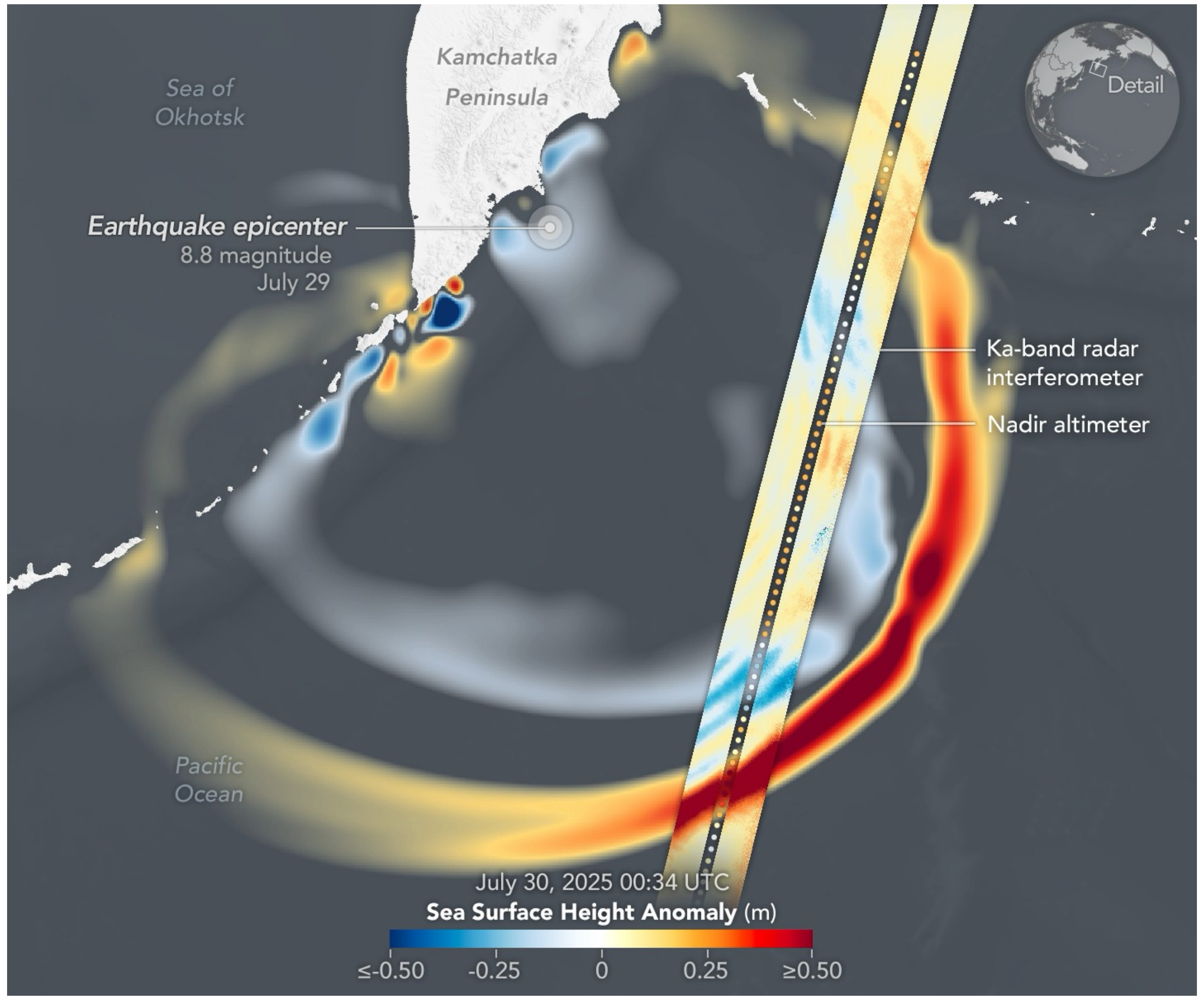


SWOT - Surface Water and Ocean Topography

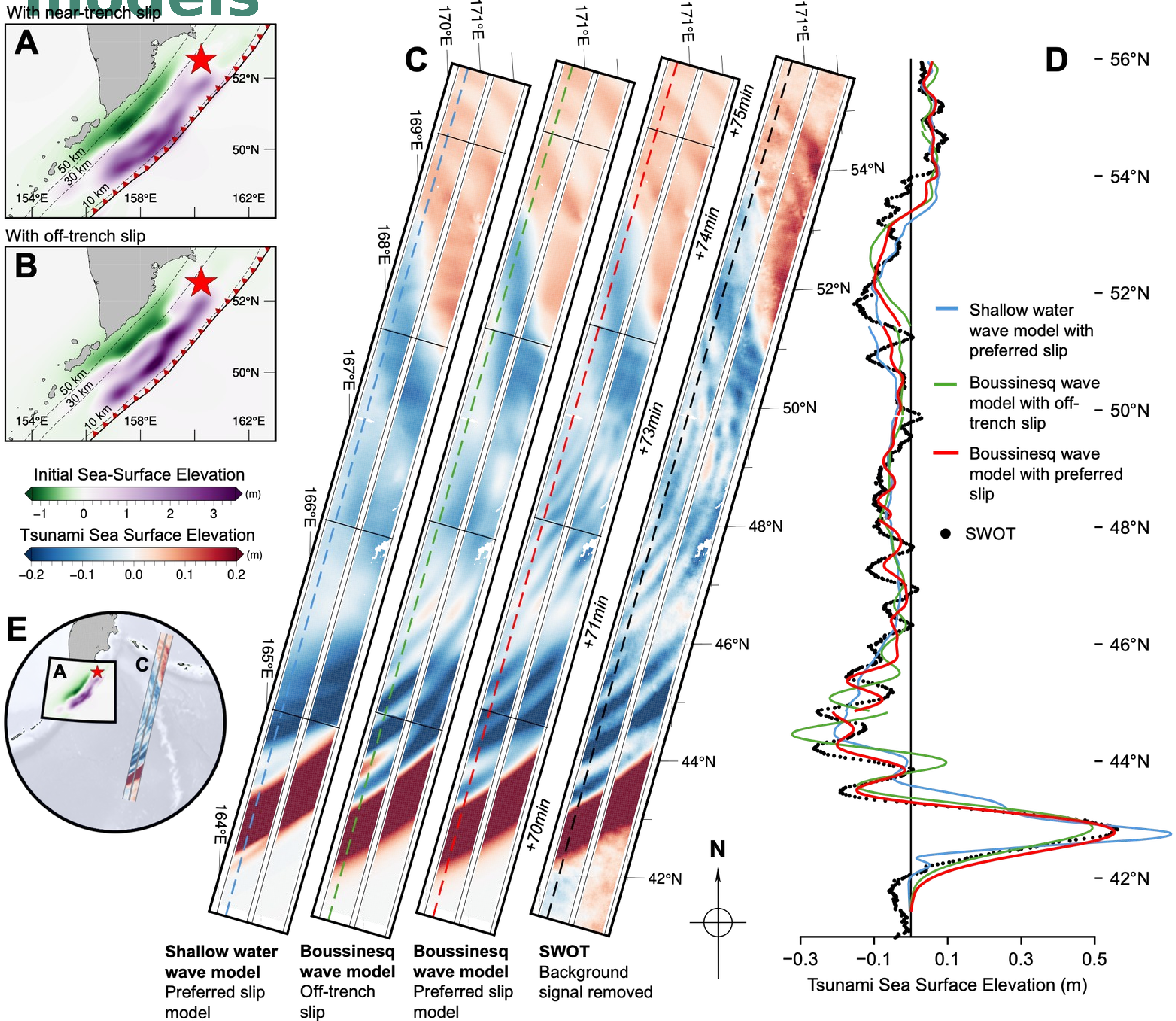
- Joint NASA / CNES mission (with CSA & UKSA); launched Dec 2022
- Carries KaRIn, a Ka-band radar interferometer with two antennas on a 10-m boom
- Images sea surface height across a 120-km-wide swath, with ~2 km posting and ~3 cm accuracy over the ocean
- **First altimeter to capture 2D images of the sea surface**, conventional altimeters (Jason, etc.) only sample directly beneath the satellite
- 21-day repeat orbit at ~890 km altitude

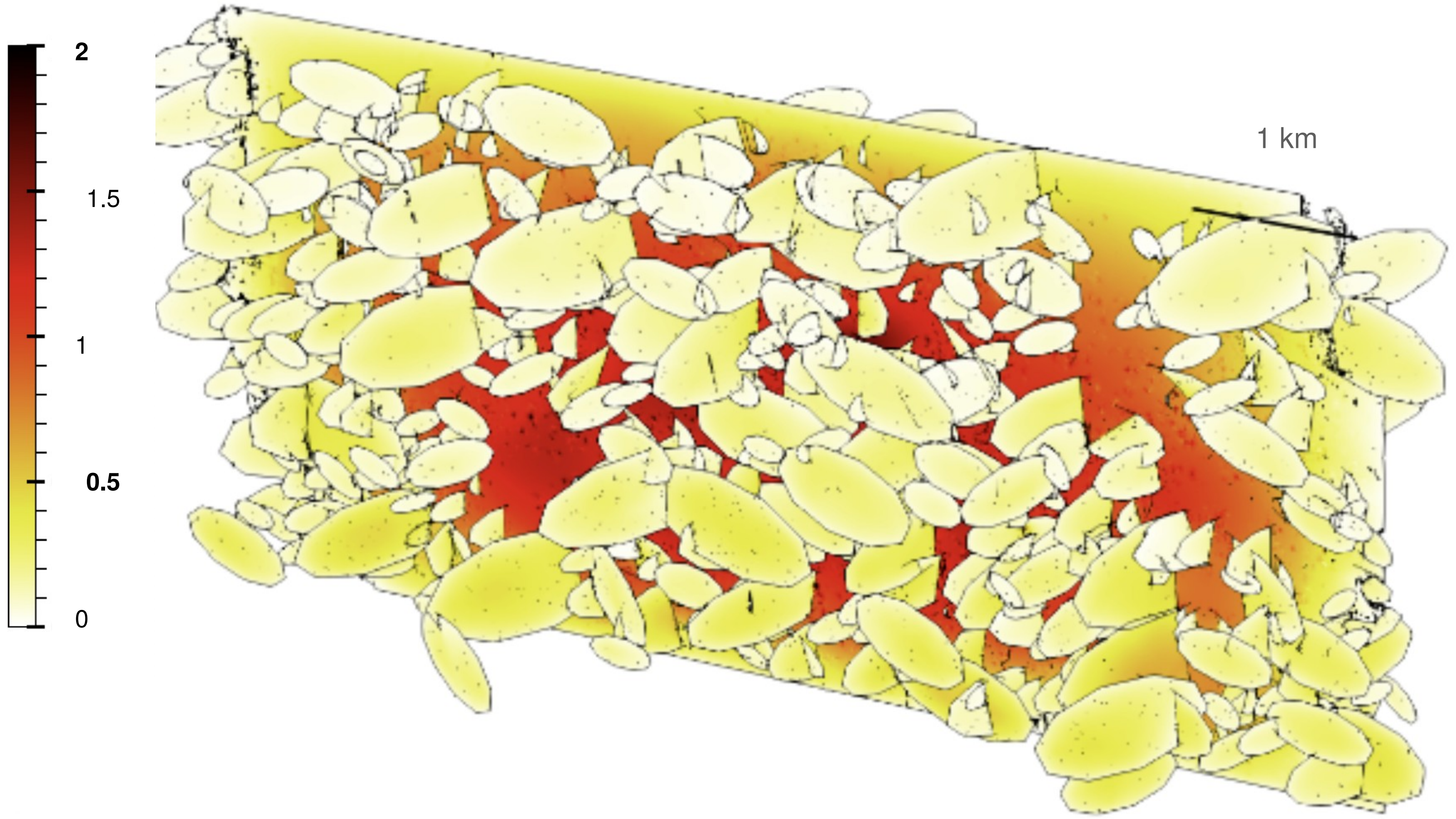
Integrating interdisciplinary data requires physics-based, high-resolution models

- **SWOT reveals <u>dispersive</u> tsunami waves - linked to near-trench tsunami-genesis!** *(I. Sepulveda, B. Nilsson, Y. Yao, M. Carvajal, M. Brandin, A.-A. Gabriel, D. Sandwell, Science, 2026)*

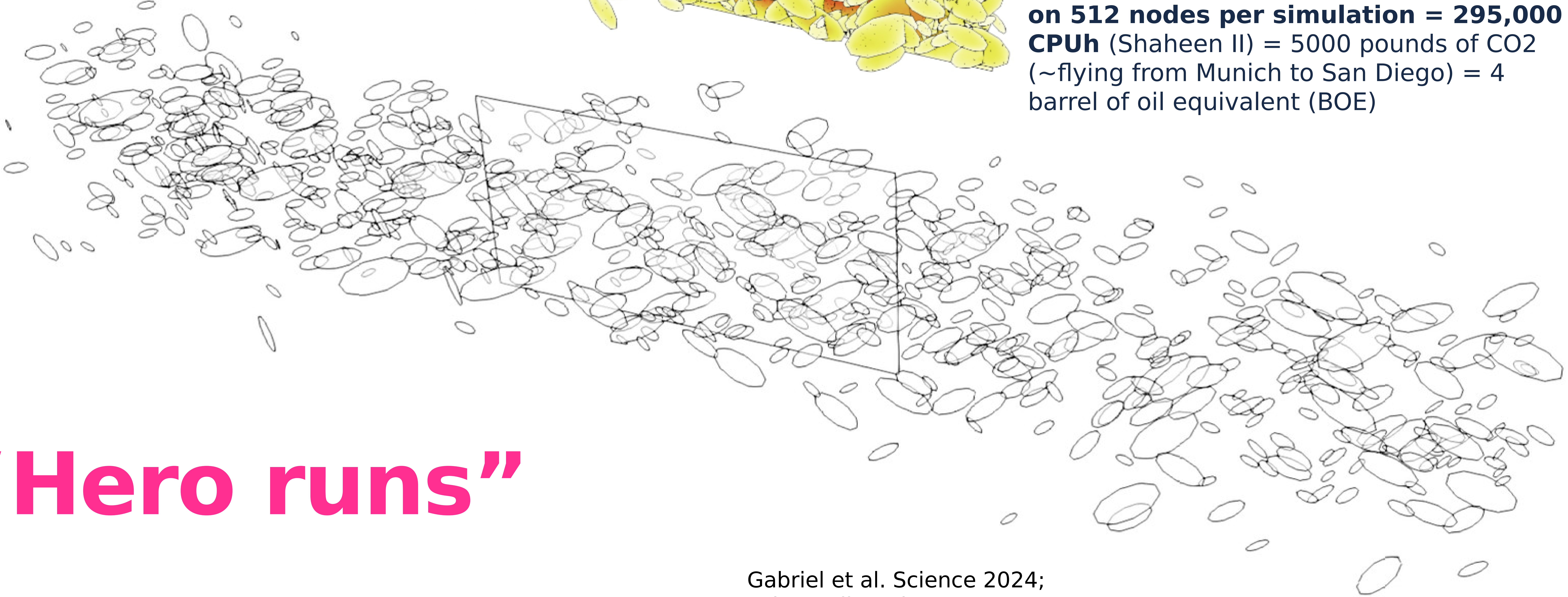


US-French SWOT Satellite Spots Tsunami Wave after 2025 Kamchatka Earthquake (NASA/JPL)



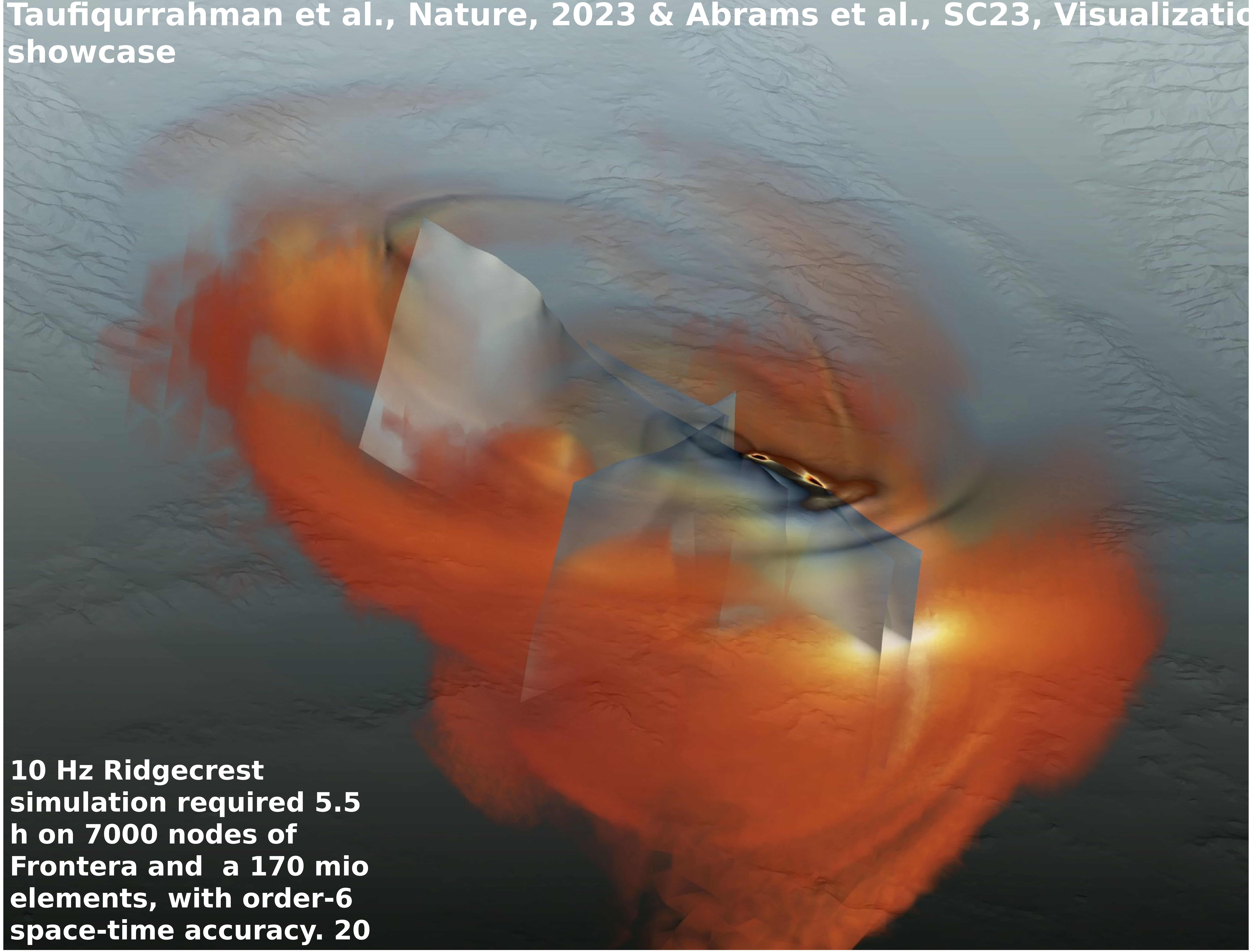


Resolving multi-scale rupture dynamics and wave propagation up to 13 Hz required **18h on 512 nodes per simulation = 295,000 CPUh** (Shaheen II) = 5000 pounds of CO_2 (~flying from Munich to San Diego) = 4 barrel of oil equivalent (BOE)

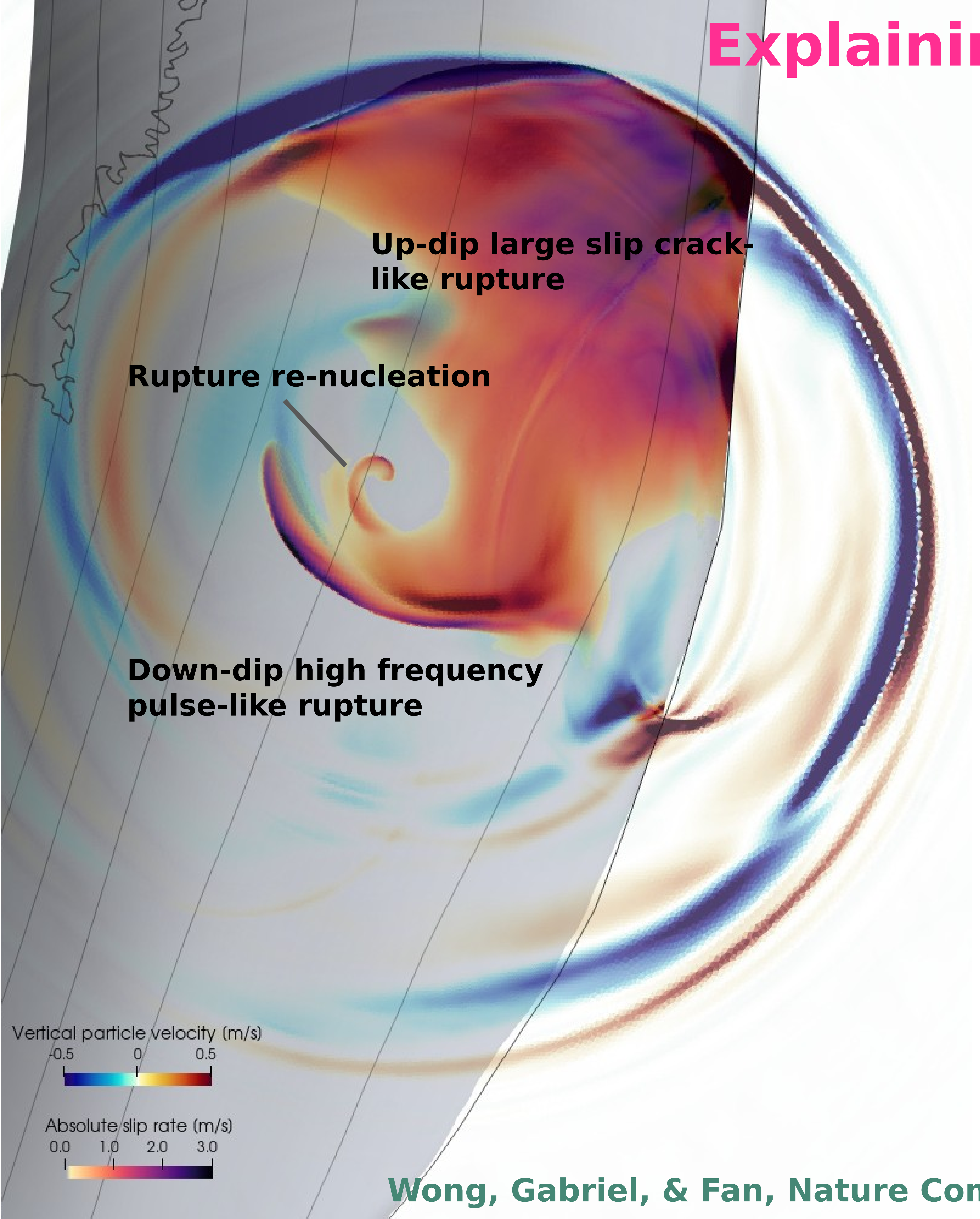


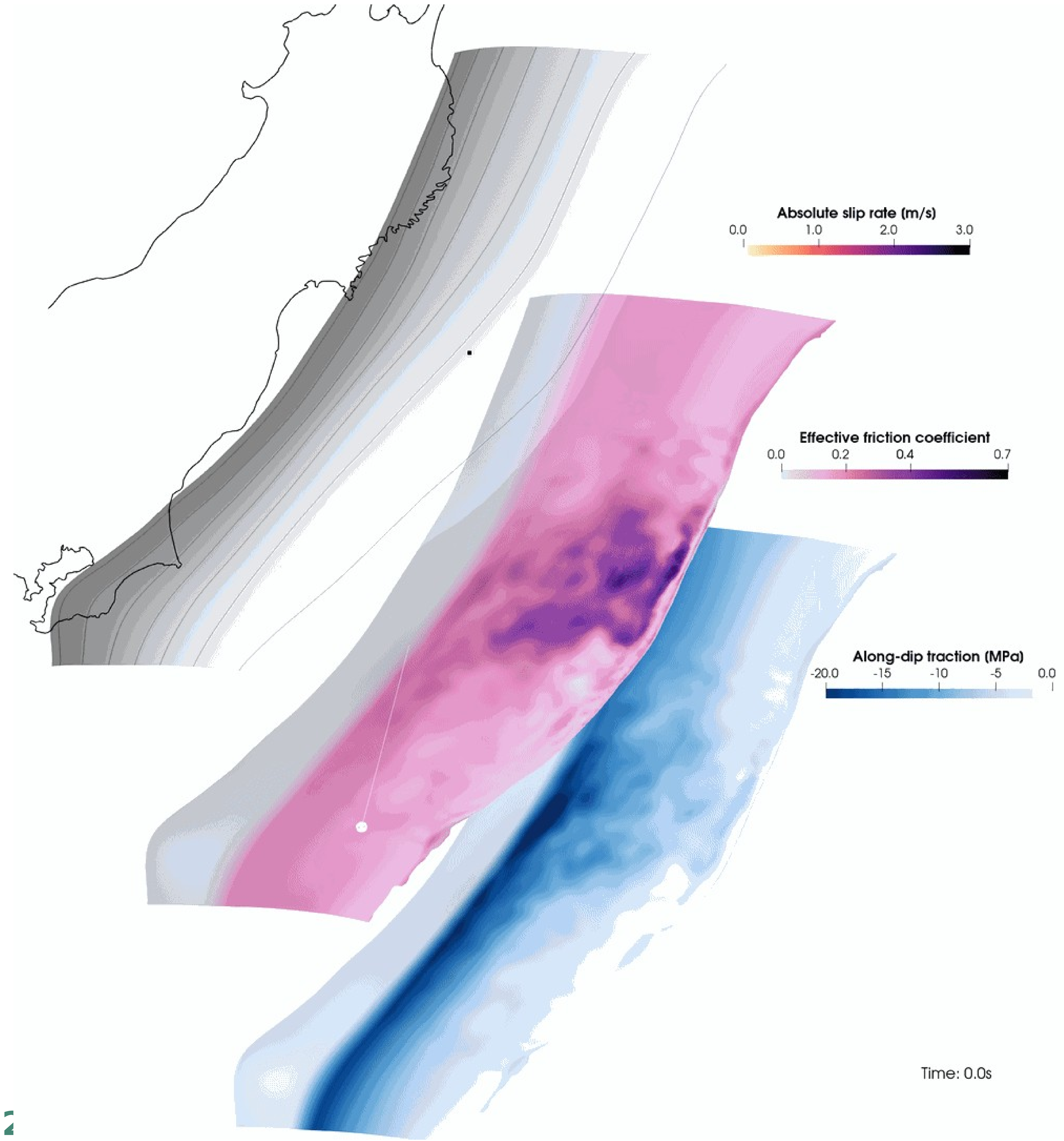
“Hero runs”

Gabriel et al. Science 2024;
Palgunadi et al., JGR 2023 & BSSA 2025


Taufiqurrahman et al., Nature, 2023 & Abrams et al., SC23, Visualizatio
showcase
10 Hz Ridgecrest simulation required 5.5 h on 7000 nodes of Frontera and a 170 mio elements, with order-6 space-time accuracy. 20

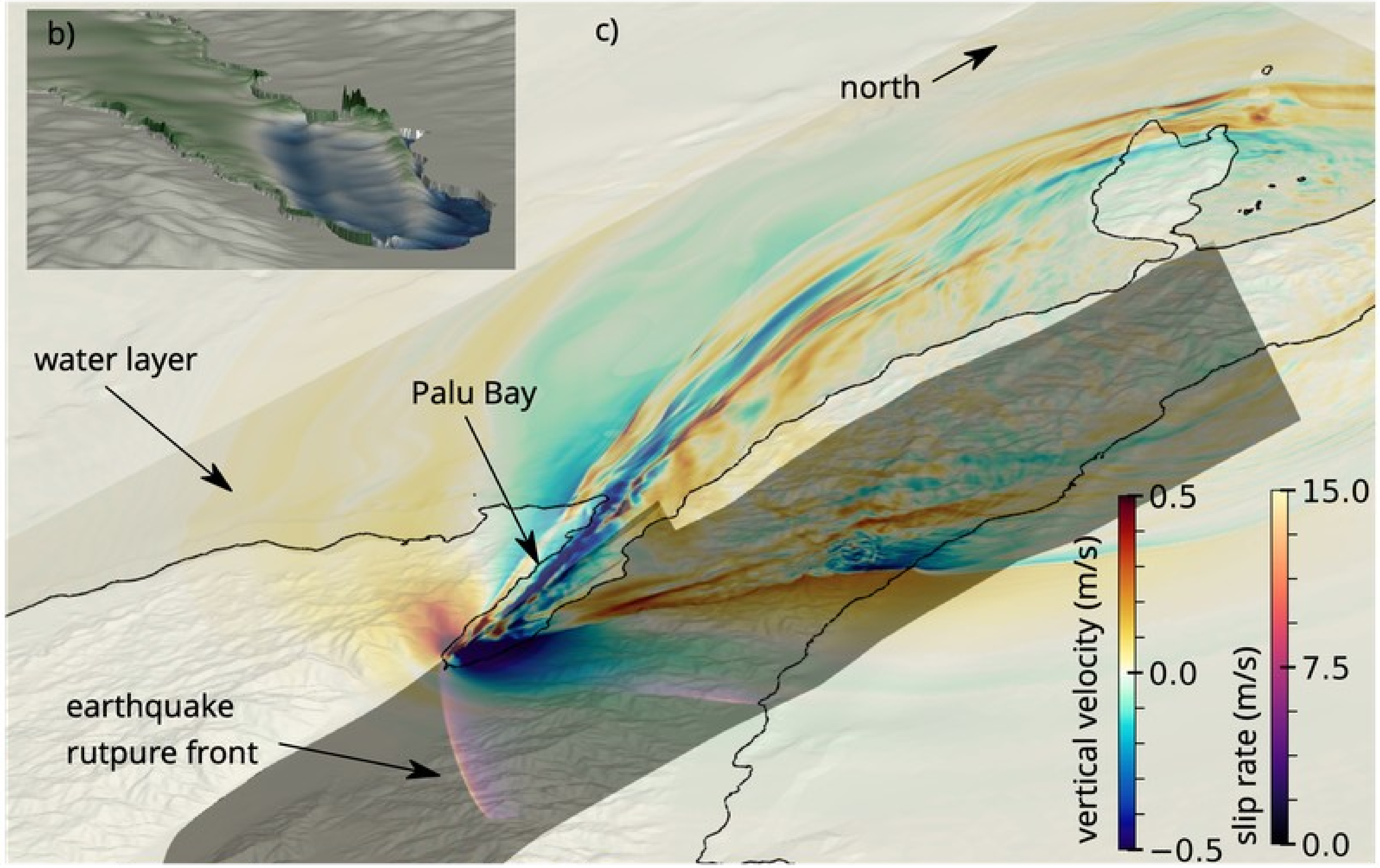
Explaining the 2011 Tohoku, Japan, earthqua





Wong, Gabriel, & Fan, Nature Comms, 2

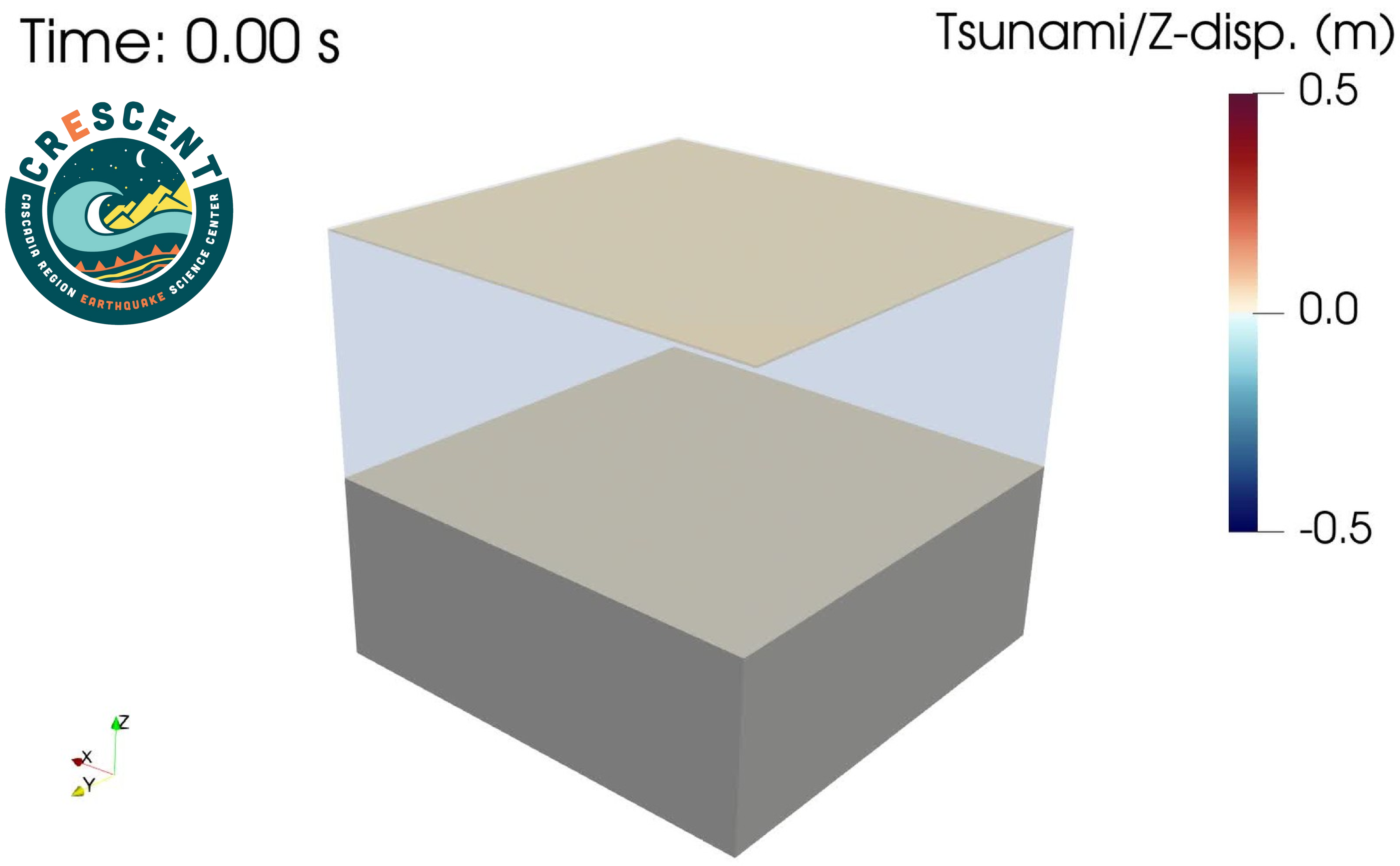
Fully-coupled earthquake-tsunami simulations



Krenz et al., SC'2021

- Large differences [redacted] sizes and wave speeds can be mitigated by **local time-stepping** in a high-order accurate ADER-DG discretization
- **Multi-petaflop simulation:** 518 mio. elements = 261 billion degrees of freedom, required 5.5 hours on 3,072 nodes of SuperMUC-NG, sustained performance at 3.1 PFLOPS, with excellent parallel efficiency, >70% on SuperMUC-NG (Germany) and Mahti (Finland) and >95% efficiency on Frontera
- Using **SeisSol** a NSF/TACC Leadership Class Computing Facility application

Fully-coupled earthquake-tsunami simulations

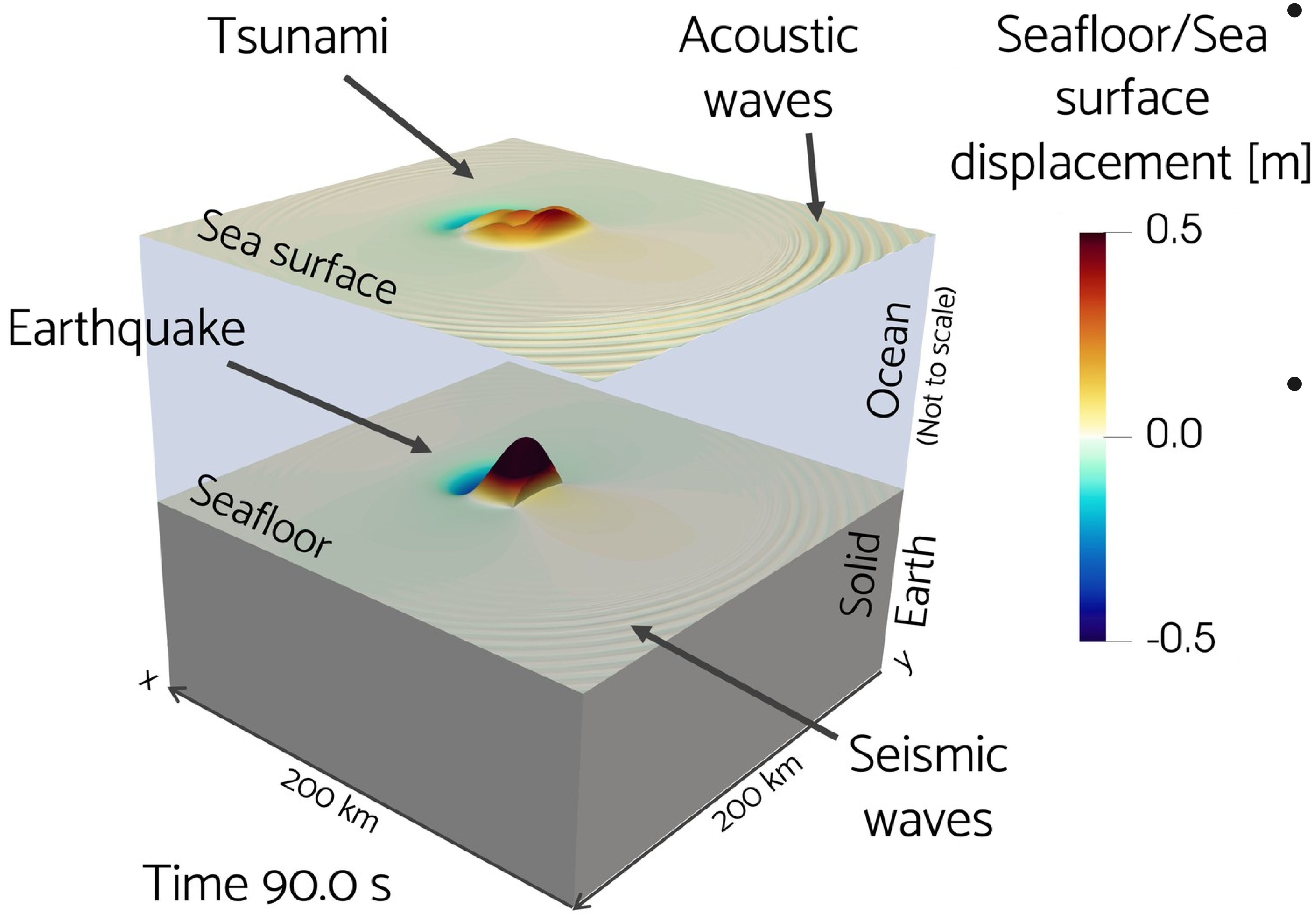


CRESCENT Tsunami Community Benchmark (animation by Fabian Kutschera)

- Fully-coupled sim[redacted] combine earthquake dynamic rupture and tsunami generation **into a single simulation**
- We can **capture 3D elastic, acoustic, and tsunami waves, including oceanic Rayleigh waves and dispersion effects, simultaneously** (e.g., Lotto & Dunham, 2015)

Fully-coupled earthquake-tsunami simulations

[redacted]



CRESCENT Tsunami Community Benchmark (animation by Fabian Kutschera)

- Fully-coupled sim[redacted] combine earthquake dynamic rupture and tsunami generation **into a single simulation**

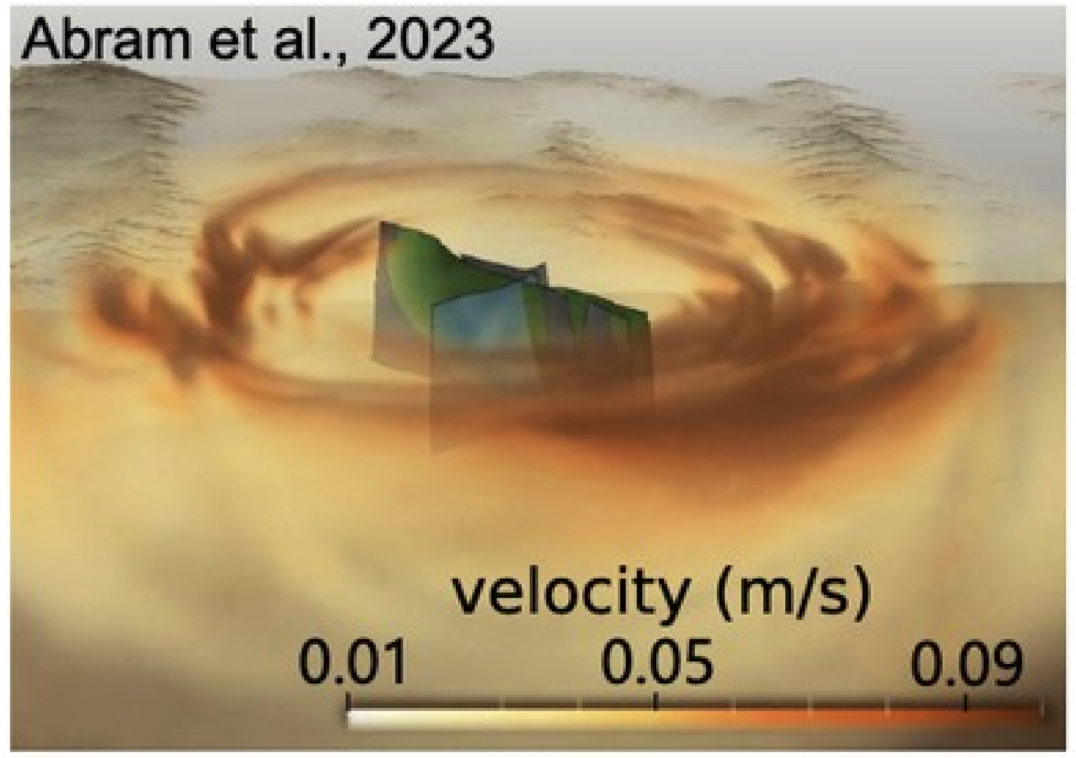
- We can **capture 3D elastic, acoustic, and tsunami waves, including oceanic Rayleigh waves and dispersion effects, simultaneously** (e.g., Lotto & Dunham, 2015)

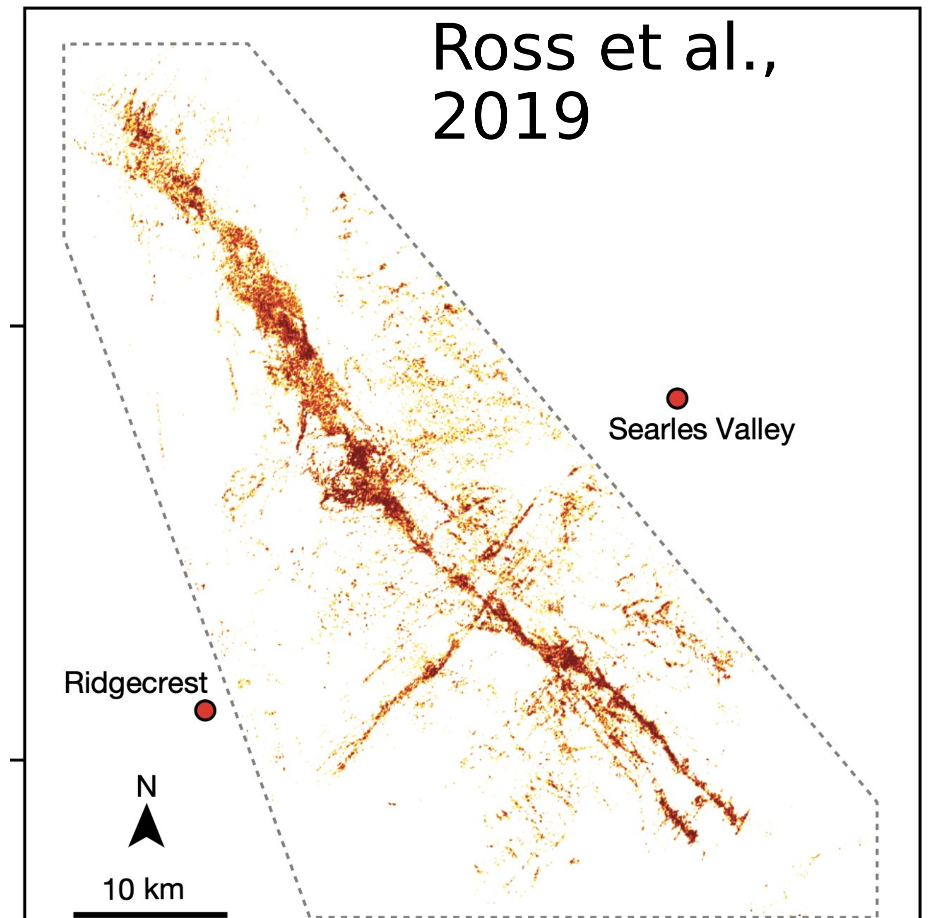
Thanks to high-performance computing, earthquake science is increasingly inverse model-rich

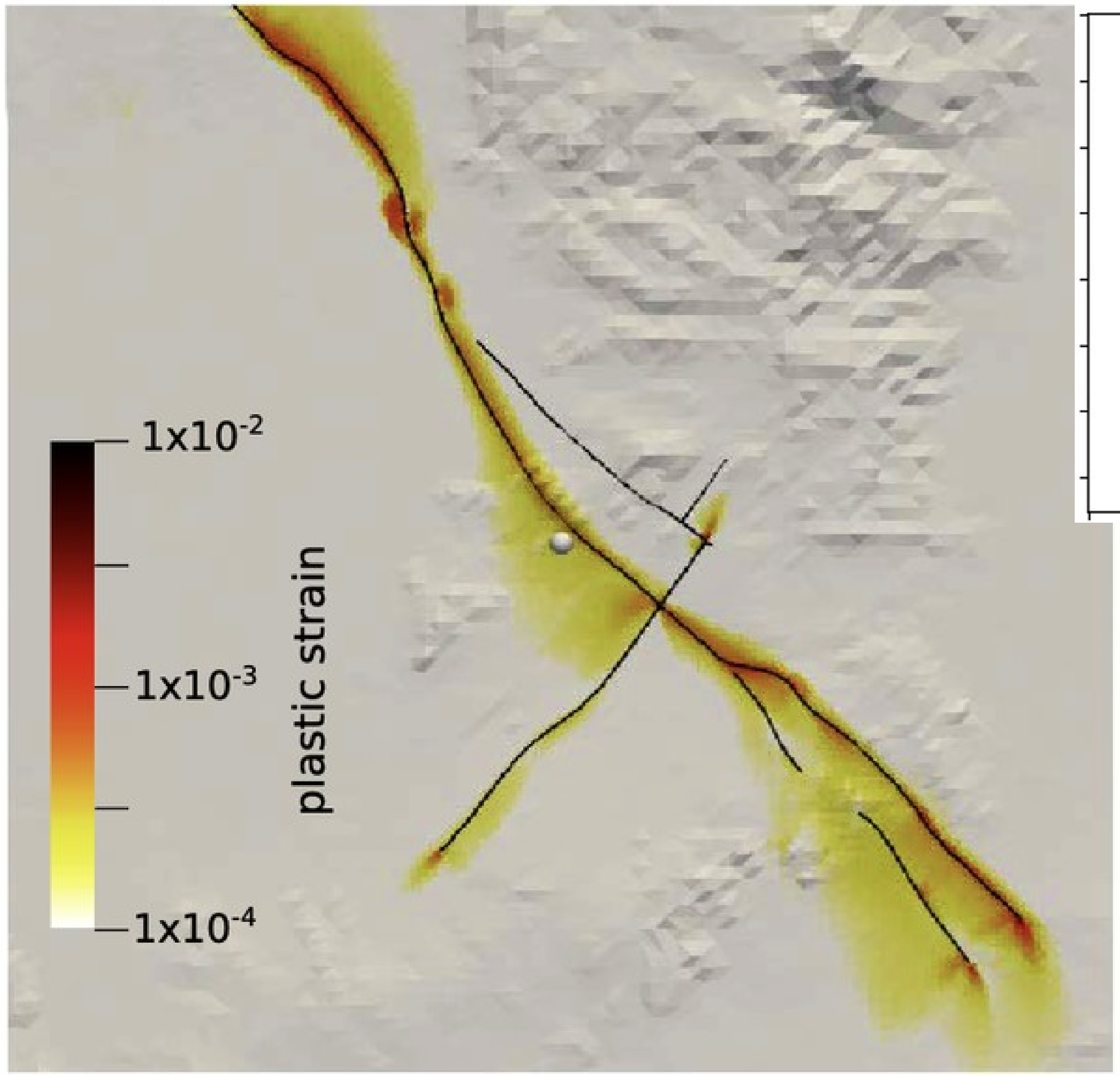
- **Dynamic source inversion** usually requires simple dynamic rupture model setups and thousands to millions of runs (e.g., Premus et al., 2022; Schliwa et al., 2024)

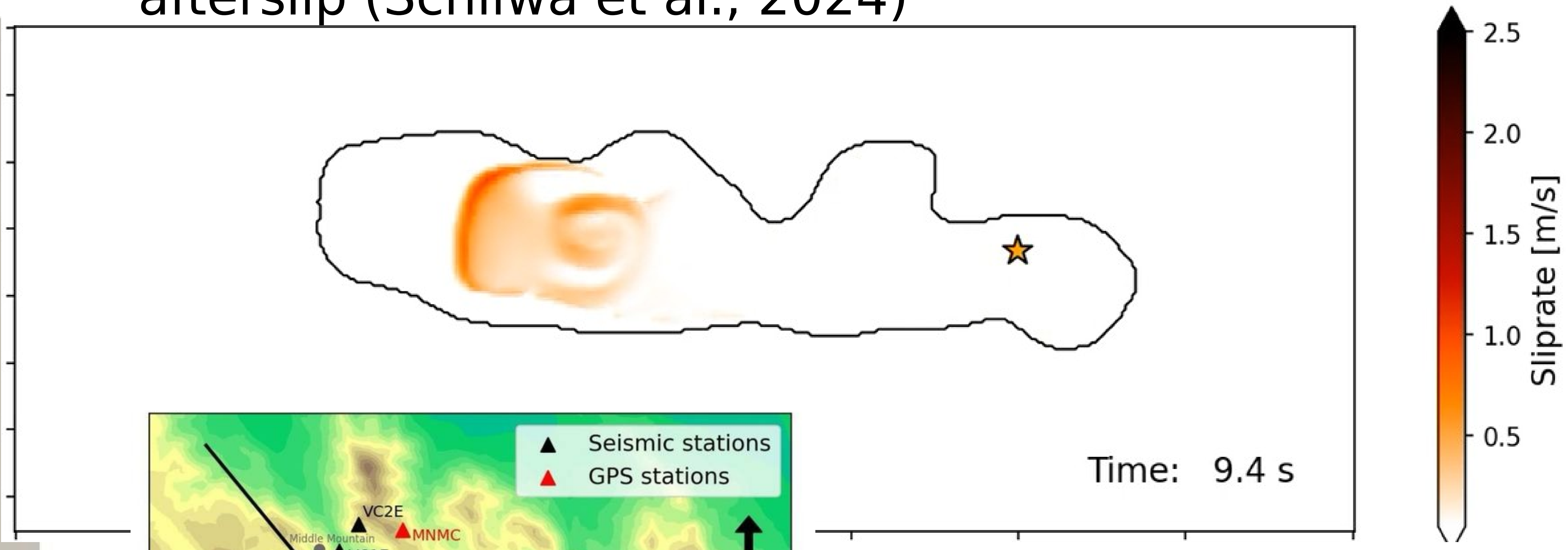
Modelled off-fault Drucker-Prager plasticity of the 2019 Ridgecrest mainshock (Taufiqurrahman et al., 2023)

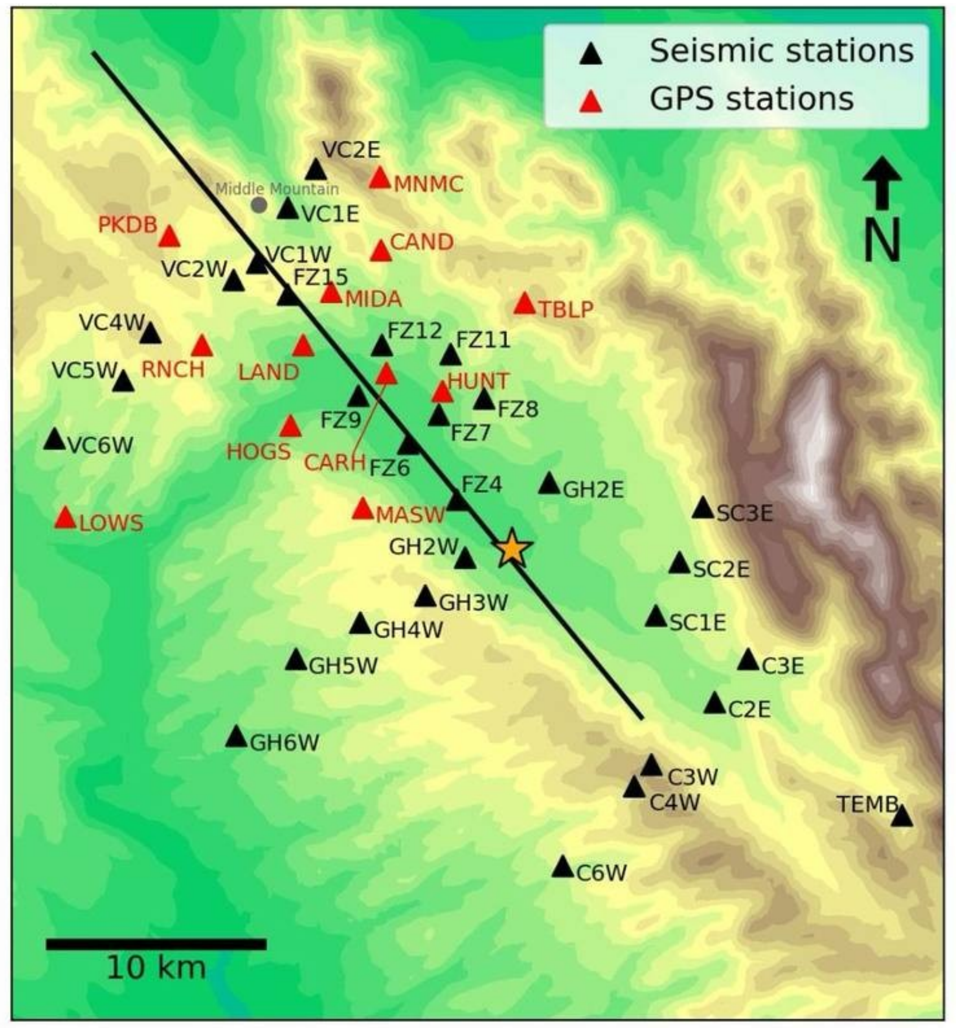
2 million forward models for a dynamic source inversion of the 2004 Parkfield earthquake and afterslip (Schliwa et al., 2024)











Inversion

Kruse at al., PASC25; Niu et al., EPSL, in pre

- Serial MCMC is slow to converge and has high sample correlation

- MLDA: Hierarchy of models, reduces the correlation between our samples by **90%**

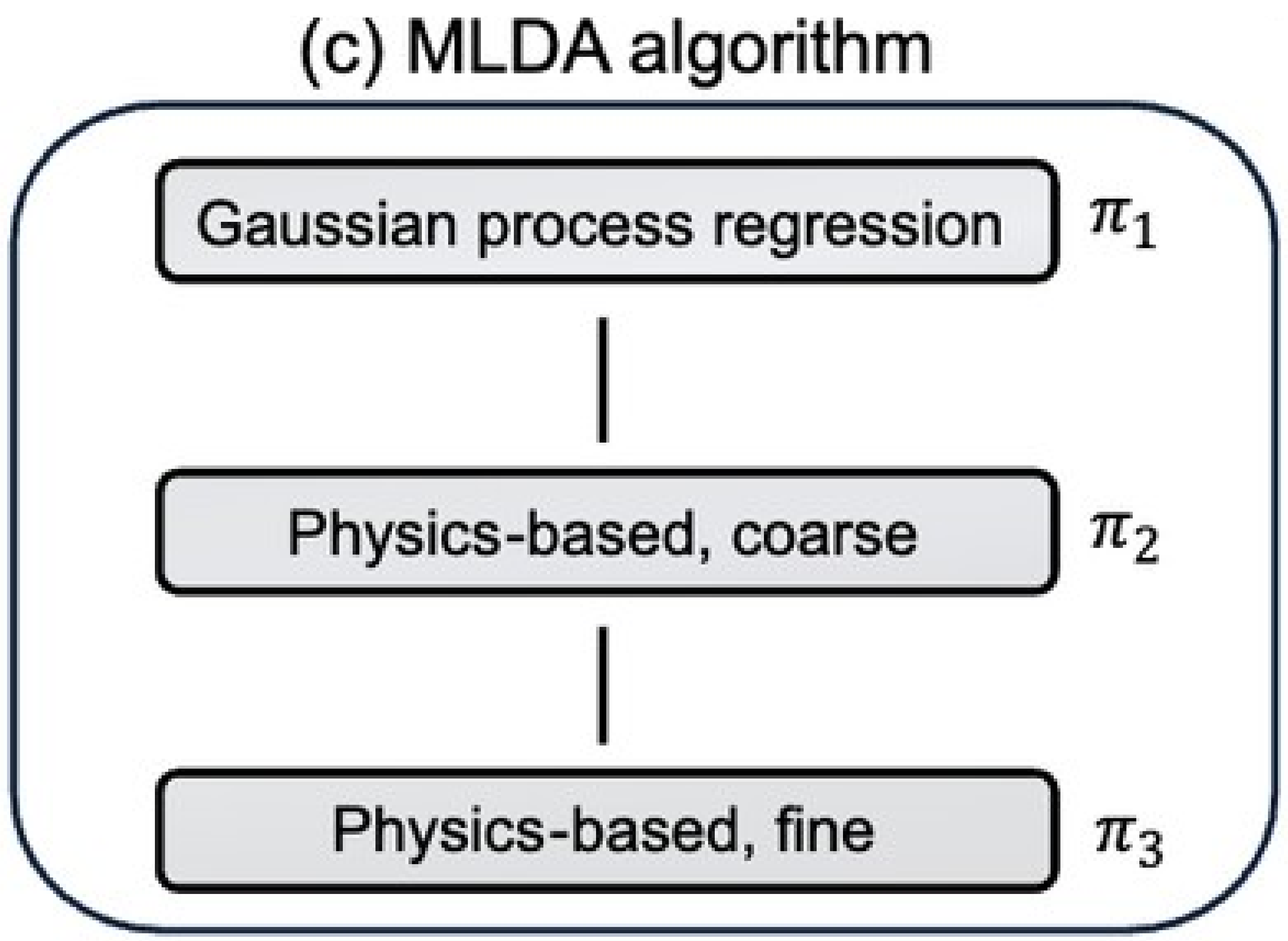
- (1) **Gaussian Process Surrogate,** discards 99% of 'bad' models; (2) **Coarse Mesh Simulation;** (3) **High-resolution,** only runs when first two levels agree that a parameter set is promising

RESEARCH-ARTICLE | OPEN ACCESS

Scalable Bayesian Inference of Large Simulations via Asynchronous Prefetching Multilevel Delayed Acceptance

Authors: Maximilian Kruse, Zihua Niu, Sebastian Wolf, Mikkel Lykkegaard, Michael Bader, Alice-Agnes Gabriel, Linus Seelinger | Authors Info & Claims

PASC '25: Proceedings of the Platform for Advanced Scientific Computing Conference • Pages 1 - 13
https://doi.org/10.1145/3732775.3733581



- **Asynchronous Prefetching,** builds Markov tree of accept/reject steps
- **MLDA** ,~4 million CPU hours on Frontera (TACC), = **19h,** a 200% speedup on 4096 nodes of Frontera compared to a single model evaluation

Fused ensembles of dynamic-rupture earthquake simulations to accelerate Bayesian inference

Kurapati at al., GEM, 2026

- **Fusing** many simulations into one run adds an ensemble dimension to the element-local degrees of freedom (Breuer et al., ISC’2017)
- Turning many small matrix operations into **tensor contractions** that use CPU SIMD registers more efficiently
- Up **to 5.54× more simulations per node-hour** at order 3, and 2.93× at order 4, on the CPU-based SuperMUC-NG system
- Gains arise from improved SIMD utilization and depend strongly on discretization order, problem size, nonlinear

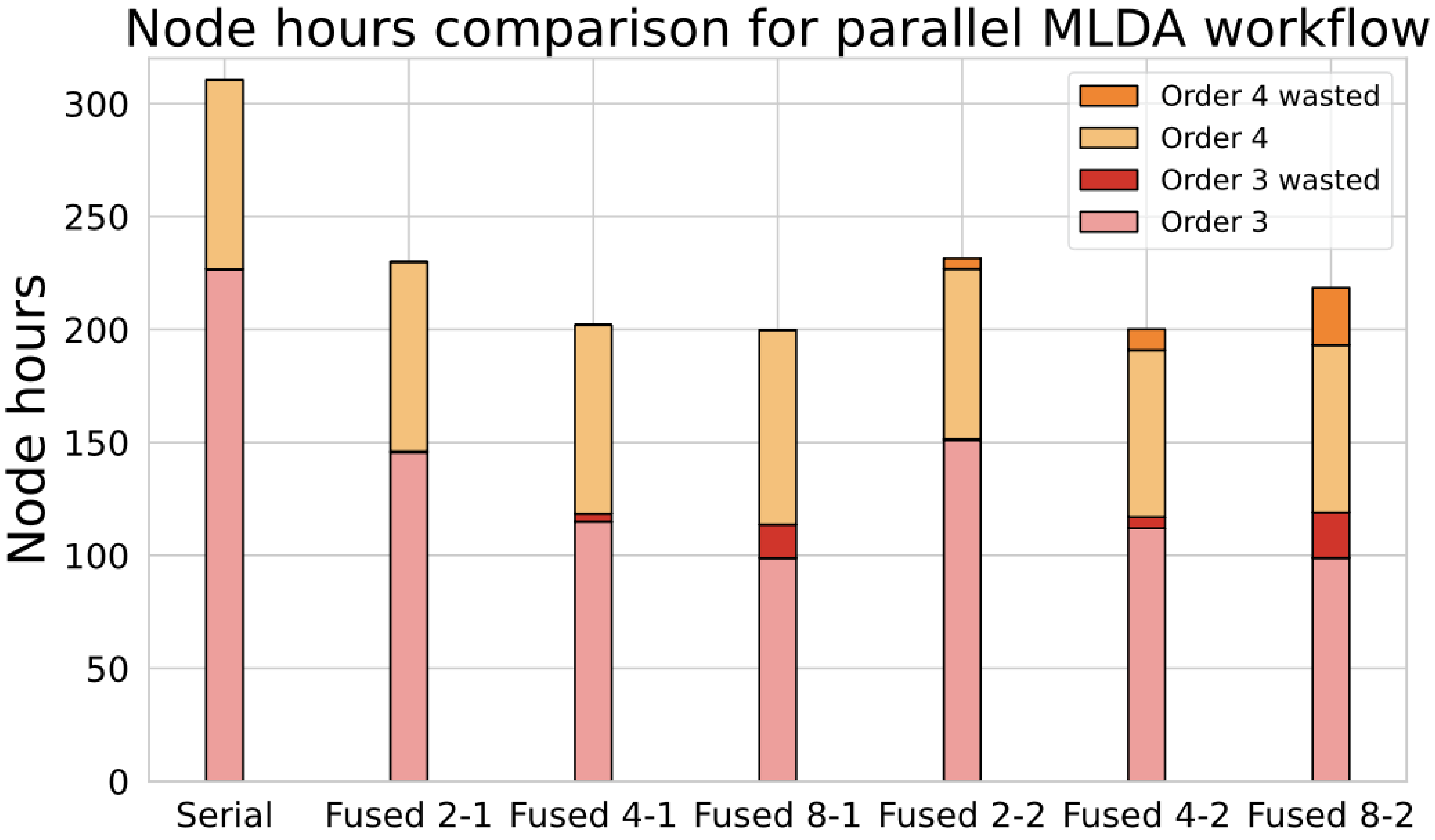
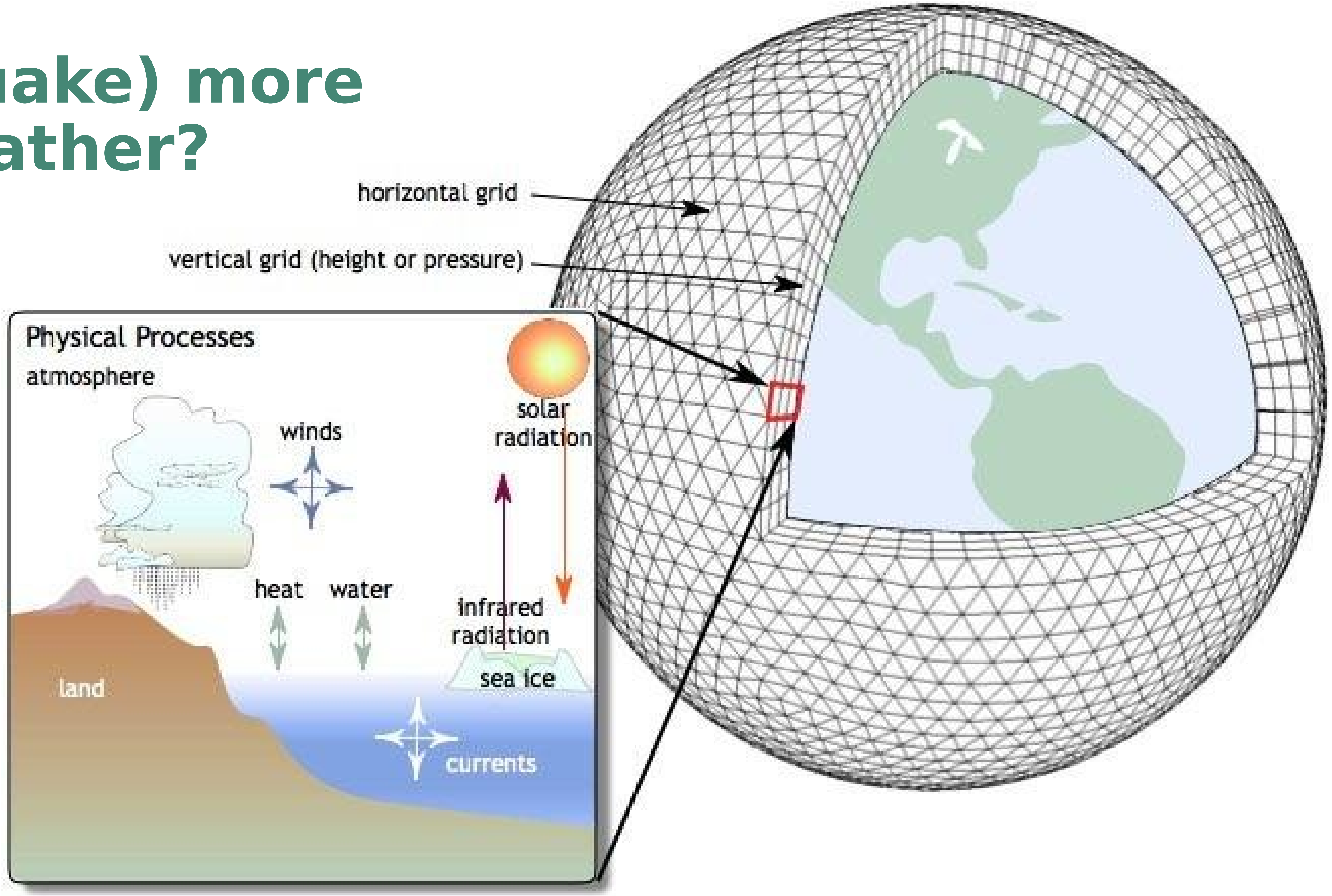
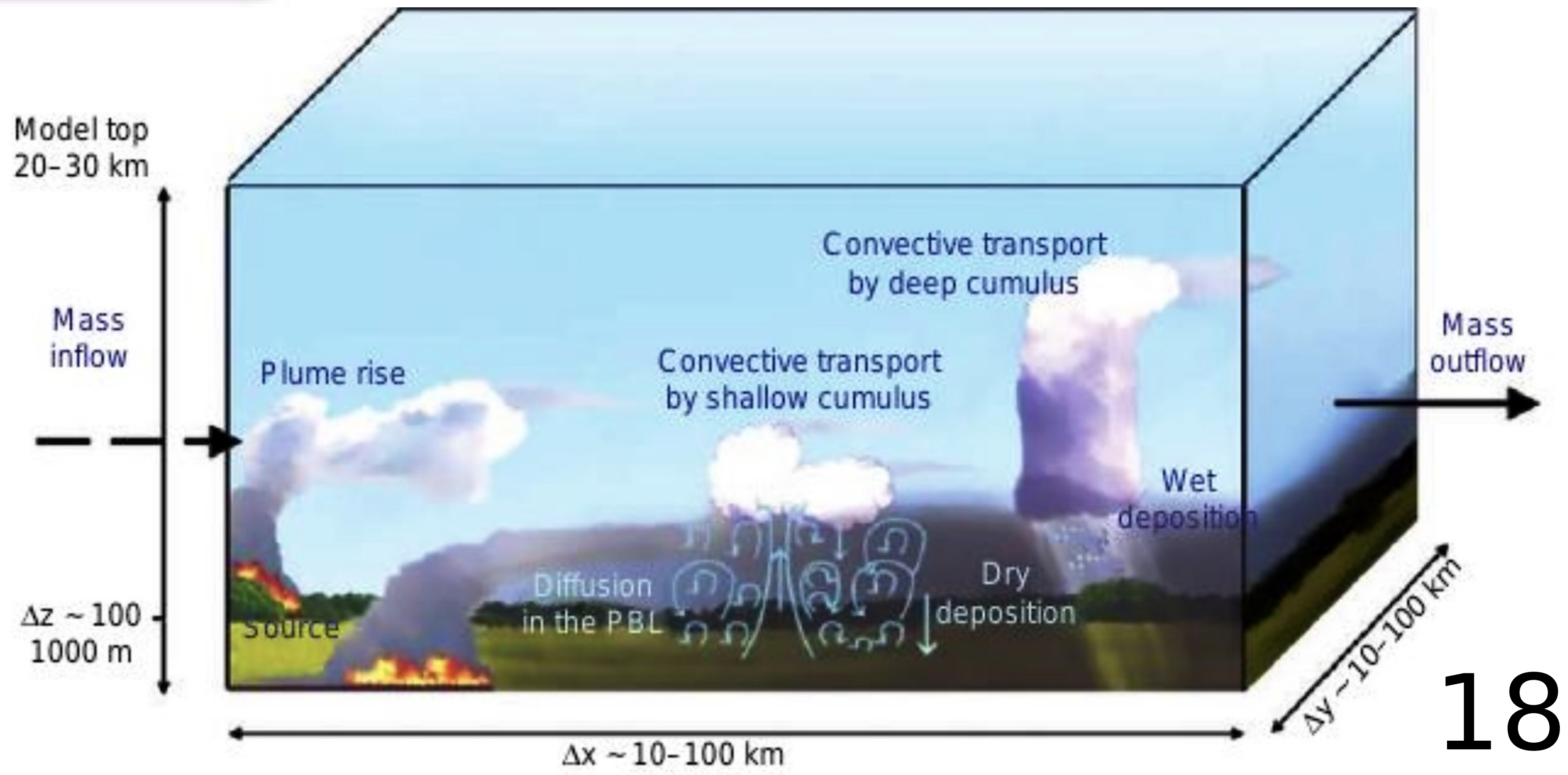


Fig. 11 Comparison of resource usage between serial and fused simulations for the MLDA workflow. “Fused-S_c-S_f” denotes that S_c simulations are fused for coarse-level simulations of the MLDA workflow, and S_f simulations are fused for the fine level. Wasted simulations denote the amount of node hours wasted due to the padding of parameter batches in fused simulations

Toward digital twins of earthquakes and tsunamis

Why are tsunami (& earthquake) more difficult to forecast than weather?

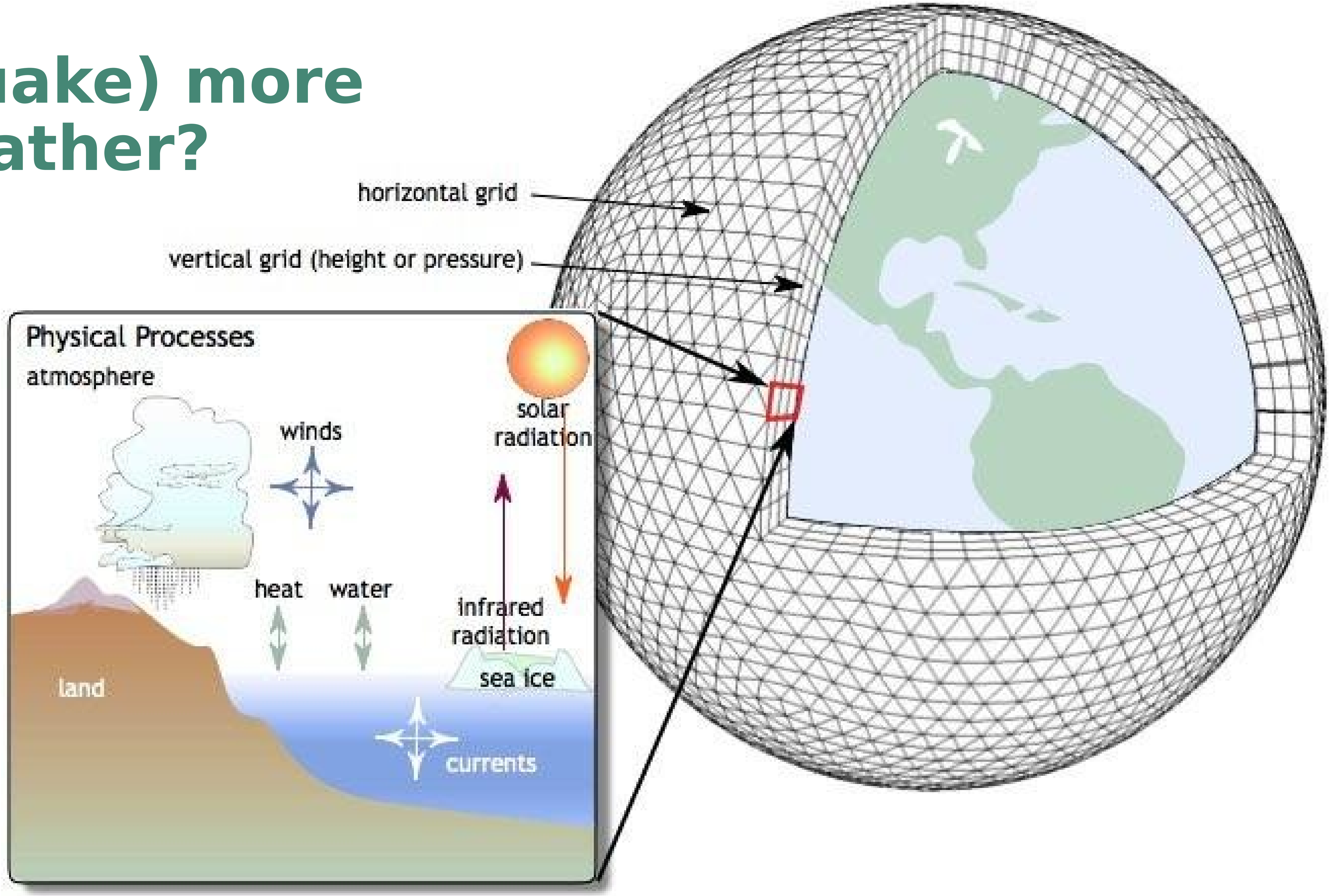


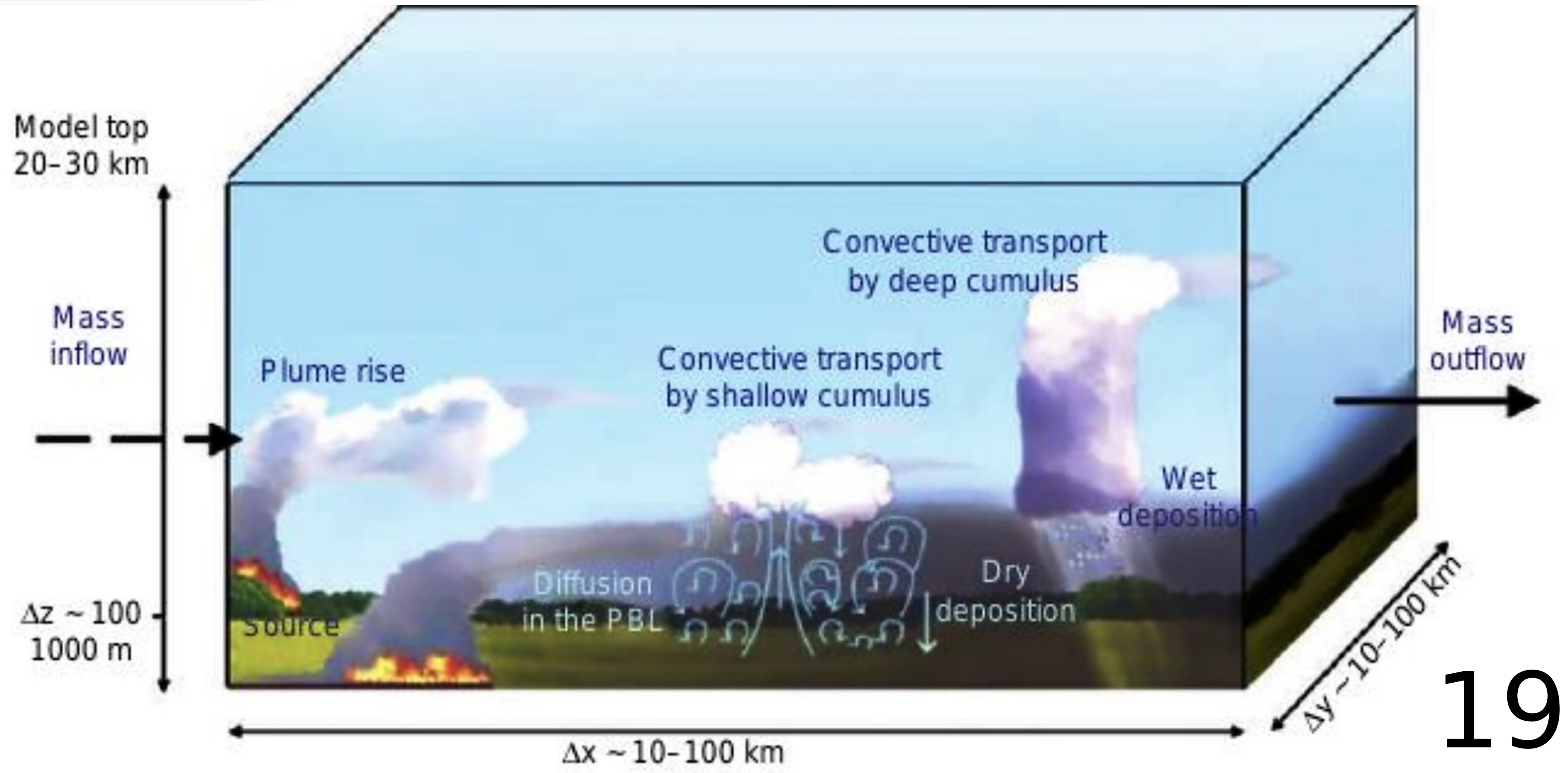


Why are tsunami (& earthquake) more difficult to forecast than weather?

- Data streams are **sparse** in space and **sparse** in time, challenging data assimilation

- **Fundamental physics are yet debated**, including rock friction; off-fault deformation; the long-term and short-term strength of faults; the energy budget of earthquakes, and tsunami generation mechanisms



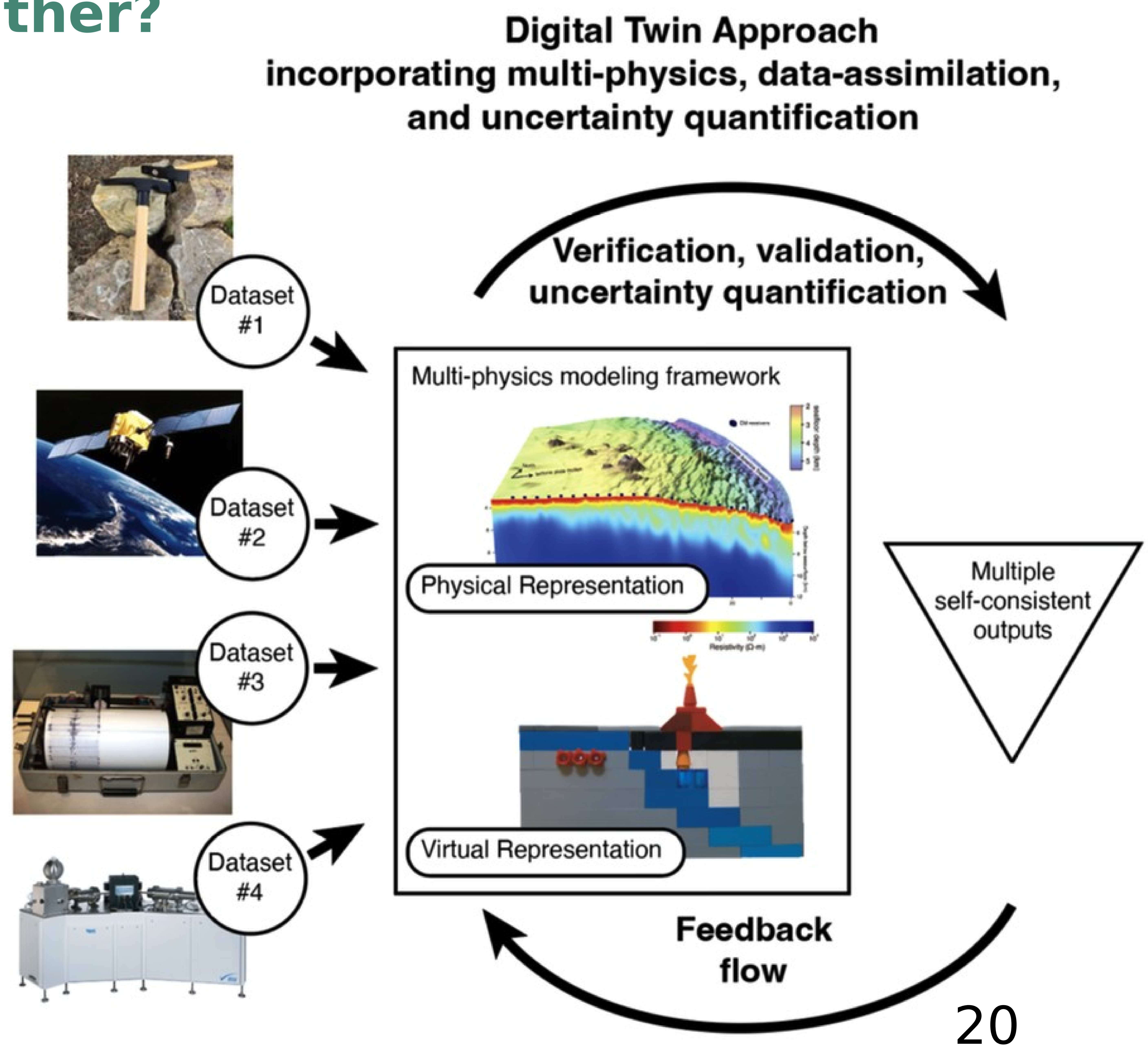


Why are tsunami (& earthquake) more difficult to forecast than weather?

- Data streams are **sparse** in space and **sparse** in time, challenging data assimilation

- **Fundamental physics are yet debated**, including rock friction; off-fault deformation; the long-term and short-term strength of faults; the energy budget of earthquakes, and tsunami generation mechanisms

- **Tsunami Digital Twins need both: Big Data and (Very) Big Physics-Based Simulations to address this sparsity and to combine rich, continuous and interdisciplinary data-sets in a core dynamic framework, allowing for hypothesis testing**



UC San Diego

Real-Time Bayesian Inference at Extreme Scale: A Digital Twin for Tsunami Early Warning Applied to the Cascadia Subduction Zone

Stefan Henneking, Sreeram Venkat, Veselin Dobrev, John Camier, Tzanio Kolev, Milinda Fernando, Alice-Agnes Gabriel, Omar Ghattas


Stefan Henneking (UTA)


Sreeram Venkat (UTA)


Milinda Fernando (UTA)


Omar Ghattas (UTA)


Alice Gabriel (UCSD)


Veselin Dobrev (LLNL)


John Camier (LLNL)


Tzanio Kolev (LLNL)

ACM Gordon Bell Prize: Inverse problems

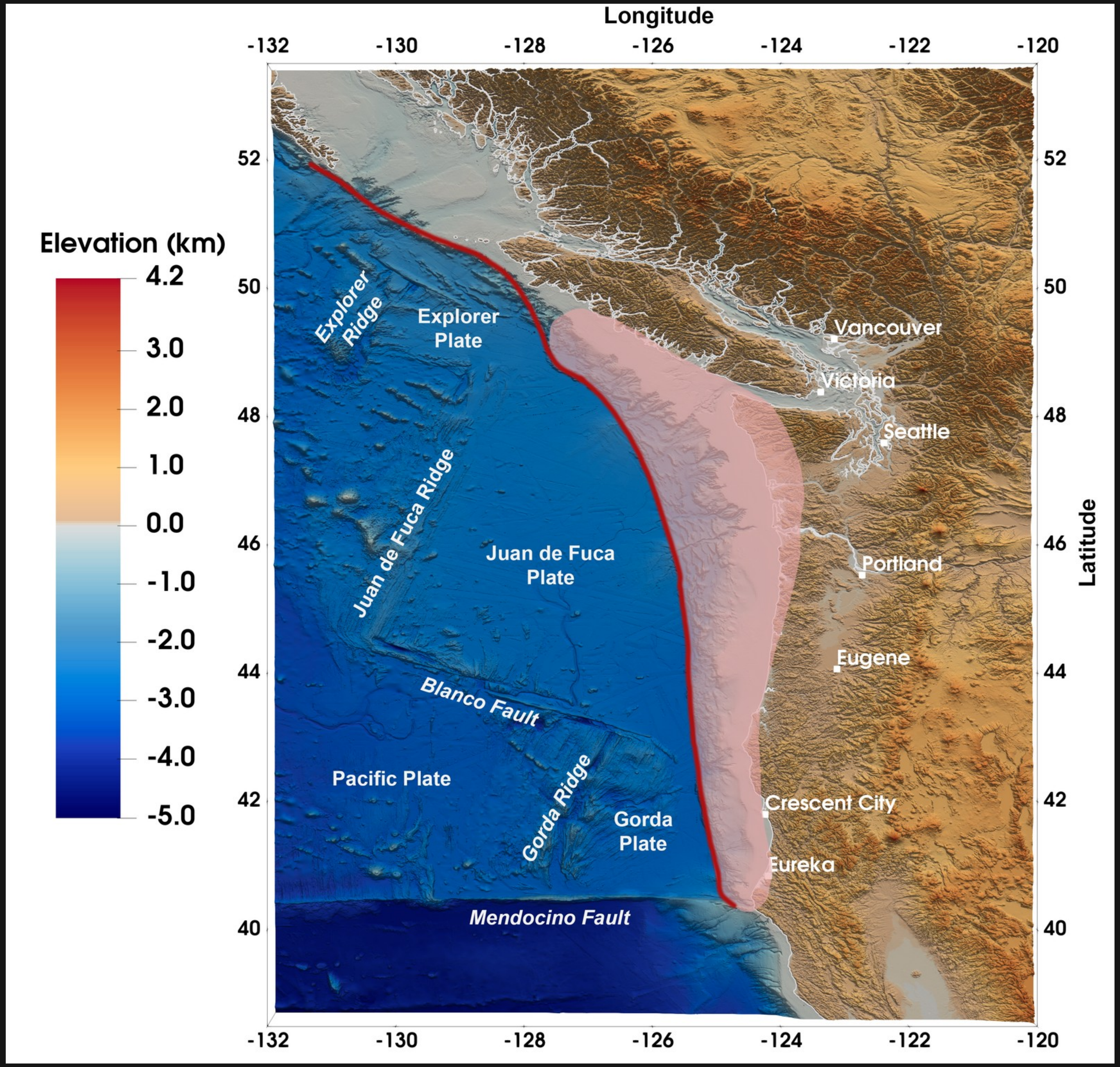
<u>Previous Bell Prize Winners related to inverse problems</u>

- Akçelik et al. "High resolution forward and inverse earthquake modeling on terascale computers" (2003)
- Das et al. "Large-scale materials modeling at quantum accuracy: Ab initio simulations of quasicrystals and interacting extended defects in metallic alloys" (2023) [Inverse DFT]

<u>This work</u>

- <u>Extreme-scale:</u> Inverse problem for > 1 billion parameters
- <u>High-fidelity:</u> Full-physics PDE model (no surrogates!)
- <u>UQ-equipped:</u> Bayesian framework quantifies uncertainties
- <u>Real-time:</u> Inference and prediction in seconds

Cascadia subduction zone



- Last major earthquake and tsunami: January 26, 1700
- Paleoseismic evidence indicates a 250–500 years recurrence interval
- Current estimate: 37% probability of M8.2+ event within 50 years; 10-15% probability that entire Cascadia subduction zone will rupture with an M9+

Image credit: Rob DeGraff/Flickr, 2008

Proposed sensor network

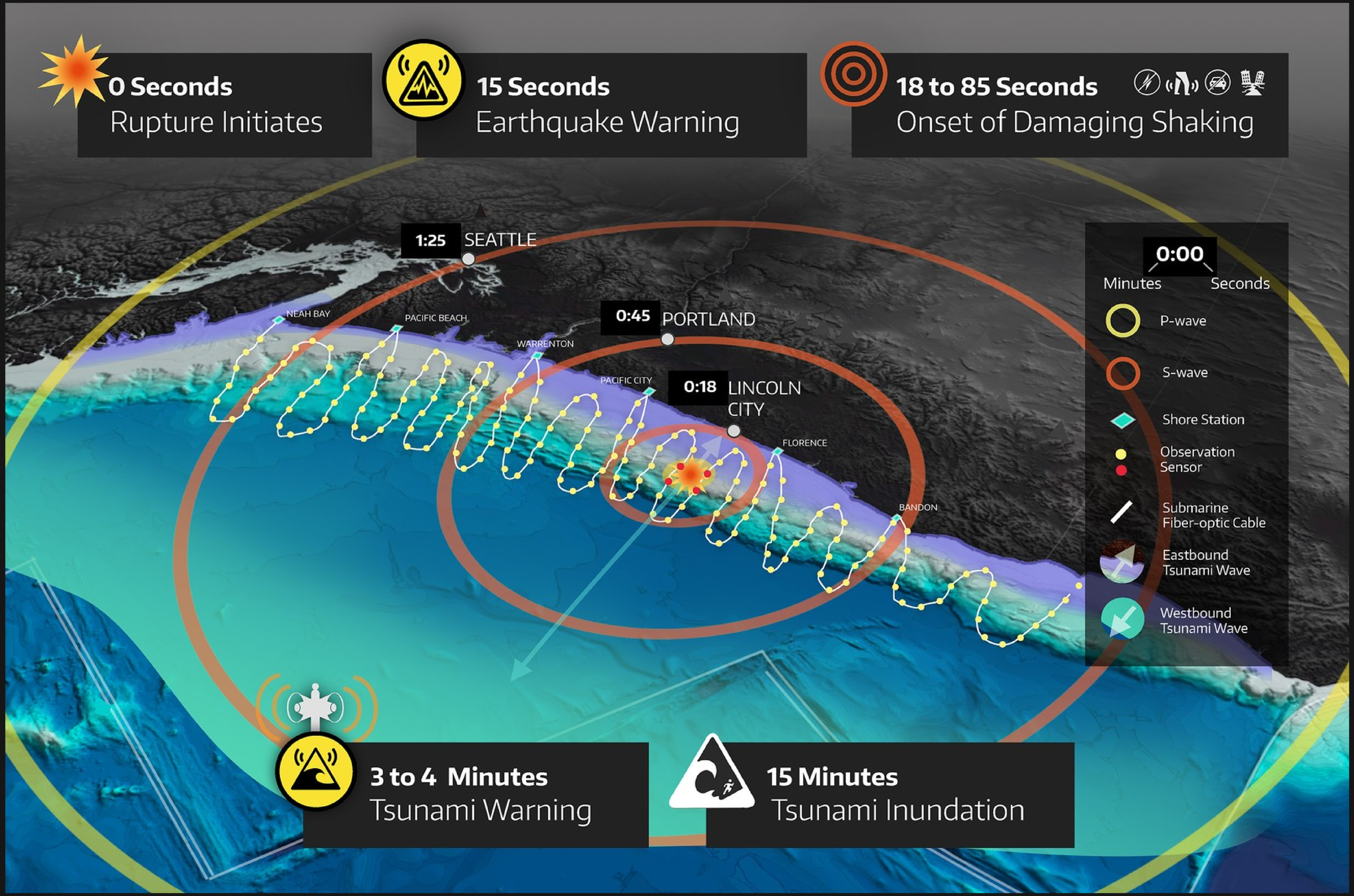


Image credit: D. Schmidt et al., 2019

Digital twin for tsunami forecasting

Current forecasting models rely on shallow water equations for tsunami propagation

- Efficient in the far-field; does not make use of near-field pressure transients from hydroacoustic waves

Our digital twin approach

- Use near-field pressure observations to **infer** the seafloor motion and **forward predict** the tsunami propagation
- Employ high-fidelity, **full-physics** model (coupled acoustic–gravity wave)
- Quantify **uncertainty** via Bayesian inference
- Solve inverse problem in **real time** (order of seconds)

The forward problem

<u>Forward model</u>

- Parameters: $m(\vec{x}, t)$ — model inputs (sources, initial conditions, coefficients, ...)
  - *<u>Here:</u> spatiotemporal seafloor motion*
- Data: $d(\vec{x}, t)$ — observables (state variables, quantities of interest, ...)
  - *<u>Here:</u> pressure observations taken from seafloor sensors*
- Parameter-to-observable map $F: m(\vec{x}, t) \mapsto d(\vec{x}, t)$
  - *<u>Here:</u> coupled acoustic–gravity wave propagation model*

<u>The forward problem</u>

- Given model parameters m, solve forward model F to yield observables d
- Well-posed: solution exists, is unique, and is stable to perturbations in inputs

The inverse problem

- Given observations d and forward model F, infer model parameters m
- **Inverse problem is ill-posed**
  - Observations are usually sparse and noisy
  - Many different parameters may be consistent with the data

Bayesian approach to ill-posedness

- Characterize the probability that any parameter field is consistent with the data and prior knowledge
- The uncertain parameter is treated as a random field, and the solution of the inverse problem becomes a probability density, the posterior distribution
- Bayes' theorem: $\pi_{\text{post}}(m|d) \propto \pi_{\text{prior}}(m)\, \pi_{\text{like}}(d|m)$

Bayesian framework, linear inverse problem

Gaussian prior with Matérn covariance:

$$\pi_{\text{prior}}(m) \propto \exp\left(-\frac{1}{2}\left\|m - m_{\text{prior}}\right\|^2_{\Gamma^{-1}_{\text{prior}}}\right), \quad \Gamma_{\text{prior}} = (\alpha_1 I - \alpha_2 \Delta_{2D})^{-2}$$

Observations: $d = Fm + \nu, \quad \nu \sim \mathcal{N}(0, \Gamma_{\text{noise}})$

Likelihood: $\pi_{\text{like}}(d|m) \propto \exp\left(-\frac{1}{2}\|Fm - d\|^2_{\Gamma^{-1}_{\text{noise}}}\right)$

Posterior: $\pi_{\text{post}}(m|d) \propto \exp\left(-\frac{1}{2}\|Fm - d\|^2_{\Gamma^{-1}_{\text{noise}}} - \frac{1}{2}\left\|m - m_{\text{prior}}\right\|^2_{\Gamma^{-1}_{\text{prior}}}\right)$

$$\propto \mathcal{N}\left(m_{\text{map}}, \Gamma_{\text{post}}\right)$$

$$\Gamma_{\text{post}} = \left(F^*\Gamma^{-1}_{\text{noise}}F + \Gamma^{-1}_{\text{prior}}\right)^{-1}$$

$$m_{\text{map}} = \Gamma_{\text{post}}\left(F^*\Gamma^{-1}_{\text{noise}}d + \Gamma^{-1}_{\text{prior}}m_{\text{prior}}\right)$$

SoA: Adjoint-based matrix-free methods

- $\mathrm{H} := \Gamma_{\mathrm{post}}^{-1}$ is the Hessian; each matrix-free Hessian action requires 1 forward and 1 adjoint model solve
- Usual approach relies on (randomized) low-rank decomposition of $\mathrm{F}^*\Gamma_{\mathrm{noise}}^{-1}\mathrm{F}$ to invert H efficiently

<u>Cascadia subduction zone digital twin</u>

- $\dim(\mathrm{m}) \sim 10^9$ parameters, $\dim(\mathrm{d}) \sim 10^5$ data
- Governed by hyperbolic wave propagation model
  - Hessian has high effective $\mathrm{rank}(\mathrm{H}) \sim 10^5$
  - ~ 1 hour / Hessian action → **<u>decades</u>** just to find $\mathrm{m}_{\mathrm{map}}$
- Traditional fix of surrogates (e.g., ROMs, NNs, ...) is not feasible
  - Very large parameter dimension, Kolmogorov *N*-width, high-frequency

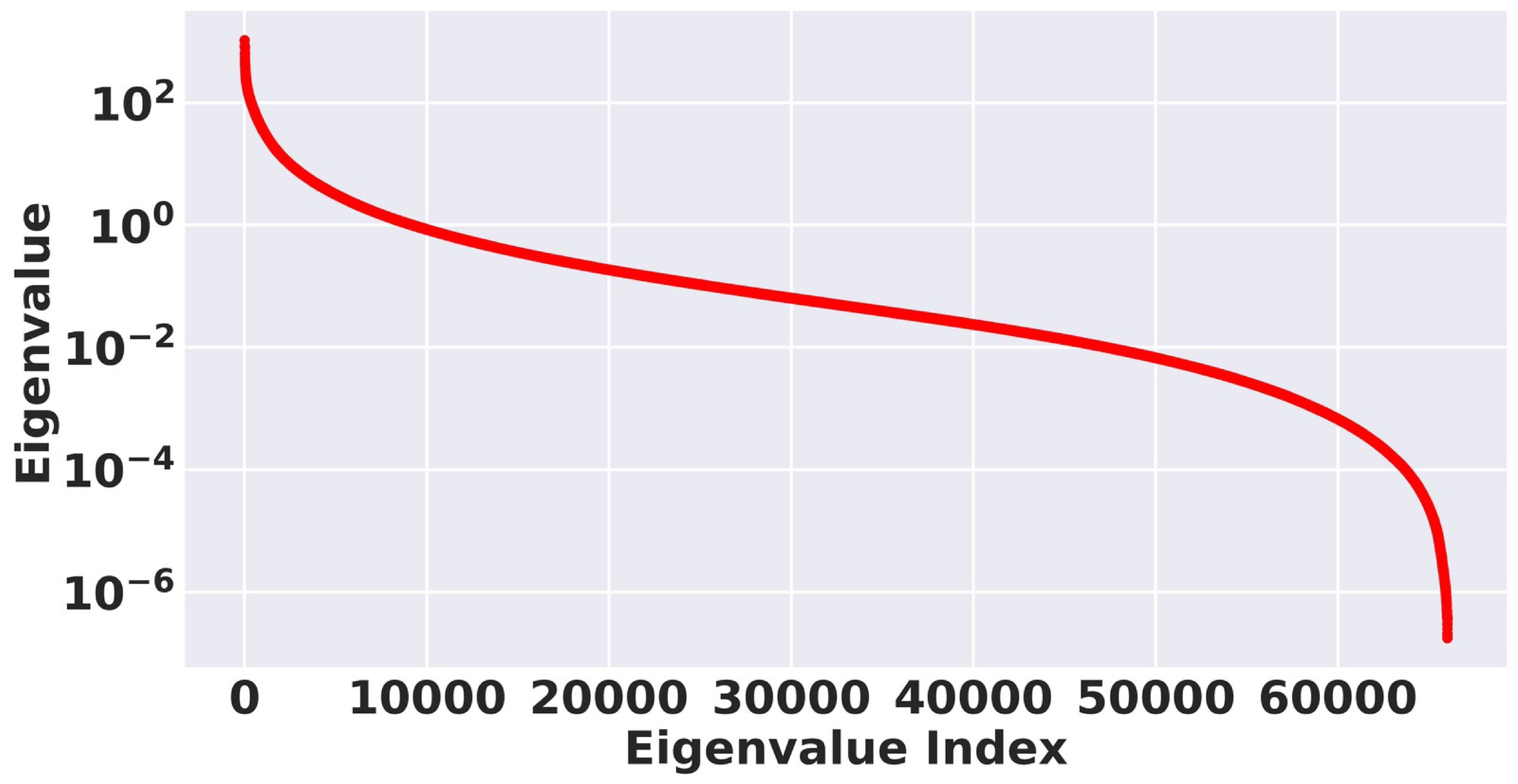
Conventional methods

- Need to solve large-scale linear system $\mathrm{Hm}_{\mathrm{map}} = \mathrm{f}(\mathrm{d}^{\mathrm{obs}})$ in real time!
- Cascadia digital twin: $\dim(\mathrm{m}) \sim 10^9$ parameters, $\dim(\mathrm{d}^{\mathrm{obs}}) \sim 2.5 \cdot 10^5$ data

Conventional methods are intractable!

- Constructing the Hessian $\mathrm{H} = \Gamma_{\mathrm{post}}^{-1}$ requires 1 model solve per column of F: $\sim 2.5 \cdot 10^5$ adjoint PDE solutions → ~25 years on 512 A100 GPUs
- Storing the Hessian in dp requires ~4 exabytes

Data misfit Hessian is not low rank!



Eigenvalues of the "inside-out" prior-preconditioned Hessian of the data misfit $F\Gamma_{\text{prior}}F^*$

Hessian <u>does not</u> readily <u>admit</u> efficient/accurate <u>low rank approximation</u>, resulting in a prohibitive cost of computing inverse solution the "usual way"

Acoustic–gravity model

- Partial differential equation (PDE) model

$$\rho \frac{\partial \vec{u}}{\partial t} + \nabla p = 0, \qquad \frac{1}{K}\frac{\partial p}{\partial t} + \nabla \cdot \vec{u} = 0, \qquad \text{in } \Omega \times (0, \mathrm{T})$$

- Boundary conditions
  - Sea surface (coupling with surface-gravity wave)

$$p = \rho g \eta, \qquad \partial \eta / \partial t = \vec{u} \cdot \vec{n}, \qquad \text{on } \Gamma_{\text{surface}} \times (0, \mathrm{T})$$

  - Seafloor velocity (boundary source)

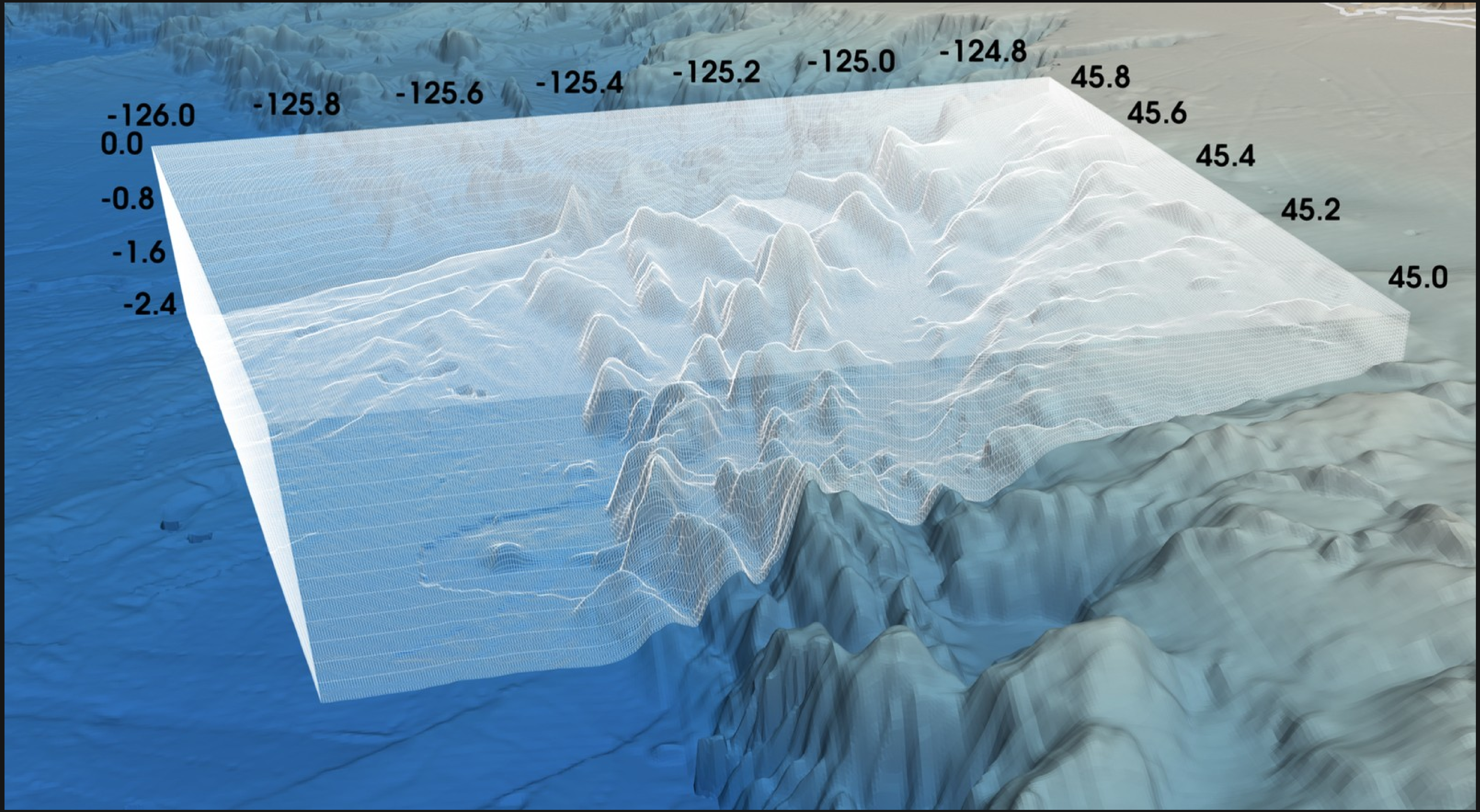
$$-\partial b / \partial t = \vec{u} \cdot \vec{n}, \qquad \text{on } \Gamma_{\text{bottom}} \times (0, \mathrm{T})$$

  - Absorbing boundary (outgoing waves)

$$\vec{u} \cdot \vec{n} = p / \rho c, \qquad \text{on } \Gamma_{\text{absorb}} \times (0, \mathrm{T})$$

- Initial conditions (homogeneous)
- Implemented with MFEM using high-order finite elements and RK4 time-stepping

Bathymetry-adapted meshing



Cascadia margin-wide rupture scenario

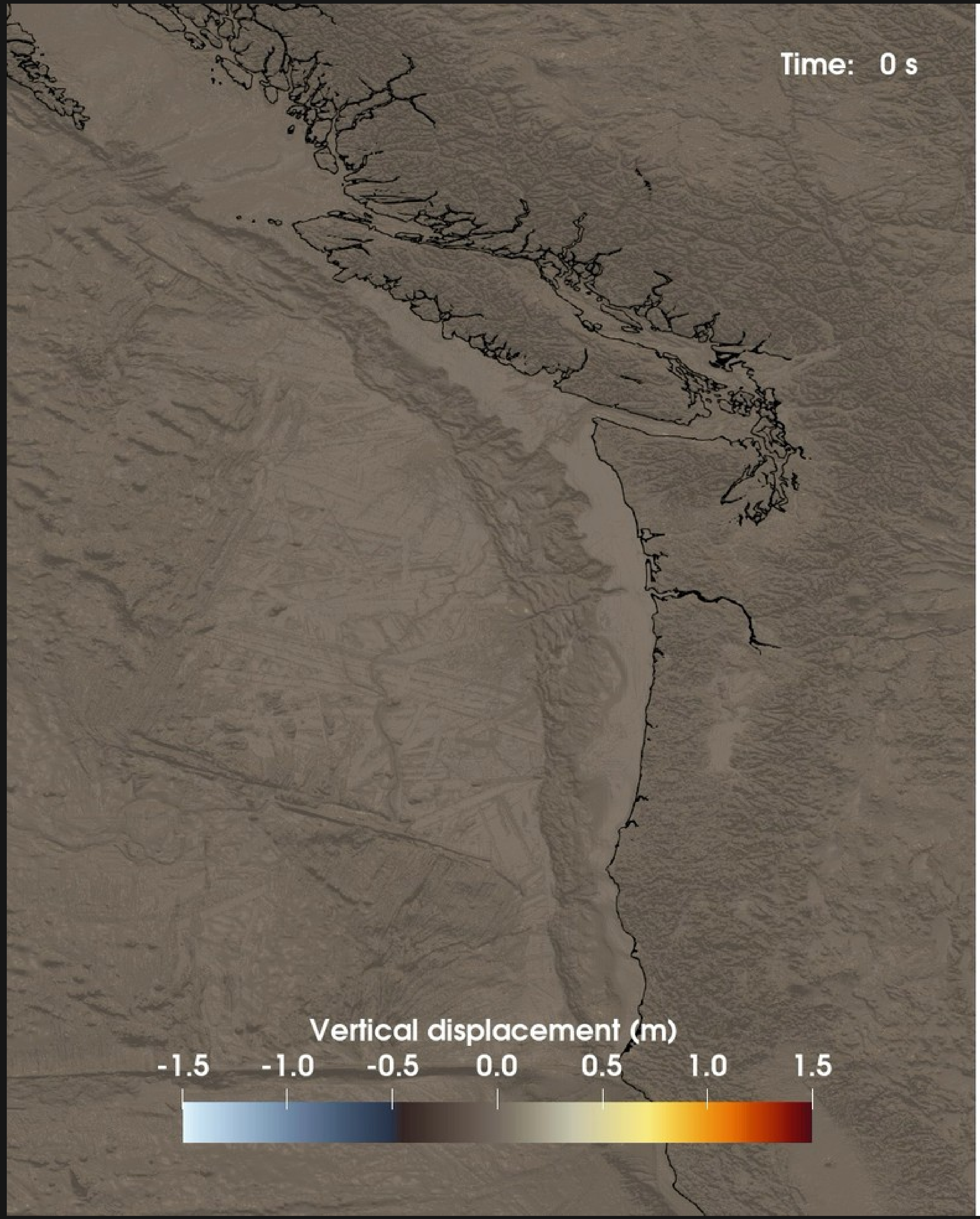
<u>Left:</u>

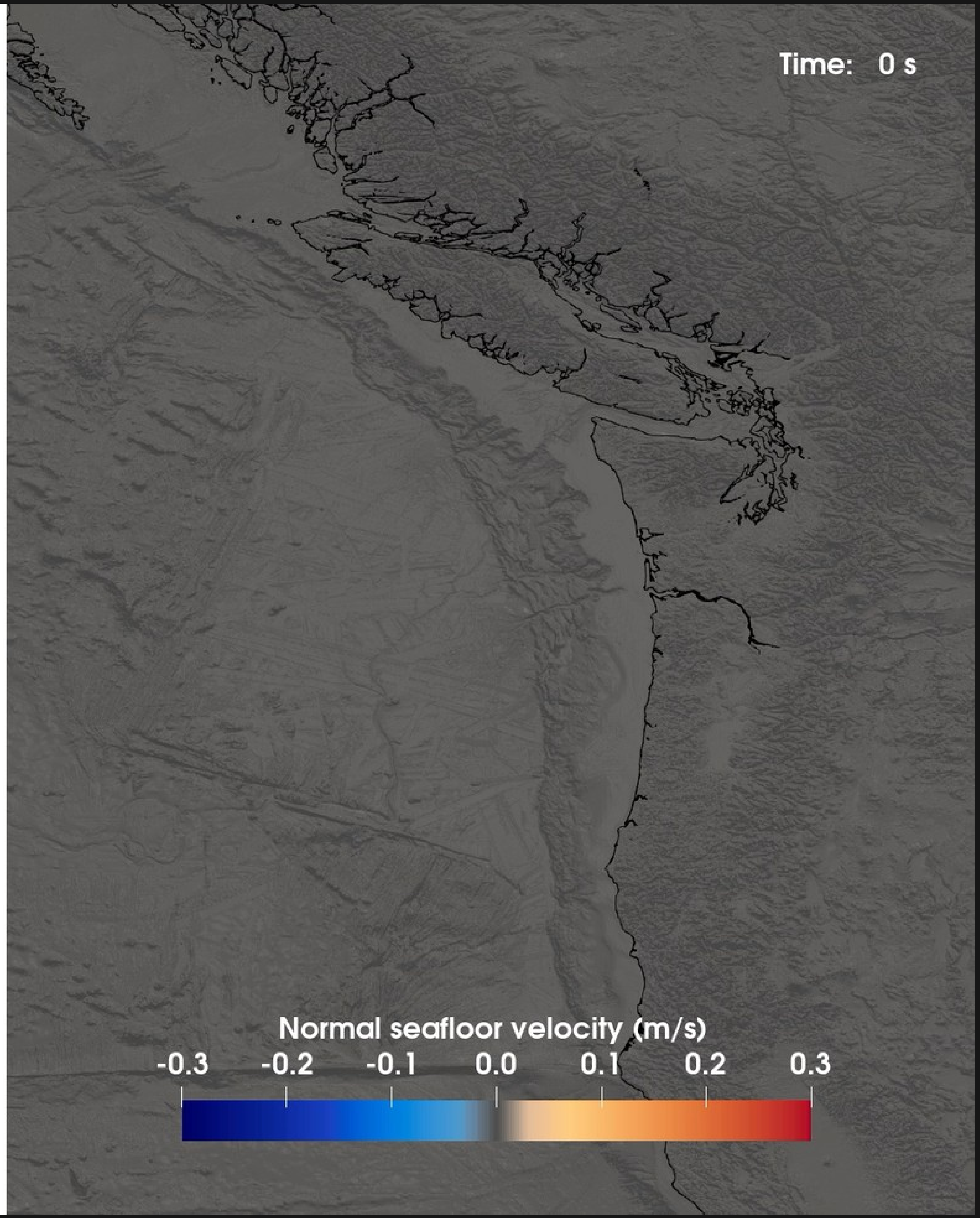
Vertical seafloor uplift

<u>Right:</u>

Seafloor normal velocity

Dynamic rupture earthquake scenarios of Cascadia earthquakes using geodetic locking models, Glehman et al., Seismica, 2025





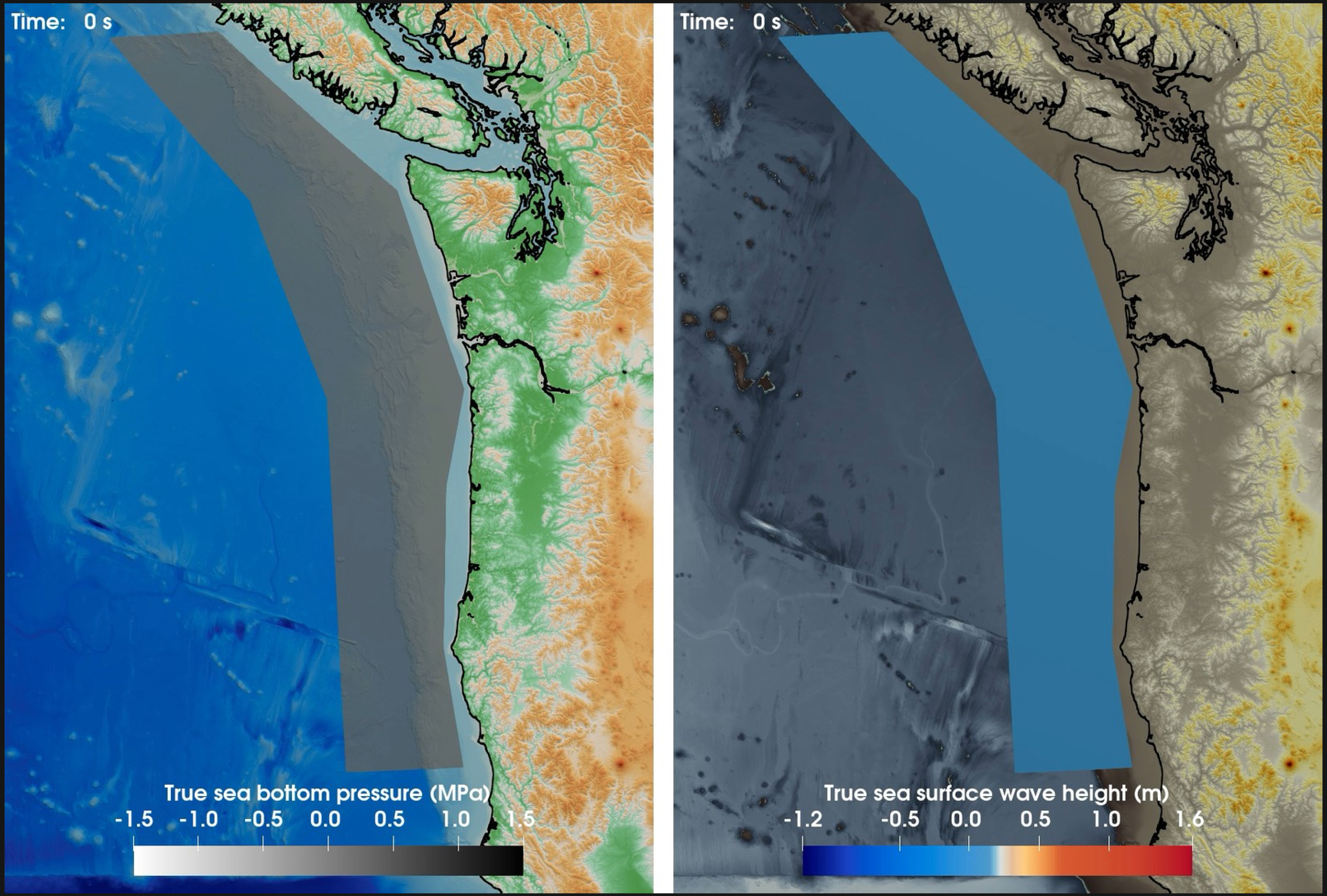
Acoustic–gravity simulation

<u>Left:</u>

Seafloor pressure change

<u>Right:</u>

Sea surface wave height



HPC systems

El Capitan:

- #1 Supercomputer (June '25 TOP500)
- 11,136 nodes with 44,544 AMD MI300A
- System peak: 2.75 ExaFLOP/s (dp)
- Sustained performance: 1.74 ExaFLOP/s (dp)
- System memory: 5.70 PB
- Vendor: HPE
- **We used 43,520 GPUs of El Capitan to do the offline phase the online inference then runs on a laptop**

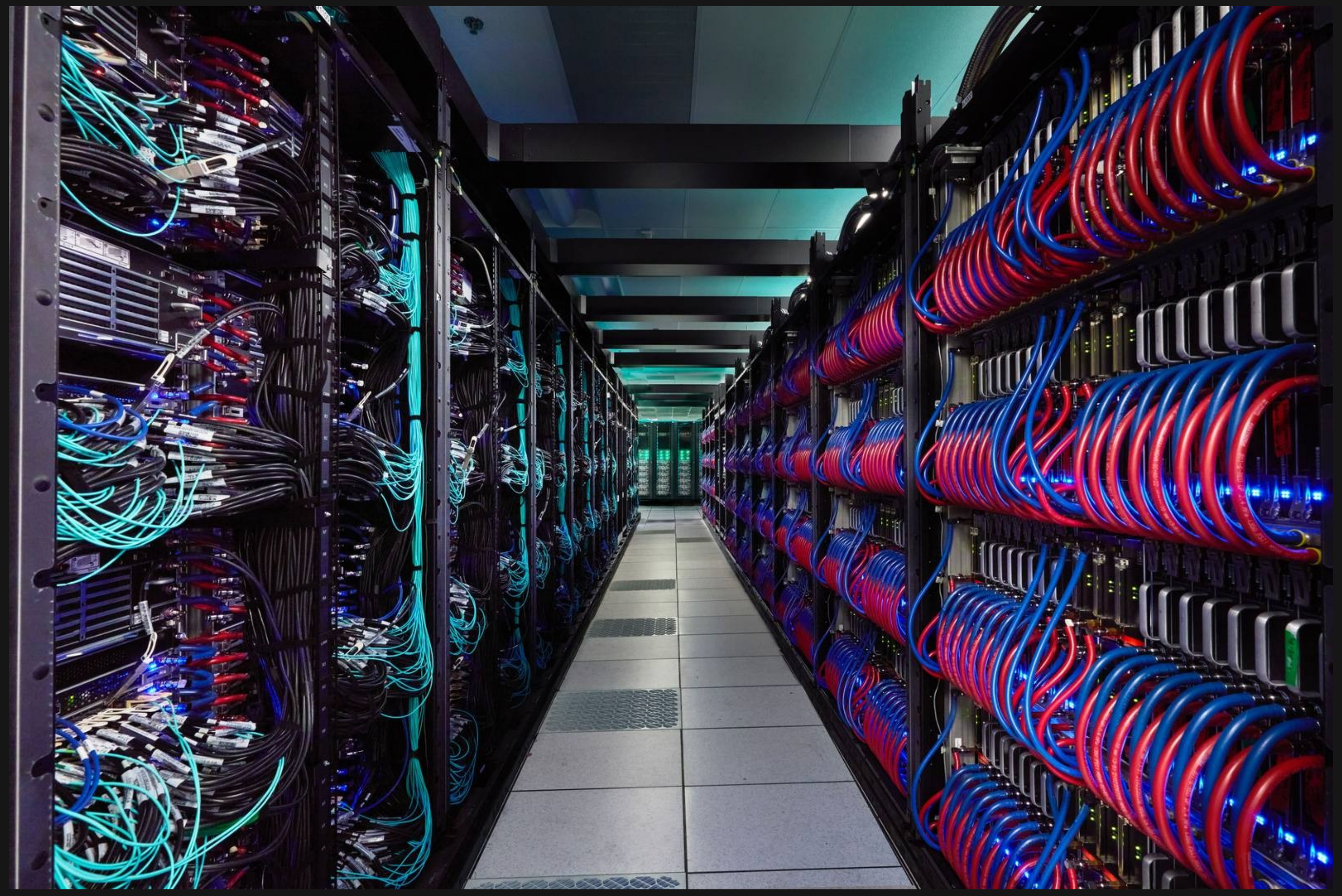


Image credit: Livermore Computing

HPC systems

Alps:
- #8 Supercomputer (June '25 TOP500)
- 2,688 nodes with 10,752 NVIDIA GH200
- System peak: 574.8 PetaFLOP/s (dp)
- Sustained: 434.9 PetaFLOP/s (dp)
- Vendor: HPE

Perlmutter:
- #25 Supercomputer (June '25 TOP500)
- 1,536 nodes with 6,144 NVIDIA A100
- System peak: 113.0 PetaFLOP/s (dp)
- Sustained: 79.2 PetaFLOP/s (dp)
- Vendor: HPE

Image credit: Swiss National Supercomputing Centre

Image credit: NERSC

Strong scaling

Weak scaling

GPU kernel optimizations

But how to solve the inverse problem?

<u>Autonomous dynamical system</u>

- Evolution of the physics model does not depend explicitly on the independent variable (e.g. time)
- Here: the mapping $m(\vec{x}, t+\tau) \mapsto d(\vec{x}, t+\tau)$ is the same as $m(\vec{x}, t) \mapsto d(\vec{x}, t)$
- The discrete p2o map F is **<u>shift-invariant</u>** with respect to time-stepping

Block Toeplitz structure, FFT-based solver

F is block Toeplitz ($N_t \times N_t$ blocks: $\mathrm{F}_{ij} \in \mathbb{R}^{N_d N_m}$, where $N_d \ll N_m$)

$$\begin{bmatrix} \mathrm{d}_1 \\ \mathrm{d}_2 \\ \mathrm{d}_3 \\ \vdots \\ \mathrm{d}_{N_t} \end{bmatrix} = \begin{bmatrix} \mathrm{F}_{11} & 0 & 0 & \cdots & 0 \\ \mathrm{F}_{21} & \mathrm{F}_{11} & 0 & \cdots & 0 \\ \mathrm{F}_{31} & \mathrm{F}_{21} & \mathrm{F}_{11} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \mathrm{F}_{N_t,1} & \mathrm{F}_{N_t-1,1} & \cdots & \mathrm{F}_{21} & \mathrm{F}_{11} \end{bmatrix} \begin{bmatrix} \mathrm{m}_1 \\ \mathrm{m}_2 \\ \mathrm{m}_3 \\ \vdots \\ \mathrm{m}_{N_t} \end{bmatrix}, \quad \begin{aligned} \mathrm{d}_i &= d(\vec{x}, t_i) \in \mathbb{R}^{N_d} \\ \mathrm{m}_i &= m(\vec{x}, t_i) \in \mathbb{R}^{N_m} \end{aligned}$$

- Precomputation of F by only N_d (number of sensors) adjoint PDE solutions
- Compact storage of F: $\mathcal{O}(N_m N_d N_t)$
- PDE-free application of F and F^* in Fourier space: $\mathcal{O}(N_m N_d N_t \log N_t)$
  Block Toeplitz matrix is block-diagonalized by FFT

Offline–online decomposition

Phase 1 (offline): Construct F and F_q maps from adjoint PDE solves

Phase 2 (offline): Compute compact representation of the Hessian

Apply Sherman–Morrison–Woodbury formula to reformulate Hessian in data space

$$\Gamma_{\text{post}} = \left(\mathrm{F}^*\Gamma_{\text{noise}}^{-1}\mathrm{F} + \Gamma_{\text{prior}}^{-1}\right)^{-1} = \Gamma_{\text{prior}}(\mathrm{I} - \mathrm{F}^*\underbrace{(\Gamma_{\text{noise}} + \mathrm{F}\Gamma_{\text{prior}}\mathrm{F}^*)}_{\mathrm{K}}{}^{-1}\mathrm{F}\Gamma_{\text{prior}})$$

Exploit that $\mathrm{G}^* \coloneqq \Gamma_{\text{prior}}\mathrm{F}^*$ is block Toeplitz

Phase 3 (offline): Compute uncertainties

Goal-oriented approach: Compute uncertainty in the quantities of interest (QoI)

$$\mathrm{q} \sim \mathcal{N}\left(\mathrm{F}_q \mathrm{m}_{\text{map}}, \mathrm{F}_q\Gamma_{\text{post}}\mathrm{F}_q^*\right), \qquad \mathrm{F}_q\Gamma_{\text{post}}\mathrm{F}_q^* = \mathrm{F}_q(\mathrm{I} - \mathrm{G}^*\mathrm{K}^{-1}\mathrm{F})\mathrm{G}_q^*$$

Phase 4 (online): Compute parameter and QoI MAP points in real time

Workflow (offline–online decomposition)

Cascadia margin-wide rupture inversion

Inverse problem configuration:

- Spatial resolution: 300 m
- Simulation time: T = 420 s
- Temporal parameter, data, and QoI dimension: N_t = 420 (1 Hz frequency)
- Spatial parameter dimension: N_d = 2,416,530
- Number of sensors: N_d = 600
- Number of QoI spatial locations: N_q = 21
- Total parameter dimension: $N_m N_t$ = 1,014,942,600
- Total data dimension: $N_d N_t$ = 252,000
- Total QoI dimension: $N_q N_t$ = 8,820
- Total state dimension: 3,744,141,189 spatial DOF; 336,000 RK4 timesteps

Compute times for each phase

Phase	Task	#GPUs	Compute time
1		512 512	
2		16 16 512 25	
3		512 512	
4		512 1	**< 0.2 seconds** **< 1 millisecond**

Phase	Task	#GPUs	Compute time
1	form $F : m \mapsto d$	512	600 x 52 m ~ 520 hours
	form $F_q : m \mapsto q$	512	21 x 52 m ~ 18 hours
2	form $G^* = \Gamma_{prior} F^*$	16	600 x 4.5 s ~ 45 minutes
	form $G_q^* = \Gamma_{prior} F_q^*$	16	21 x 4.5 s ~ 1.5 minutes
	form $K = \Gamma_{noise} + FG^*$	512	252,000 x 24 ms ~ 100 minutes
	factorize K	25	22 seconds
3	compute $\Gamma_{post(q)}$	512	8,820 x 150 ms ~ 25 minutes
	compute $Q: d \mapsto q$	512	8,820 x 150 ms ~ 25 minutes
4	**infer parameters m_{map}**	512	**< 0.2 seconds**
	predict QoI q_{map}	1	**< 1 millisecond**

Time-to-solution for the online computation is < 0.2 seconds!

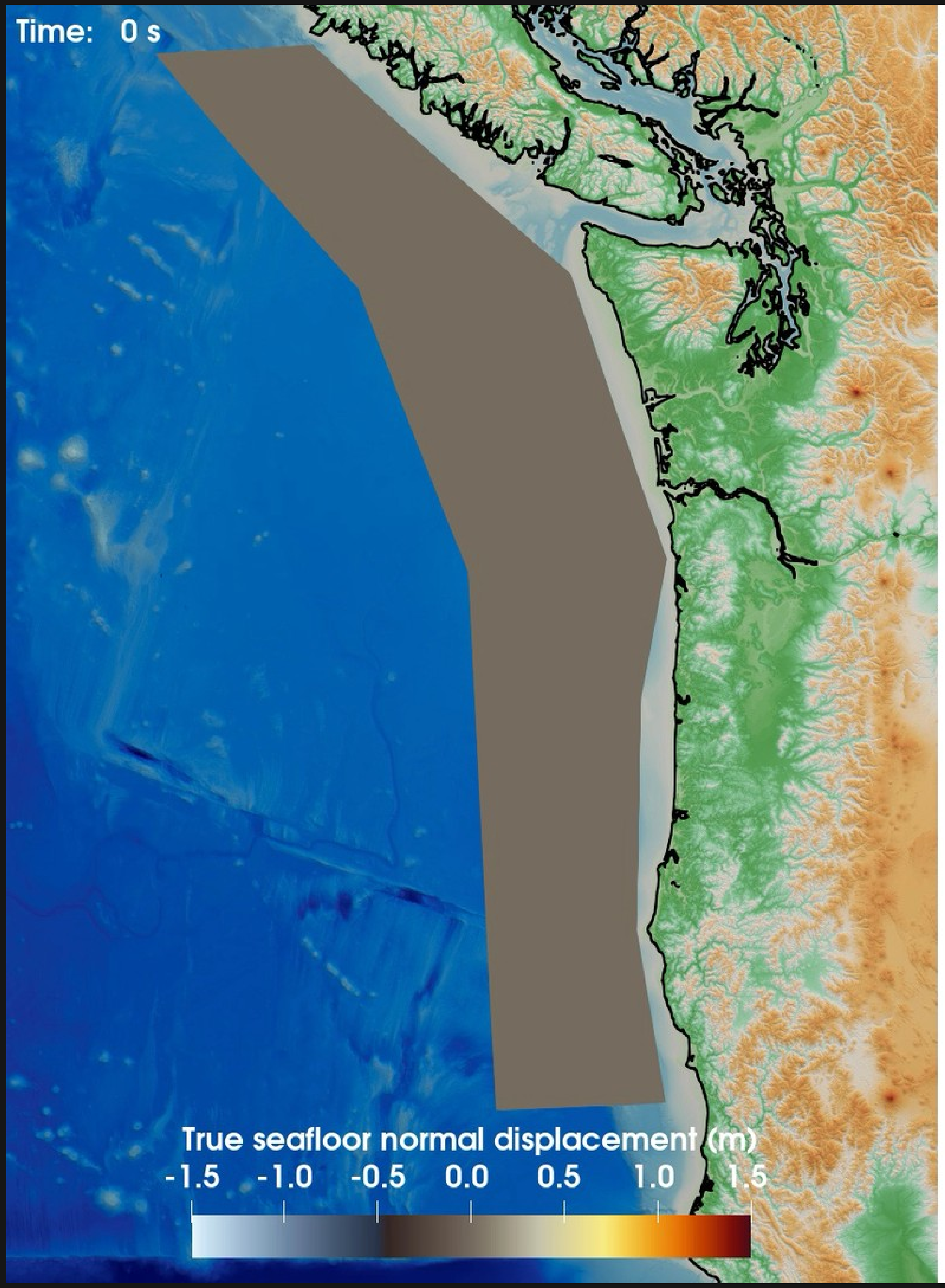
True vs. inferred spatiotemporal seafloor motion

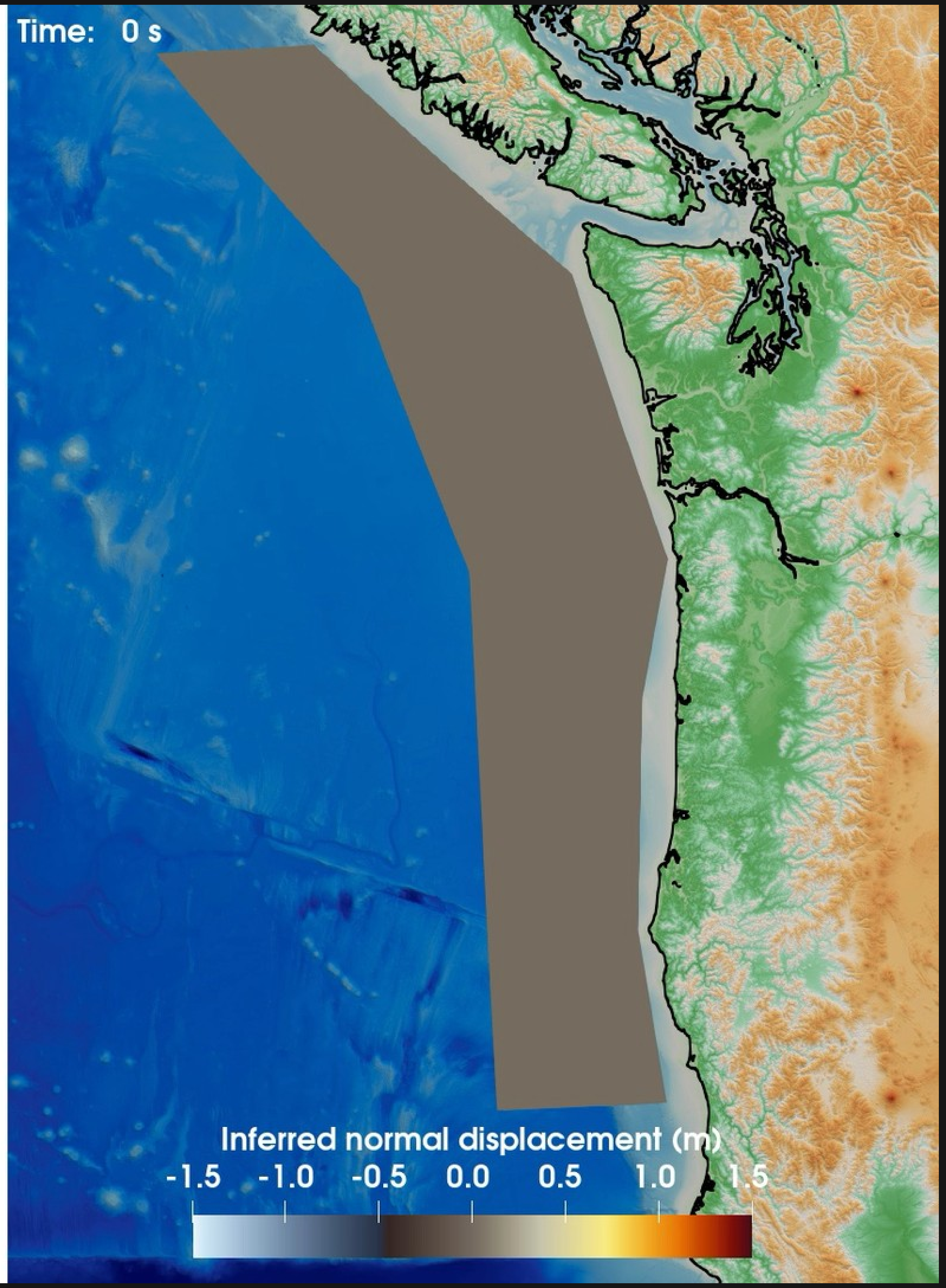
<u>Left:</u>

"True" seafloor normal uplift

<u>Middle:</u>

Mean of inferred seafloor normal uplift





600 hypothesized sensor locations

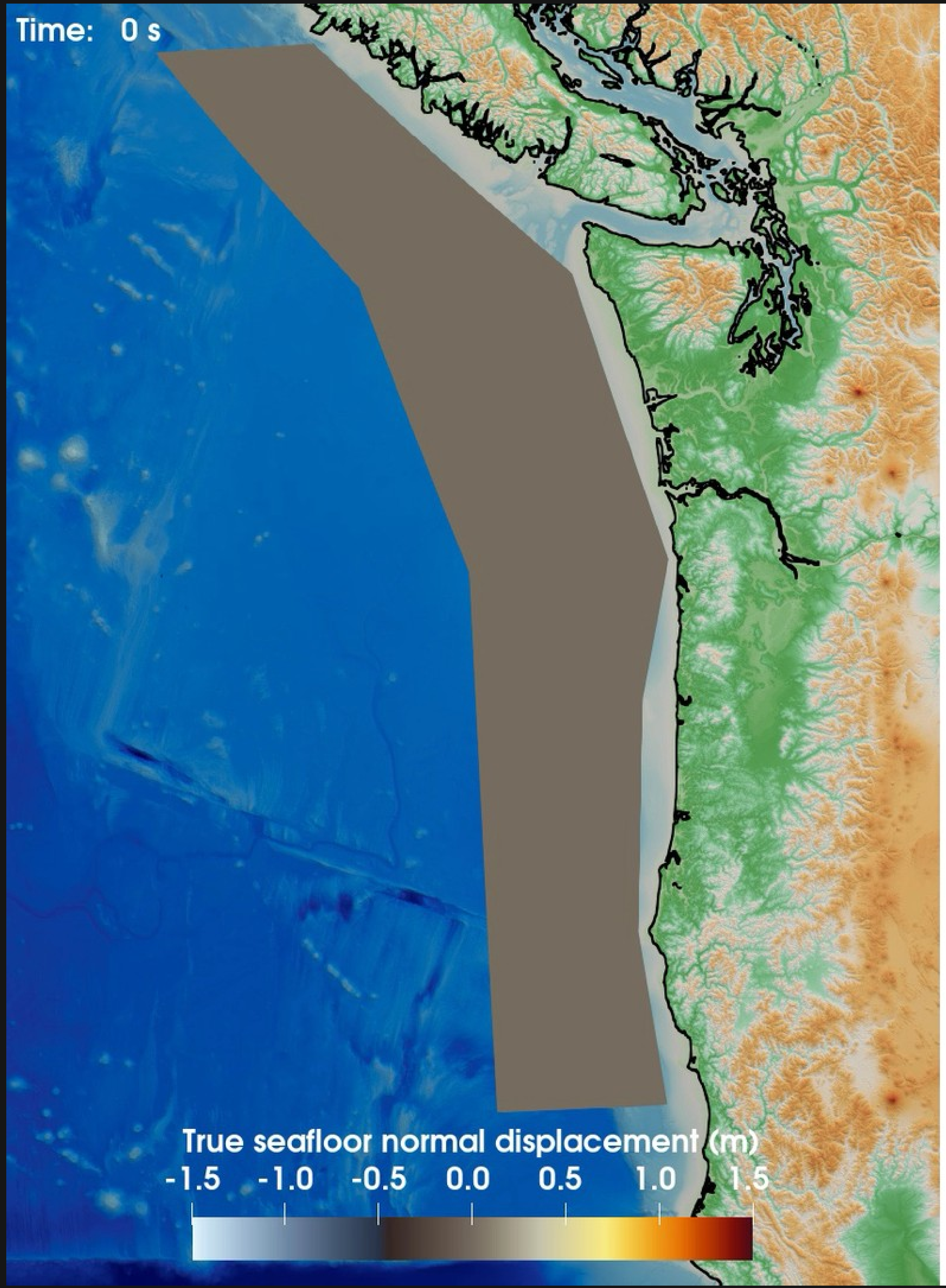
True vs. inferred spatiotemporal seafloor motion

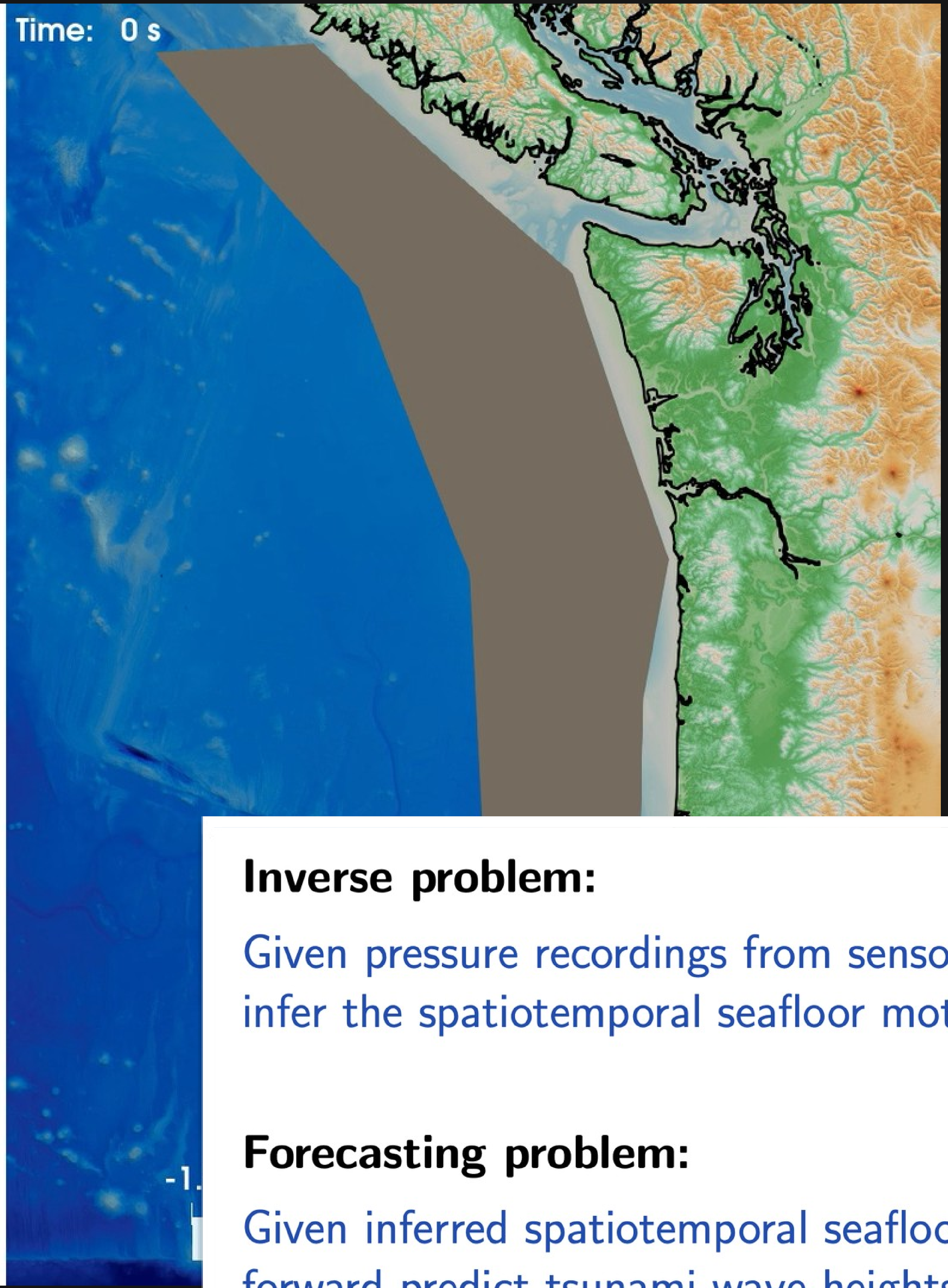
Left:

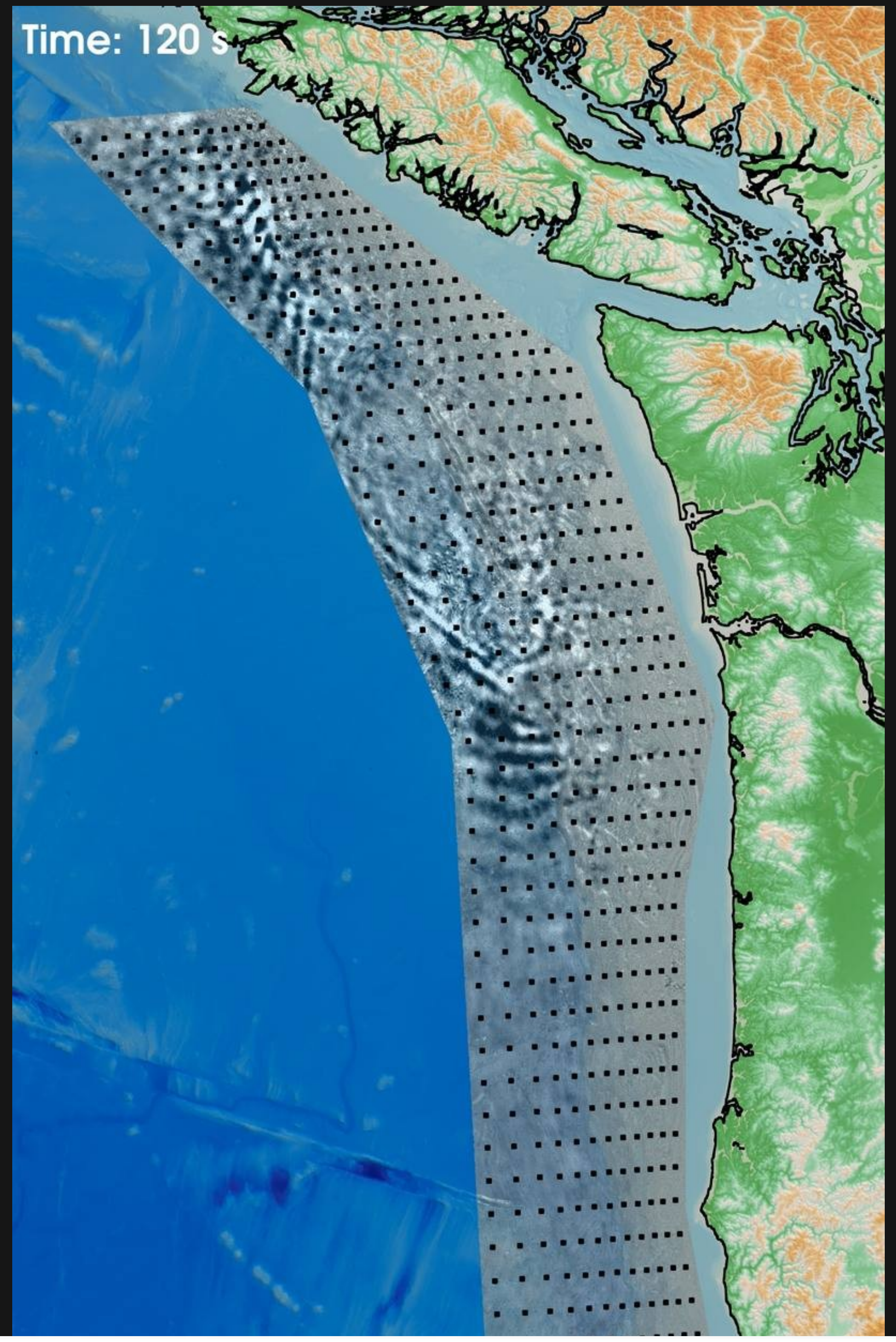
"True" seafloor normal uplift

Middle:

Mean of inferred seafloor normal uplift







Inverse problem:

Given pressure recordings from sensors on the seafloor,
infer the spatiotemporal seafloor motion in the subduction zone

Forecasting problem:

Given inferred spatiotemporal seafloor motion,
forward predict tsunami wave heights at specified coastal locations

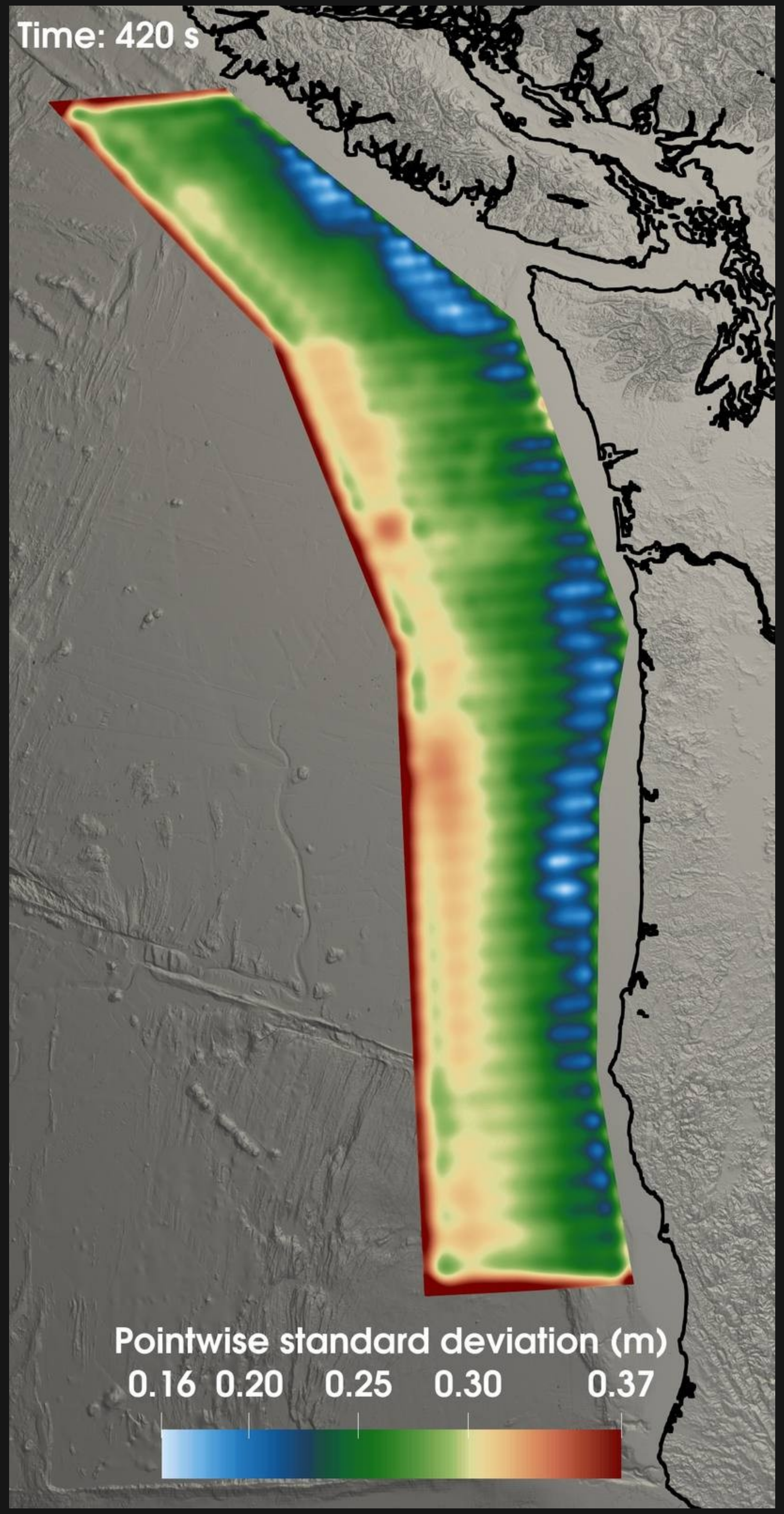
Sensor locations and parameter uncertainties

Left:

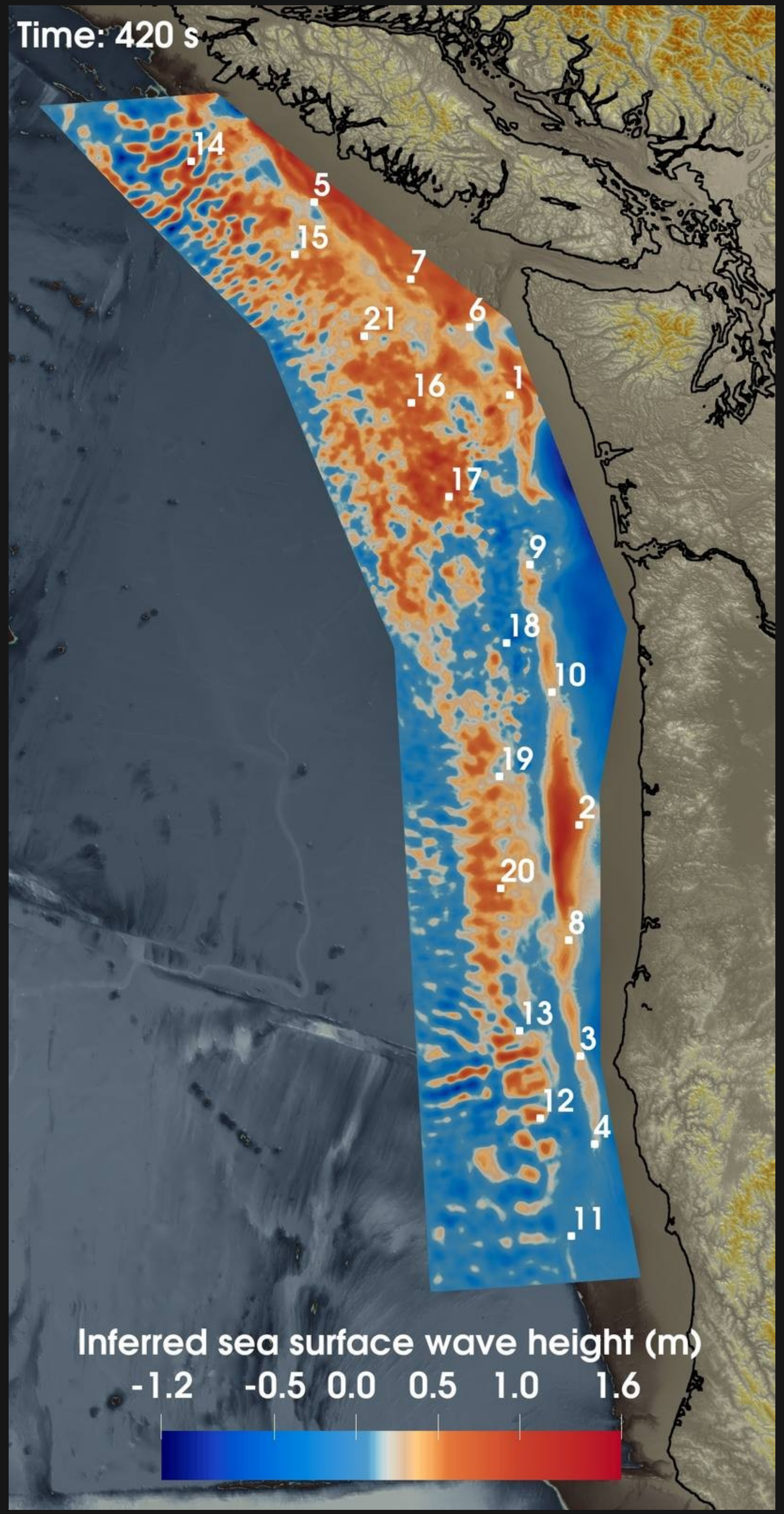
600 hypothesized sensor locations

Right:

Uncertainties illustrated as pointwise standard deviations of seafloor normal uplift



Real-time tsunami wave height forecasts with 95% credible intervals (CIs)



Highlights

- Fastest time-to-solution of a PDE-based Bayesian inverse problem with 1 billion parameters in 0.2 seconds
  Ten-billion-fold speedup for high-fidelity Bayesian-inversion-based tsunami early warning!
- Largest-to-date unstructured mesh finite element simulation with 55.5 trillion DOF on 43,520 GPUs
- 92% weak and 79% strong parallel efficiencies in scaling over a 128x increase of GPUs on the *El Capitan* system
- Performance-portable implementation; excellent scalability on *Alps*, *El Capitan*, *Frontera*, *and Perlmutter*

Literature references - I

S. Henneking, S. Venkat, V. Dobrev, J. Camier, T. Kolev, M. Fernando, A.-A. Gabriel, and O. Ghattas. “Real-time Bayesian inference at extreme scale: A digital twin for tsunami early warning applied to the Cascadia subduction zone.” In: *The International Conference for High Performance Computing, Networking, Storage and Analysis (SC ’25), November 16–21, 2025, St Louis, MO, USA. ACM, New York, NY, USA.*

S. Venkat, M. Fernando, S. Henneking, and O. Ghattas. “Fast and scalable FFT-based GPU-accelerated algorithms for block-triangular Toeplitz matrices with application to linear inverse problems governed by autonomous dynamical systems.” In: *SIAM Journal on Scientific Computing 47.5 (2025), B1201–B1226.*

S. Henneking, S. Venkat, and O. Ghattas. “Goal-oriented real-time Bayesian inference for linear autonomous dynamical systems with application to digital twins for tsunami early warning.” In: *Journal of Computational Physics 552 (2026), 114682.*

Literature references - II

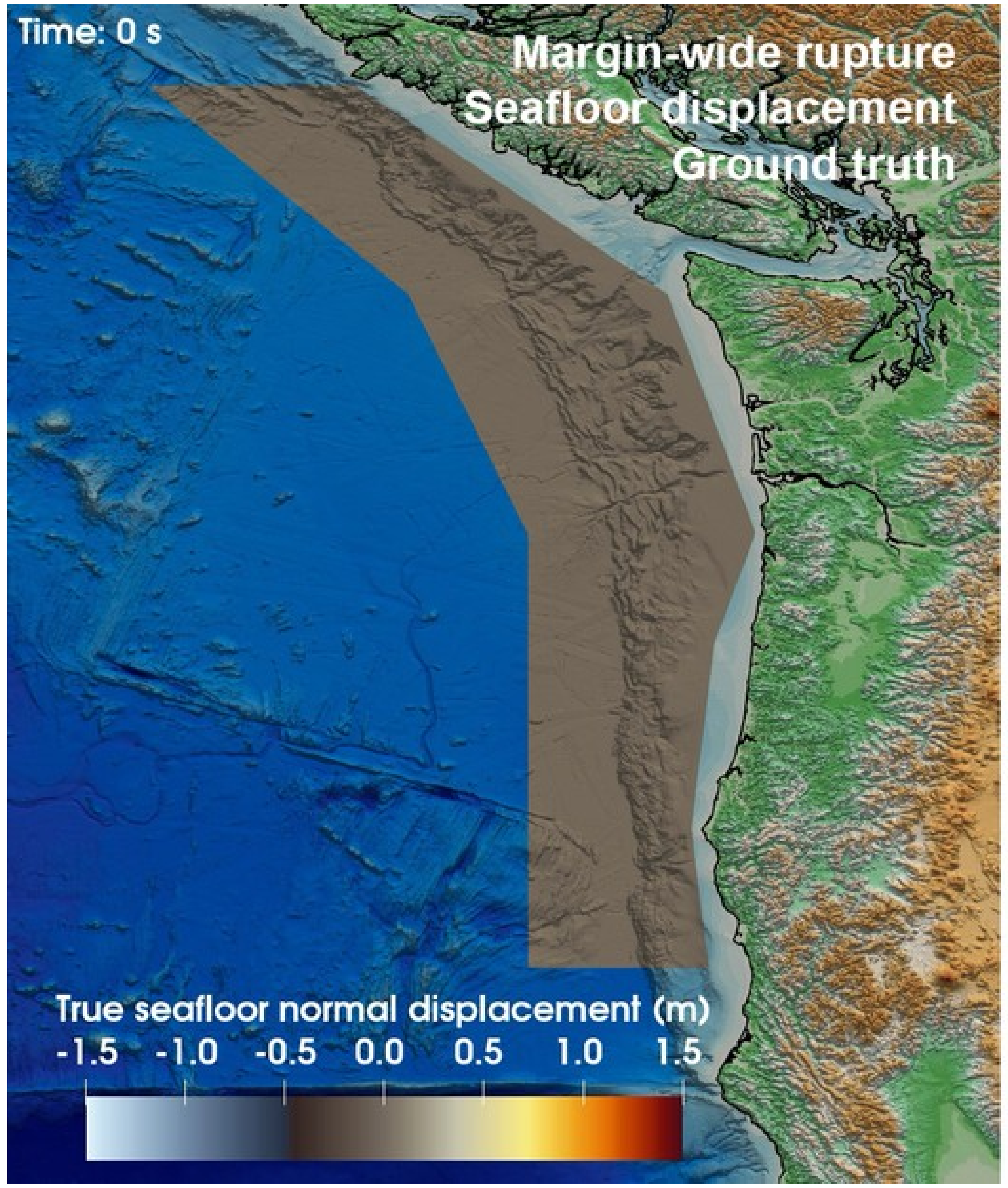
J. Tu, I. Karlin, J. Camier, V. Dobrev, T. Kolev, S. Henneking, O. Ghattas. “Accelerating high-order finite element simulations at extreme scale with FP64 tensor cores.” In: *arXiv preprint arXiv:2603.09038 (2026)*.

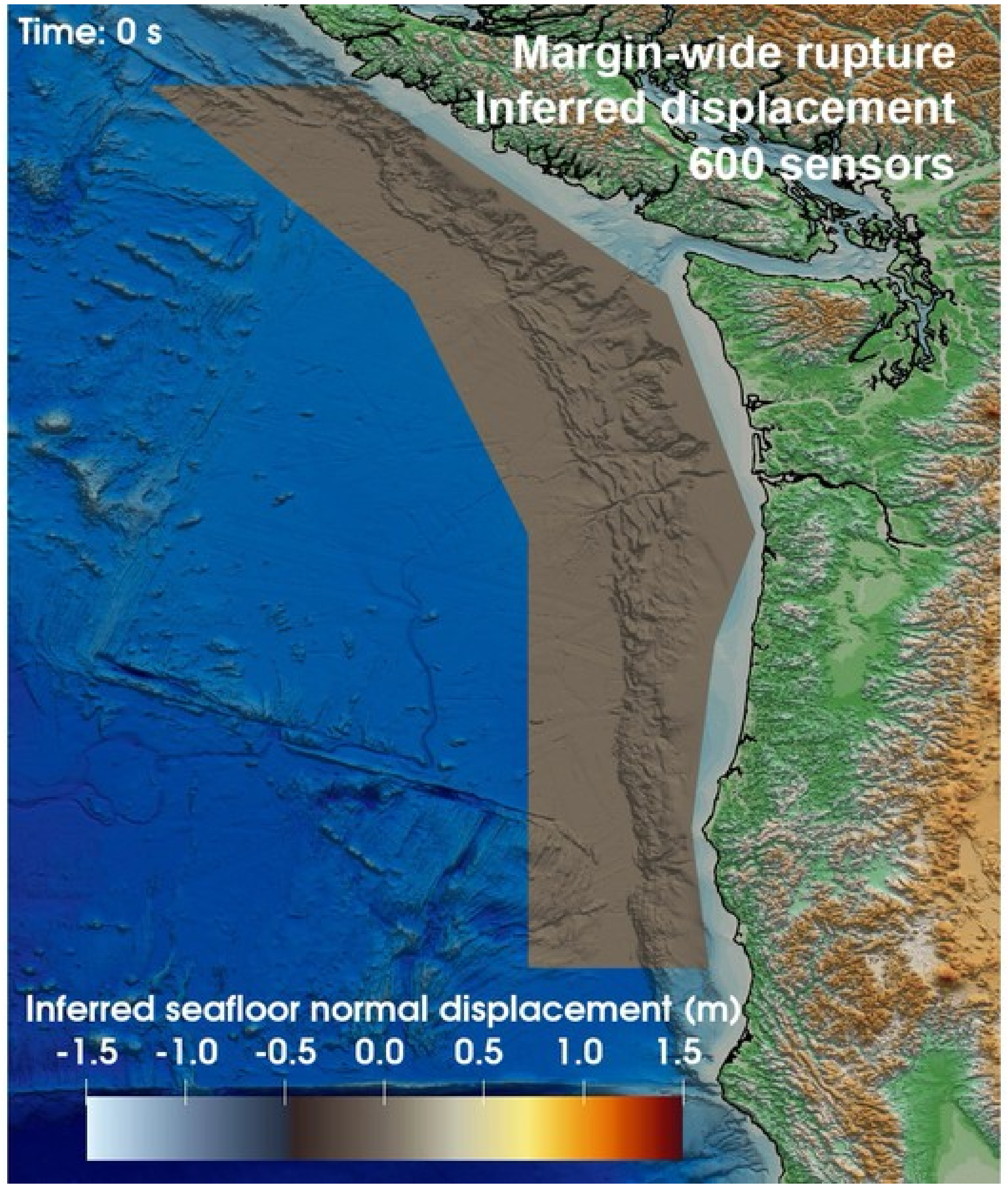
S. Henneking, F. Kutschera, S. Venkat, A.-A. Gabriel, O. Ghattas. “Real-time probabilistic tsunami forecasting in Cascadia from sparse offshore pressure observations.” In: *arXiv preprint:2603.14966 (2026)*.

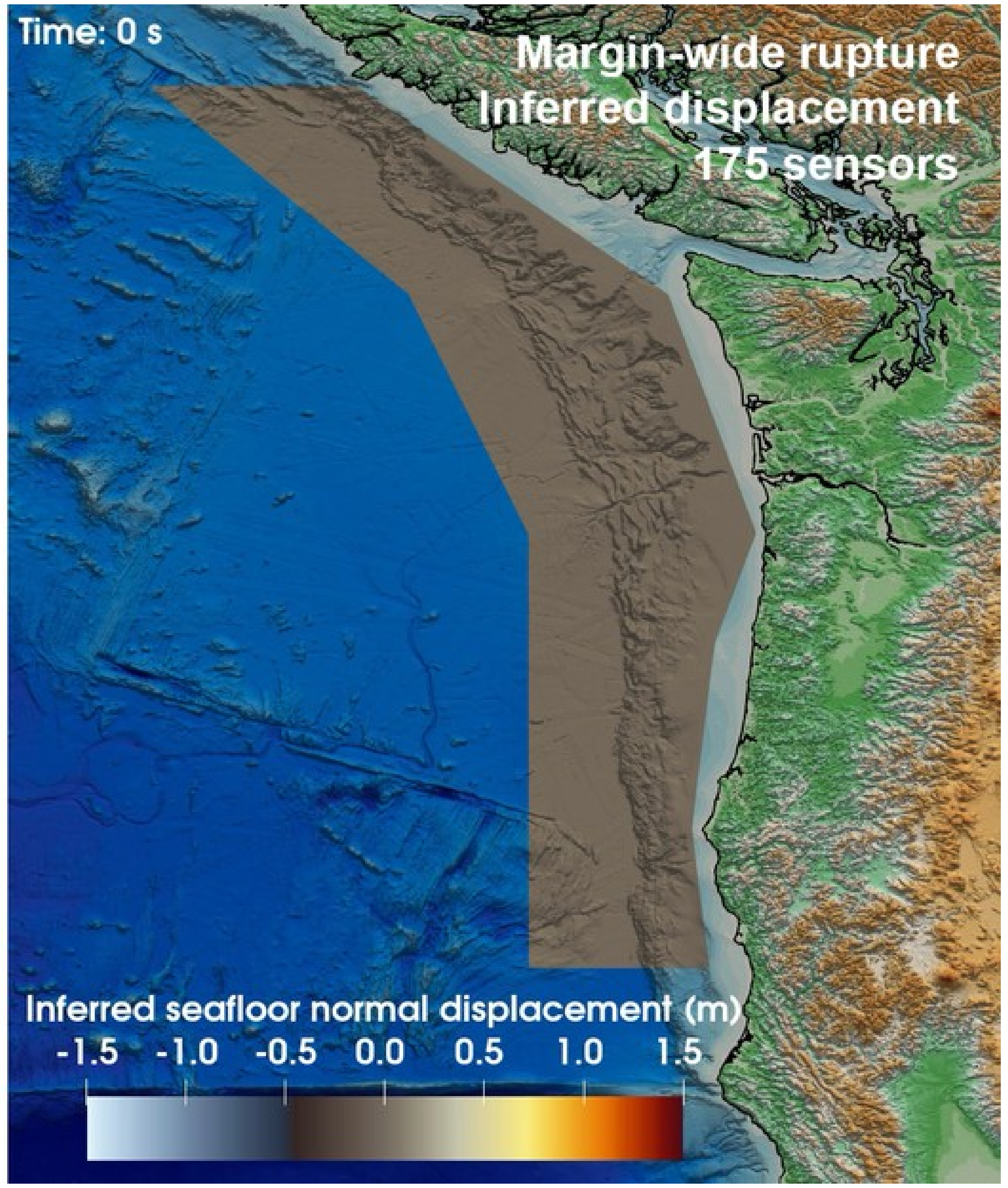
S. Venkat, S. Henneking, O. Ghattas. “Sensor placement for tsunami early warning via large-scale Bayesian optimal experimental design.” In: *arXiv preprint: 2604.08812 (2026)*.

Real-time probabilistic tsunami forecasting in Cascadia from sparse offshore pressure observations

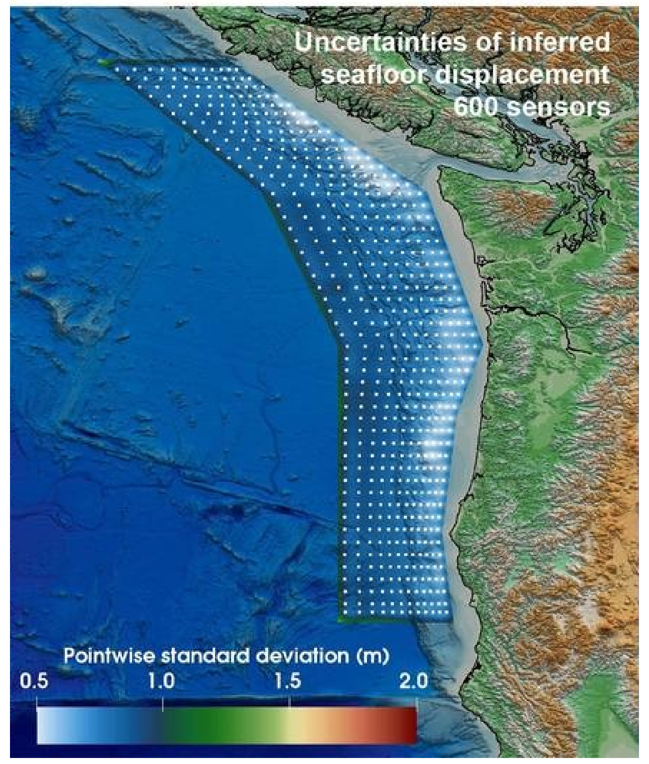
- Forecast errors **for sparse 175 sensor network** remain low at 22.1% for the margin-wide and 19.6% for the partial rupture



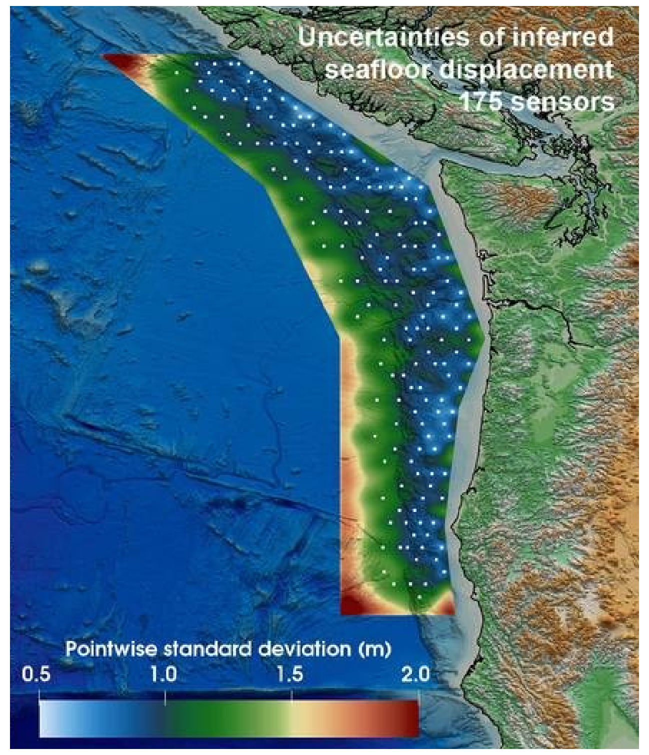




600 sensors



175 sensors



(Henneking et al., 2026,

Thank you!

Questions?

algabriel@ucsd.edu

www.alicegabriel.com

Henneking et al., Art in HPC, SC’