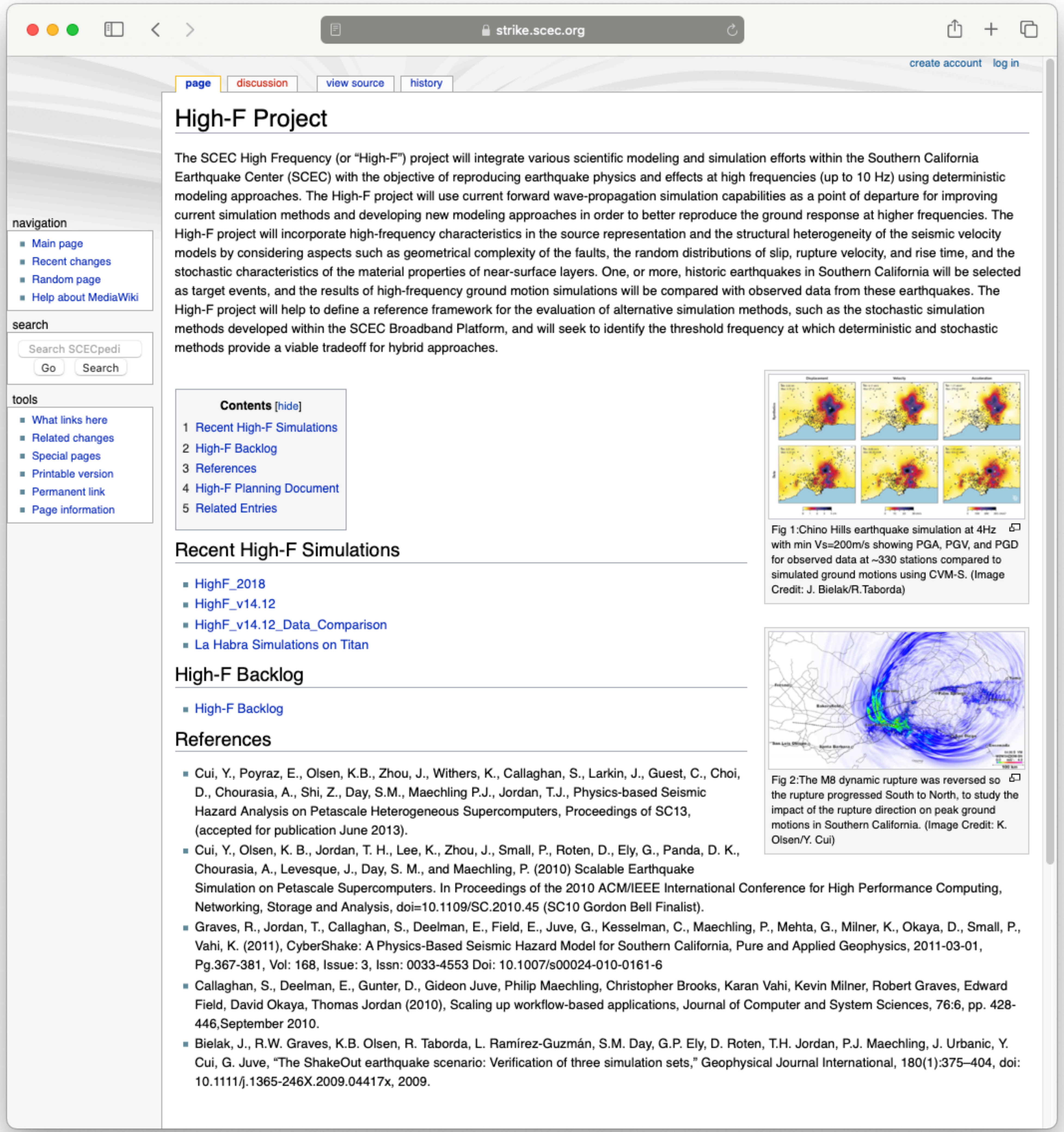
The Efficient Integration of Field Data into an Unstructured Software

SIAM GS21

Alex Breuer (alex.breuer@uni-jena.de)

High-Frequency Ground Motion Simulations



strike.scec.org

create account log in

page | discussion | view source | history

High-F Project

The SCEC High Frequency (or "High-F") project will integrate various scientific modeling and simulation efforts within the Southern California Earthquake Center (SCEC) with the objective of reproducing earthquake physics and effects at high frequencies (up to 10 Hz) using deterministic modeling approaches. The High-F project will use current forward wave-propagation simulation capabilities as a point of departure for improving current simulation methods and developing new modeling approaches in order to better reproduce the ground response at higher frequencies. The High-F project will incorporate high-frequency characteristics in the source representation and the structural heterogeneity of the seismic velocity models by considering aspects such as geometrical complexity of the faults, the random distributions of slip, rupture velocity, and rise time, and the stochastic characteristics of the material properties of near-surface layers. One, or more, historic earthquakes in Southern California will be selected as target events, and the results of high-frequency ground motion simulations will be compared with observed data from these earthquakes. The High-F project will help to define a reference framework for the evaluation of alternative simulation methods, such as the stochastic simulation methods developed within the SCEC Broadband Platform, and will seek to identify the threshold frequency at which deterministic and stochastic methods provide a viable tradeoff for hybrid approaches.

navigation
- Main page
- Recent changes
- Random page
- Help about MediaWiki

search
Search SCECpedi | Go | Search

tools
- What links here
- Related changes
- Special pages
- Printable version
- Permanent link
- Page information

Contents [hide]

Fig 1:Chino Hills earthquake simulation at 4Hz with min Vs=200m/s showing PGA, PGV, and PGD for observed data at ~330 stations compared to simulated ground motions using CVM-S. (Image Credit: J. Bielak/R.Taborda)

Recent High-F Simulations

- HighF_2018
- HighF_v14.12
- HighF_v14.12_Data_Comparison
- La Habra Simulations on Titan

High-F Backlog

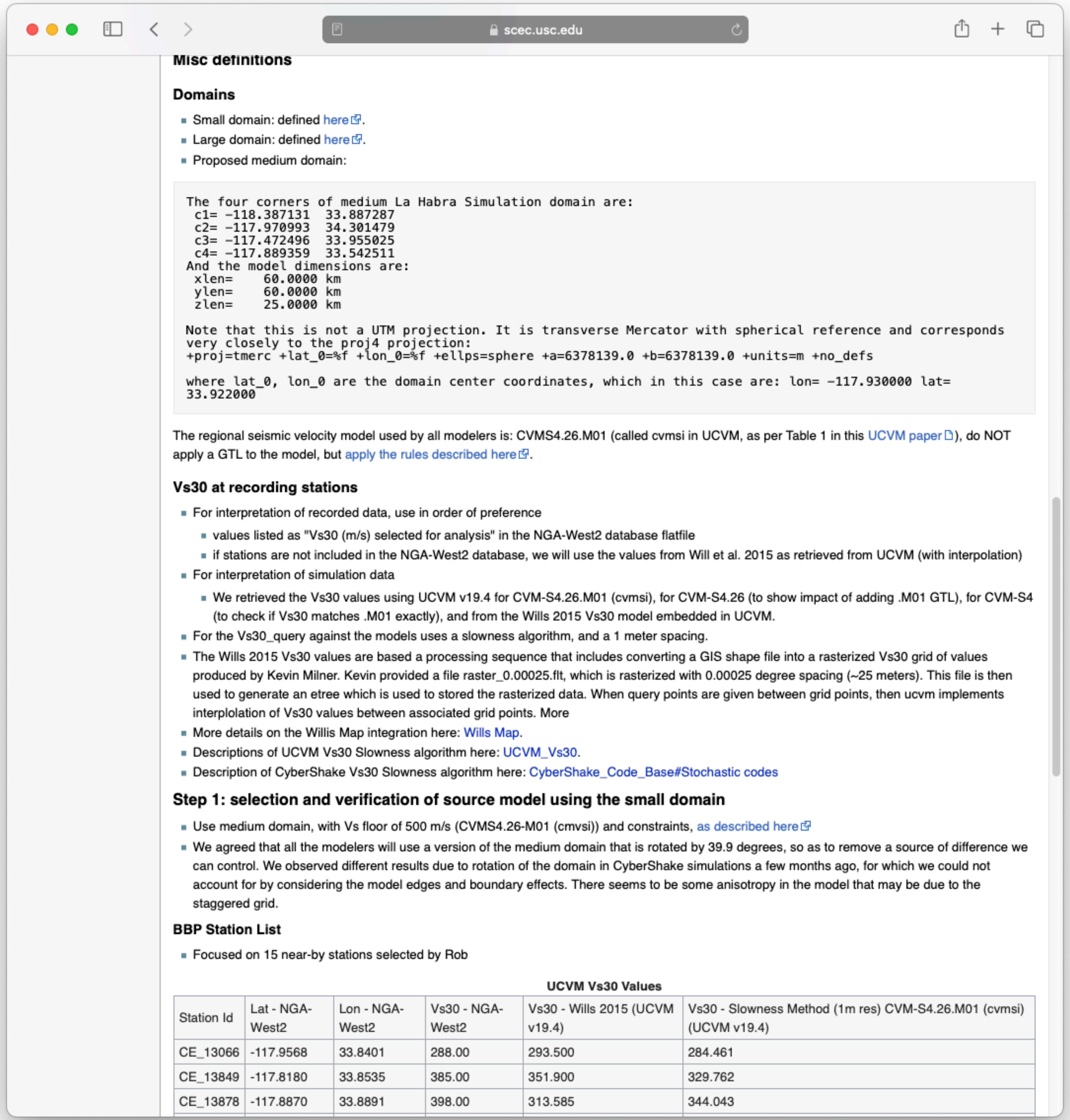
- High-F Backlog

Fig 2:The M8 dynamic rupture was reversed so the rupture progressed South to North, to study the impact of the rupture direction on peak ground motions in Southern California. (Image Credit: K. Olsen/Y. Cui)

References

- Cui, Y., Poyraz, E., Olsen, K.B., Zhou, J., Withers, K., Callaghan, S., Larkin, J., Guest, C., Choi, D., Chourasia, A., Shi, Z., Day, S.M., Maechling P.J., Jordan, T.J., Physics-based Seismic Hazard Analysis on Petascale Heterogeneous Supercomputers, Proceedings of SC13, (accepted for publication June 2013).
- Cui, Y., Olsen, K. B., Jordan, T. H., Lee, K., Zhou, J., Small, P., Roten, D., Ely, G., Panda, D. K., Chourasia, A., Levesque, J., Day, S. M., and Maechling, P. (2010) Scalable Earthquake Simulation on Petascale Supercomputers. In Proceedings of the 2010 ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis, doi=10.1109/SC.2010.45 (SC10 Gordon Bell Finalist).
- Graves, R., Jordan, T., Callaghan, S., Deelman, E., Field, E., Juve, G., Kesselman, C., Maechling, P., Mehta, G., Milner, K., Okaya, D., Small, P., Vahi, K. (2011), CyberShake: A Physics-Based Seismic Hazard Model for Southern California, Pure and Applied Geophysics, 2011-03-01, Pg.367-381, Vol: 168, Issue: 3, Issn: 0033-4553 Doi: 10.1007/s00024-010-0161-6
- Callaghan, S., Deelman, E., Gunter, D., Gideon Juve, Philip Maechling, Christopher Brooks, Karan Vahi, Kevin Milner, Robert Graves, Edward Field, David Okaya, Thomas Jordan (2010), Scaling up workflow-based applications, Journal of Computer and System Sciences, 76:6, pp. 428-446,September 2010.
- Bielak, J., R.W. Graves, K.B. Olsen, R. Taborda, L. Ramírez-Guzmán, S.M. Day, G.P. Ely, D. Roten, T.H. Jordan, P.J. Maechling, J. Urbanic, Y. Cui, G. Juve, "The ShakeOut earthquake scenario: Verification of three simulation sets," Geophysical Journal International, 180(1):375–404, doi: 10.1111/j.1365-246X.2009.04417x, 2009.

Source: https://strike.scec.org/scecpedia/High-F_Project



scec.usc.edu

Misc definitions

Domains

- Small domain: defined here.
- Large domain: defined here.
- Proposed medium domain:

```
The four corners of medium La Habra Simulation domain are:
 c1= -118.387131  33.887287
 c2= -117.970993  34.301479
 c3= -117.472496  33.955025
 c4= -117.889359  33.542511
And the model dimensions are:
 xlen=    60.0000 km
 ylen=    60.0000 km
 zlen=    25.0000 km

Note that this is not a UTM projection. It is transverse Mercator with spherical reference and corresponds
very closely to the proj4 projection:
+proj=tmerc +lat_0=%f +lon_0=%f +ellps=sphere +a=6378139.0 +b=6378139.0 +units=m +no_defs

where lat_0, lon_0 are the domain center coordinates, which in this case are: lon= -117.930000 lat=
33.922000
```

The regional seismic velocity model used by all modelers is: CVMS4.26.M01 (called cvmsi in UCVM, as per Table 1 in this UCVM paper), do NOT apply a GTL to the model, but apply the rules described here.

Vs30 at recording stations

- For interpretation of recorded data, use in order of preference
  - values listed as "Vs30 (m/s) selected for analysis" in the NGA-West2 database flatfile
  - if stations are not included in the NGA-West2 database, we will use the values from Will et al. 2015 as retrieved from UCVM (with interpolation)
- For interpretation of simulation data
  - We retrieved the Vs30 values using UCVM v19.4 for CVM-S4.26.M01 (cvmsi), for CVM-S4.26 (to show impact of adding .M01 GTL), for CVM-S4 (to check if Vs30 matches .M01 exactly), and from the Wills 2015 Vs30 model embedded in UCVM.
- For the Vs30_query against the models uses a slowness algorithm, and a 1 meter spacing.
- The Wills 2015 Vs30 values are based a processing sequence that includes converting a GIS shape file into a rasterized Vs30 grid of values produced by Kevin Milner. Kevin provided a file raster_0.00025.flt, which is rasterized with 0.00025 degree spacing (~25 meters). This file is then used to generate an etree which is used to stored the rasterized data. When query points are given between grid points, then ucvm implements interpolation of Vs30 values between associated grid points. More
- More details on the Willis Map integration here: Wills Map.
- Descriptions of UCVM Vs30 Slowness algorithm here: UCVM_Vs30.
- Description of CyberShake Vs30 Slowness algorithm here: CyberShake_Code_Base#Stochastic codes

Step 1: selection and verification of source model using the small domain

- Use medium domain, with Vs floor of 500 m/s (CVMS4.26-M01 (cmvsi)) and constraints, as described here
- We agreed that all the modelers will use a version of the medium domain that is rotated by 39.9 degrees, so as to remove a source of difference we can control. We observed different results due to rotation of the domain in CyberShake simulations a few months ago, for which we could not account for by considering the model edges and boundary effects. There seems to be some anisotropy in the model that may be due to the staggered grid.

BBP Station List

- Focused on 15 near-by stations selected by Rob

UCVM Vs30 Values

Station Id	Lat - NGA-West2	Lon - NGA-West2	Vs30 - NGA-West2	Vs30 - Wills 2015 (UCVM v19.4)	Vs30 - Slowness Method (1m res) CVM-S4.26.M01 (cvmsi) (UCVM v19.4)
CE_13066	-117.9568	33.8401	288.00	293.500	284.461
CE_13849	-117.8180	33.8535	385.00	351.900	329.762
CE_13878	-117.8870	33.8891	398.00	313.585	344.043

Source: https://scec.usc.edu/scecpedia/HighF_La_Habra_Verification

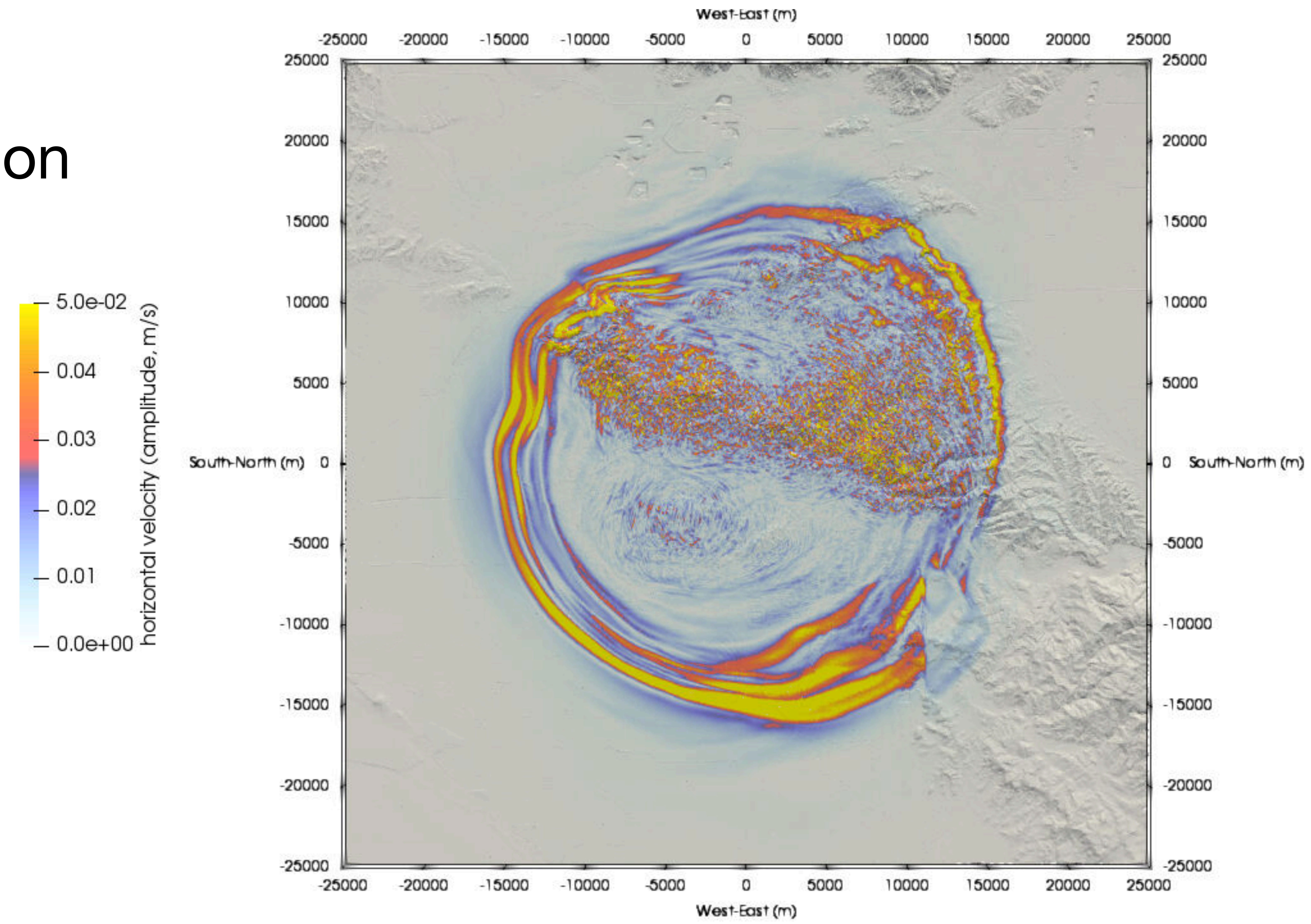
Study area of the Southern California Earthquake Center's High Frequency project simulating the 2014 M5.1 La Habra, California earthquake. Shown are the "small domain" through the inner box and partially the "large domain" through the outer box. Additionally the locations of the epicenter and three stations are given.

You are able interact with this map at:
http://umap.openstreetmap.fr/en/map/high-f_560152

High-F Project

Ground Motion Simulation

Visualization of the seismic wave field for a simulation of the 2014 Mw 5.1 La Habra Earthquake. Shown are the amplitudes of the horizontal particle velocities after seven seconds of simulated time.



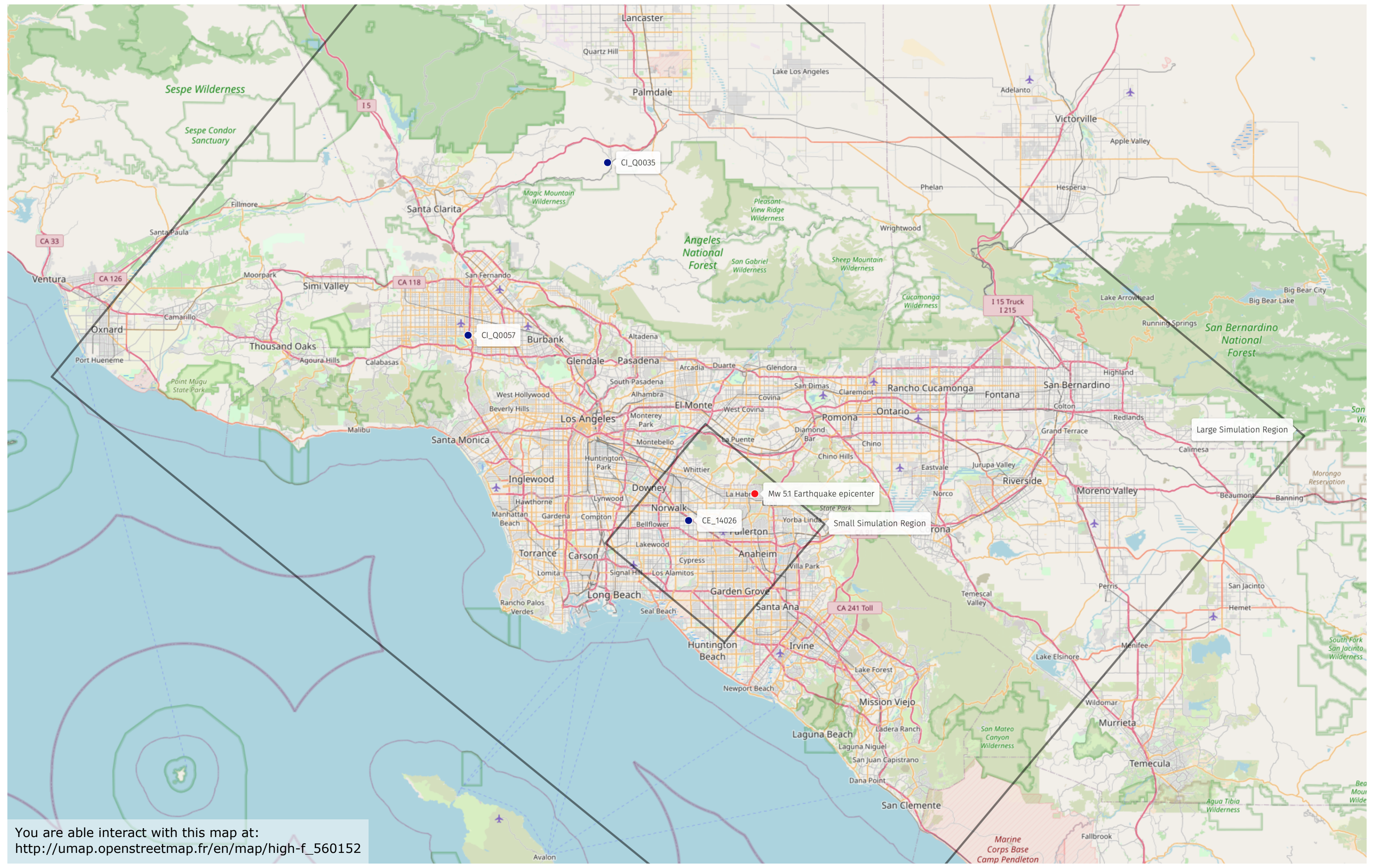
FRIEDRICH-SCHILLER-UNIVERSITÄT JENA

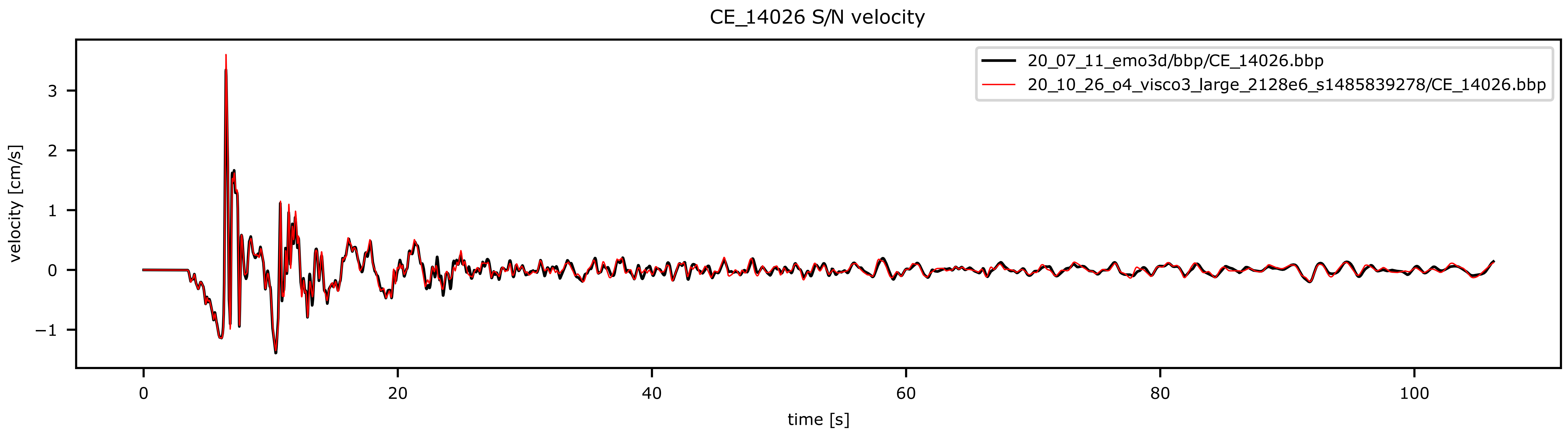
First fully-featured earthquake simulation which ran on Ampere's new Altra processor in the Oracle Cloud Infrastructure. Our results show that we are not only able to fully utilize the Armv8.2 cores' vector pipelines but also all of the 160 cores offered by a single dual-socket cloud instance. A brief overview of this work is featured by Oracle Live as part of the "Advancing the Future of Cloud with Arm-Based Computing" event.

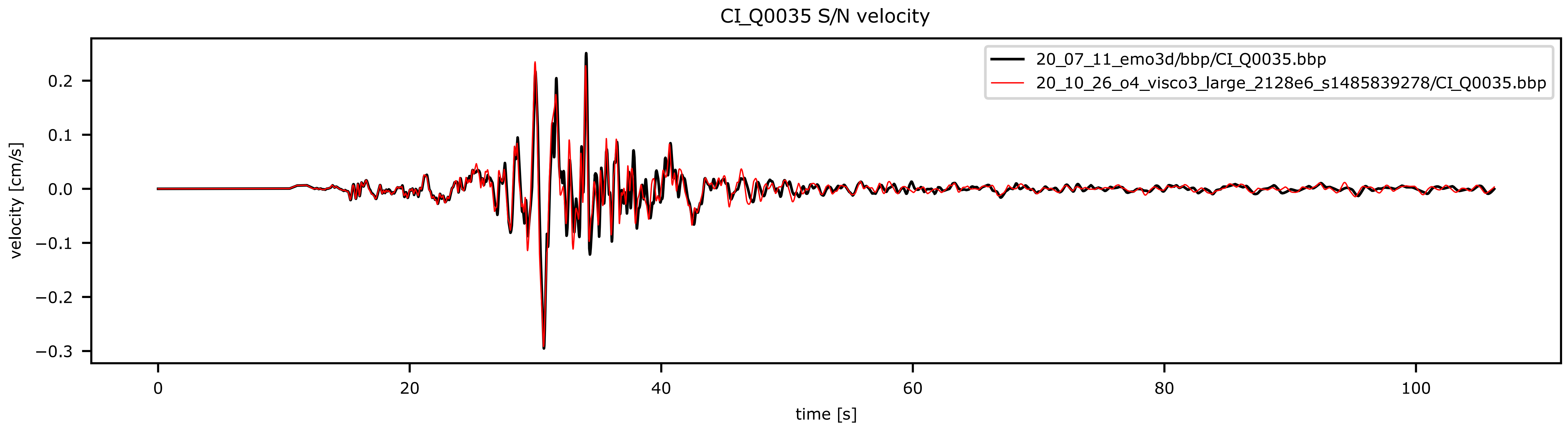
Verification

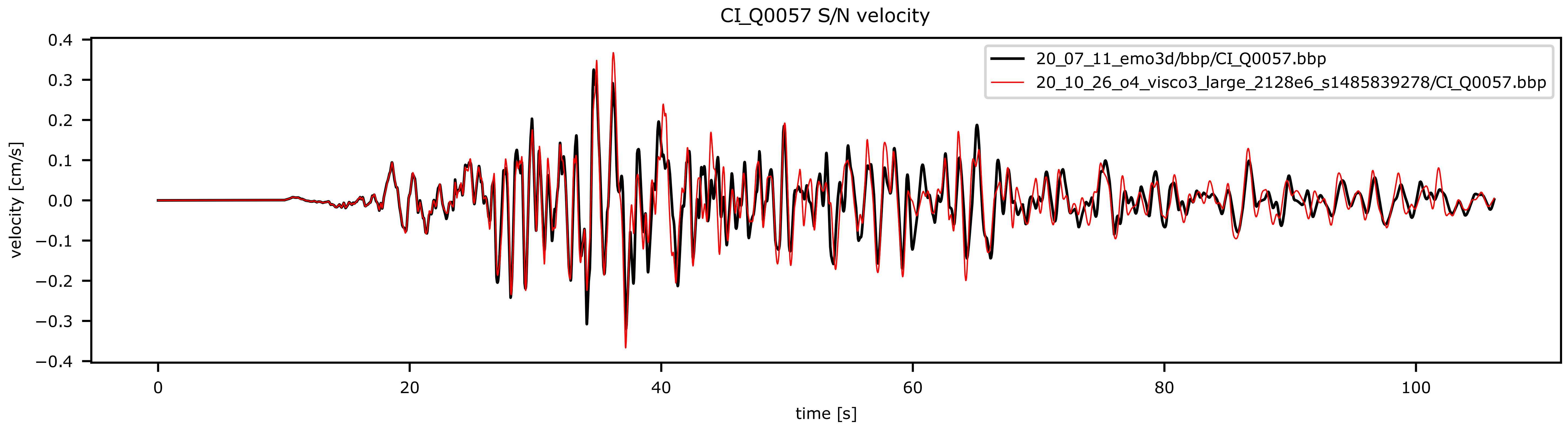
Comparing Synthetics

Right figure: Comparison of EDGE's South-North velocity component (red) to another solver of the High Frequency project (black). Shown are synthetic seismograms for the three stations depicted in the figure below. The seismograms were low-pass filtered at 5Hz. EDGE's respective ground motion simulation harnessed 1,536 nodes of the Frontera machine for a total of 48 hours to advance the used 2.1 billion tetrahedral element mesh in time.





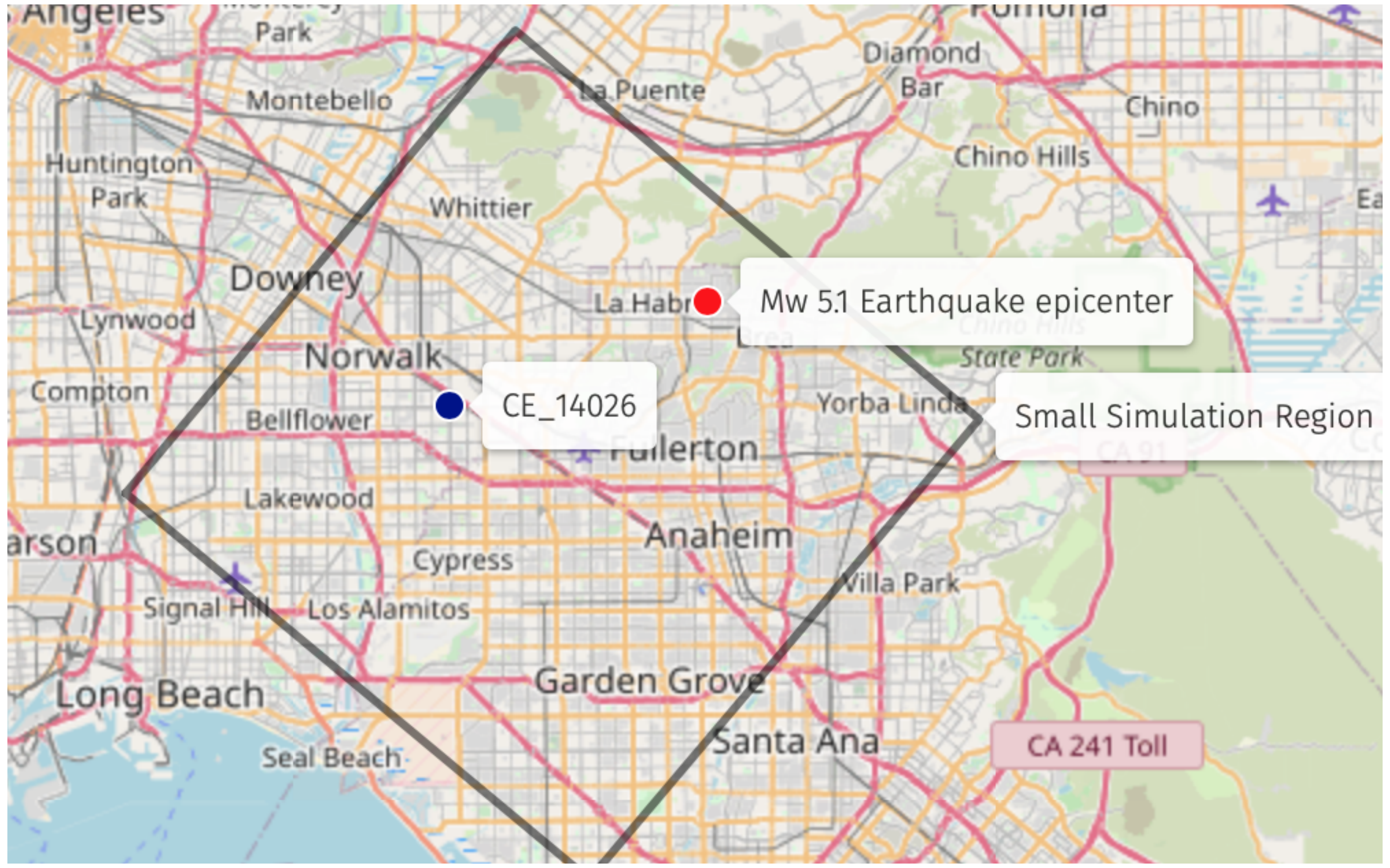


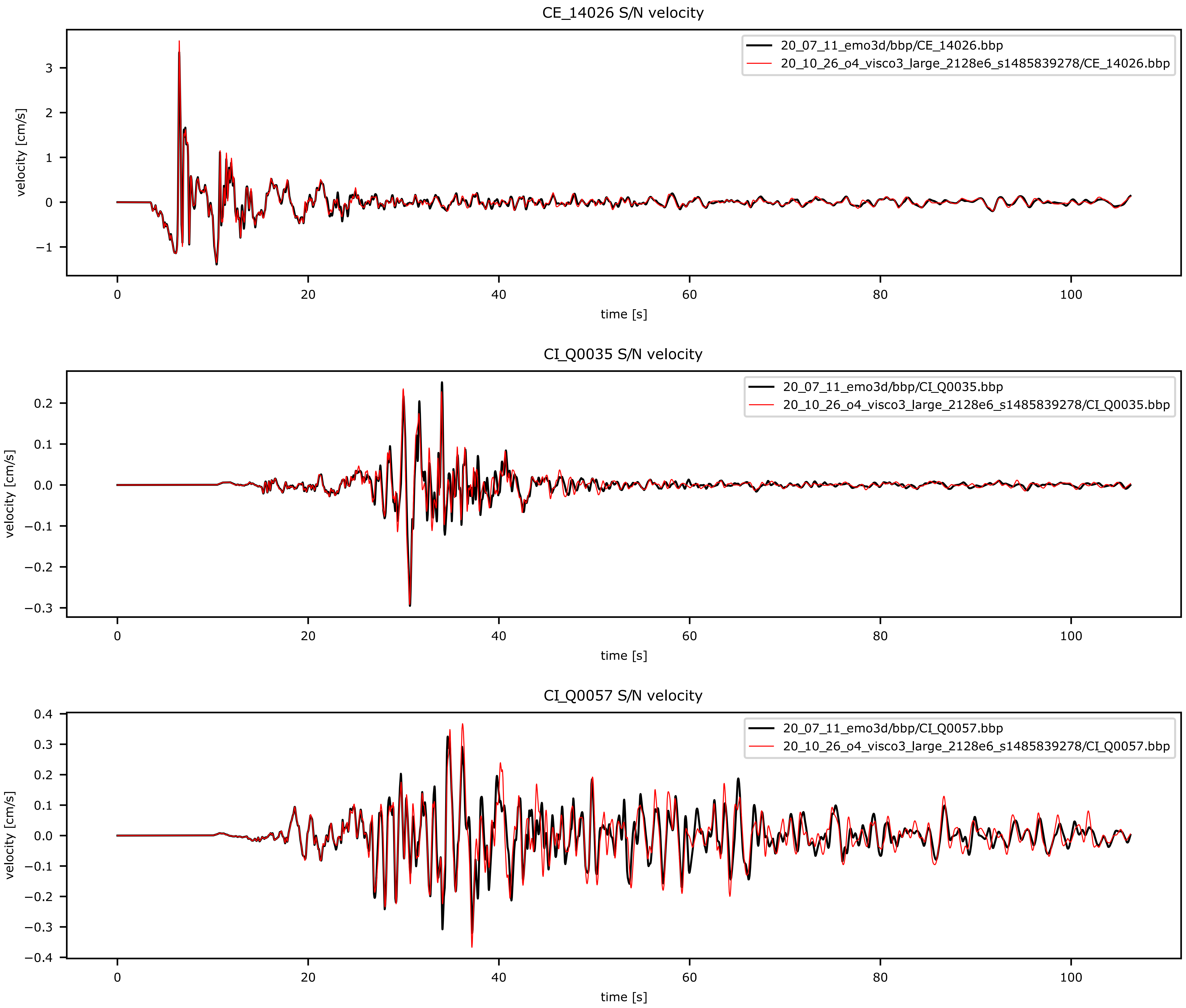


Verification

Comparing Synthetics

Right figure: Comparison of EDGE's South-North velocity component (red) to another solver of the High Frequency project (black). Shown are synthetic seismograms for the three stations depicted in the figure below. The seismograms were low-pass filtered at 5Hz. EDGE's respective ground motion simulation harnessed 1,536 nodes of the Frontera machine for a total of 48 hours to advance the used 2.1 billion tetrahedral element mesh in time.



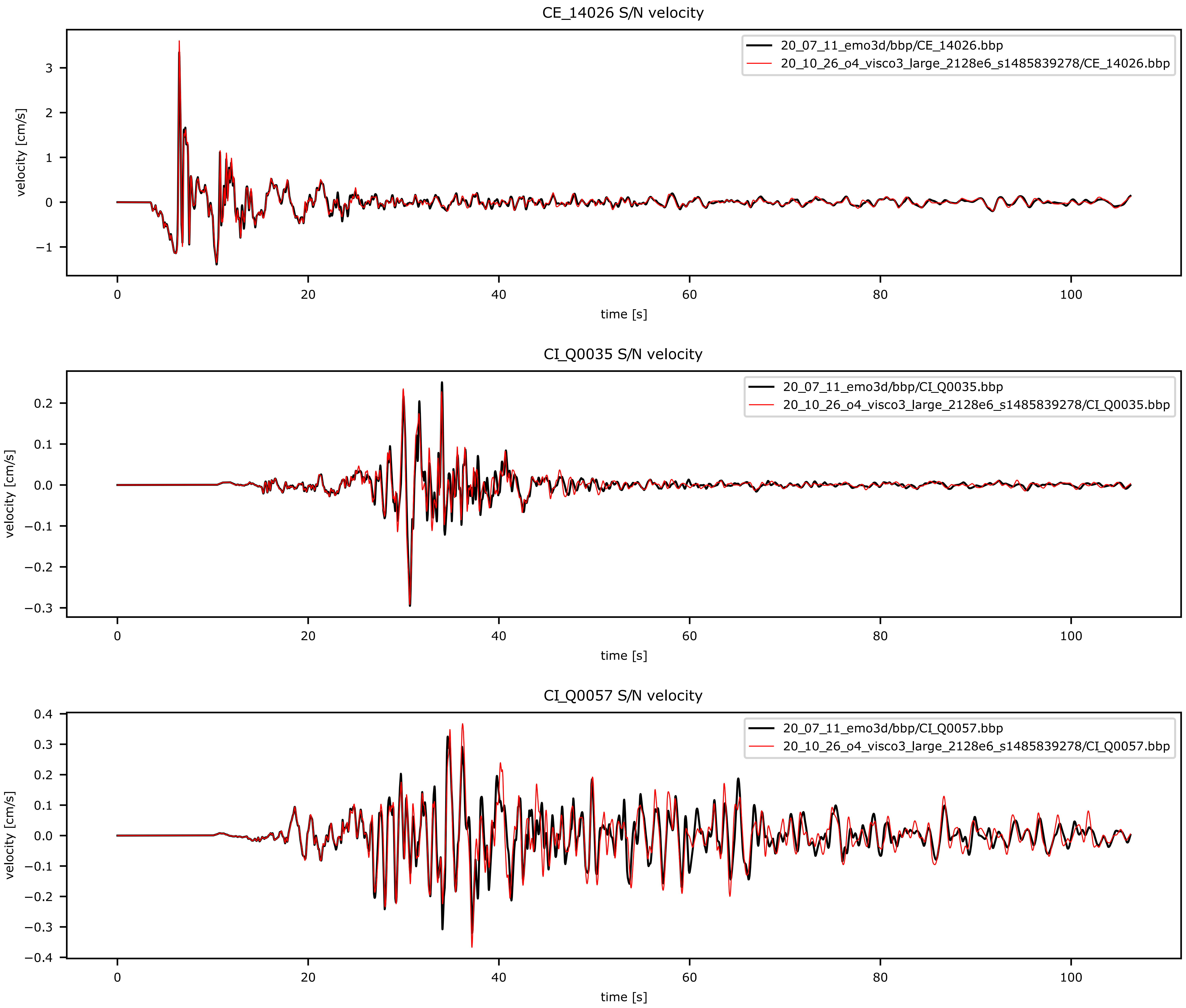


FRIEDRICH-SCHILLER-UNIVERSITÄT JENA

Verification

Comparing Synthetics

Right figure: Comparison of EDGE's South-North velocity component (red) to another solver of the High Frequency project (black). Shown are synthetic seismograms for the three stations depicted in the figure below. The seismograms were low-pass filtered at 5Hz. EDGE's respective ground motion simulation harnessed 1,536 nodes of the Frontera machine for a total of 48 hours to advance the used 2.1 billion tetrahedral element mesh in time.

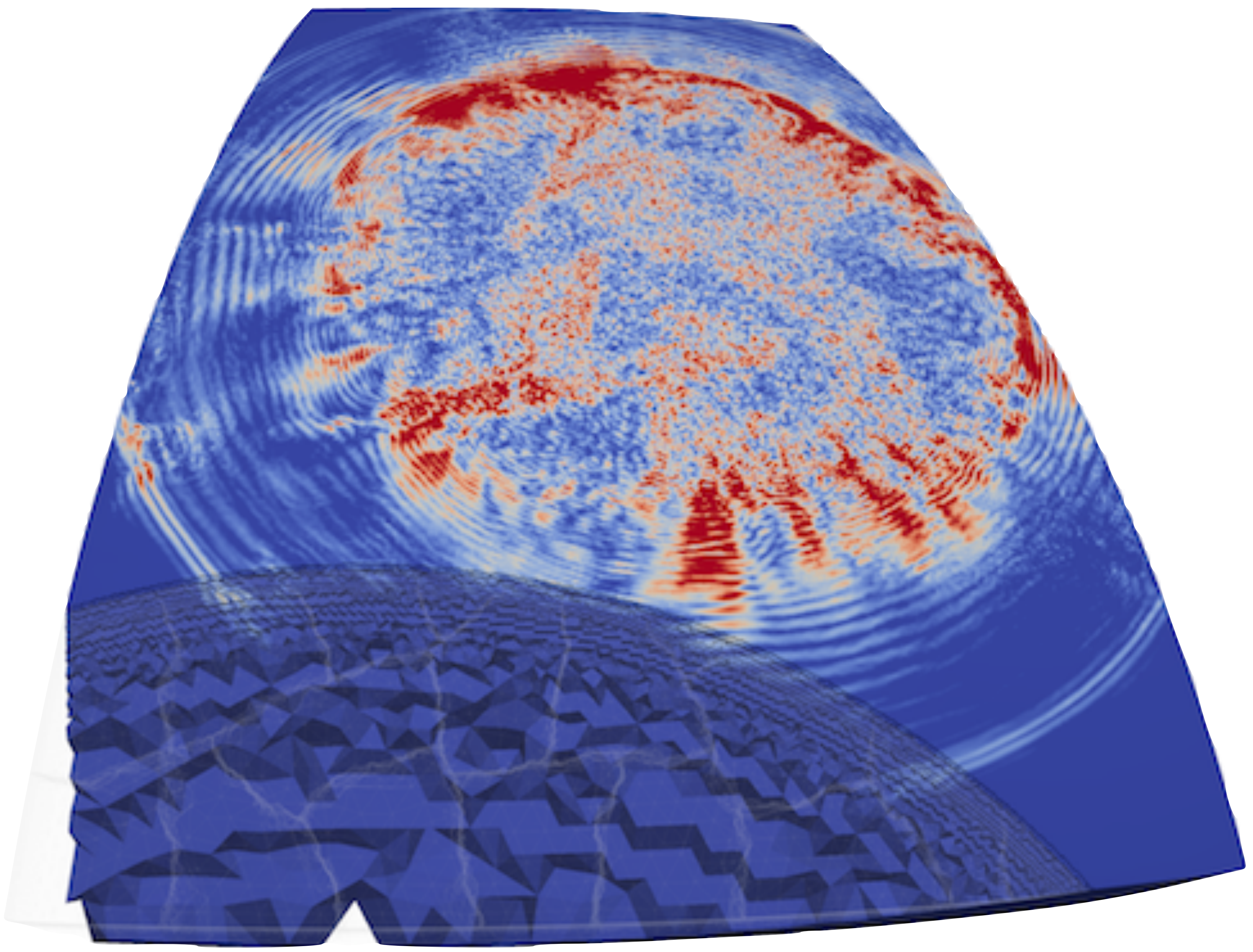


Solver

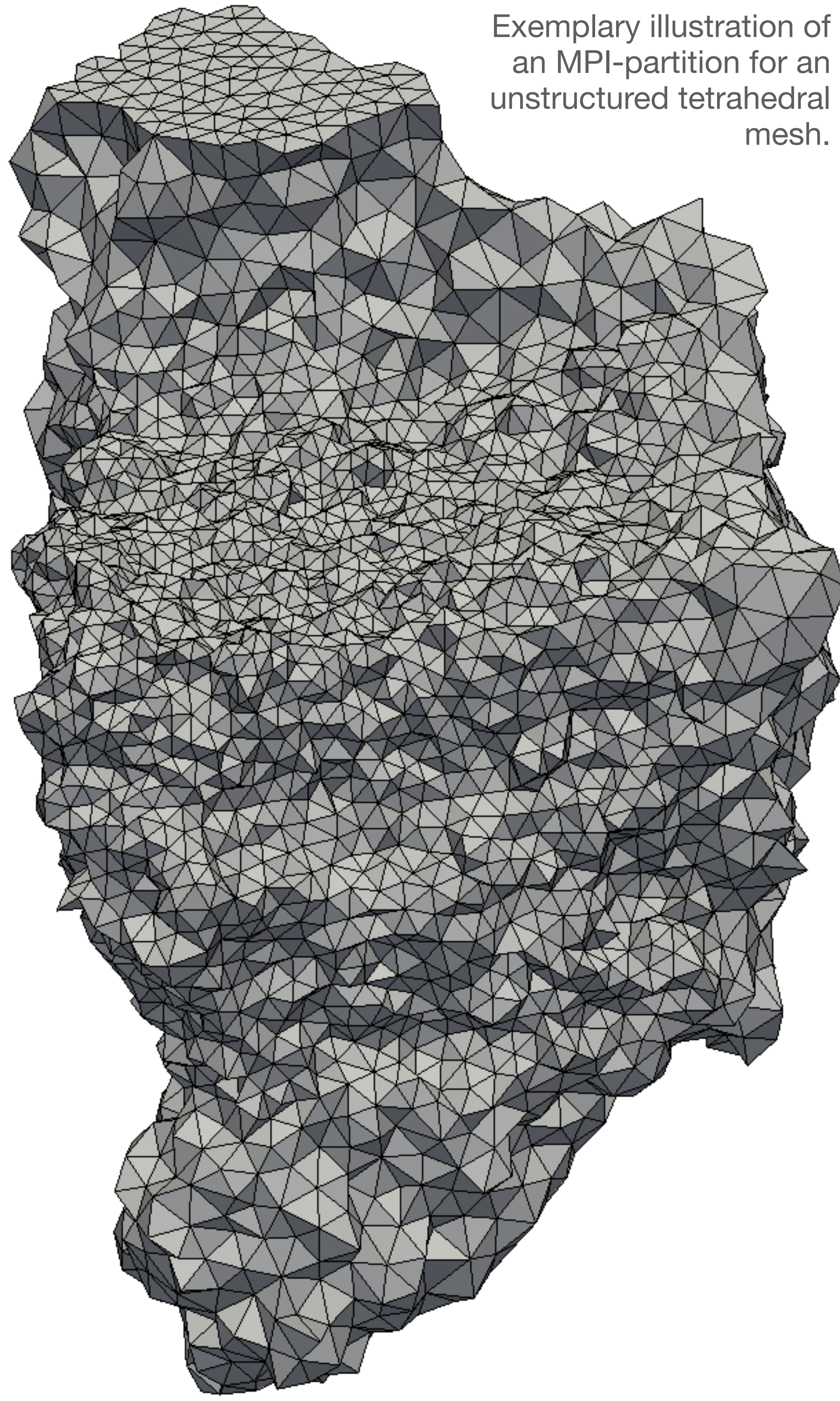
Extreme-scale Discontinuous Galerkin Environment (EDGE)

- Focus: static meshes with high geometric complexity
- Novel scalable local time stepping scheme with a full integration in EDGE's preprocessing tools
- Unique support for fused simulations exploiting inter-simulation parallelism
- Rapid prototyping through support for: line elements, quads, triangles, hexes, tets
- Parallelization: assembly kernels for highest performance on many recent CPU architectures (AVX, AVX2, AVX512, ASIMD, SVE); advanced OpenMP+MPI for hidden communication
- Extremely scalable with sustained petascale performance: 10.4 DP-PFLOPS on Cori II, 1.1 SP-PFLOPS in AWS
- Supporting tools for surface meshing, constrained velocity-aware volume meshing and partitioning
- Core solver and all tools are open source software (BSD-3), standard modeling and simulation pipeline relies exclusively 0 on open source software

Visualization of the absolute particle velocities for a simulation of the 2009 L'Aquila earthquake.



Exemplary illustration of an MPI-partition for an unstructured tetrahedral mesh.



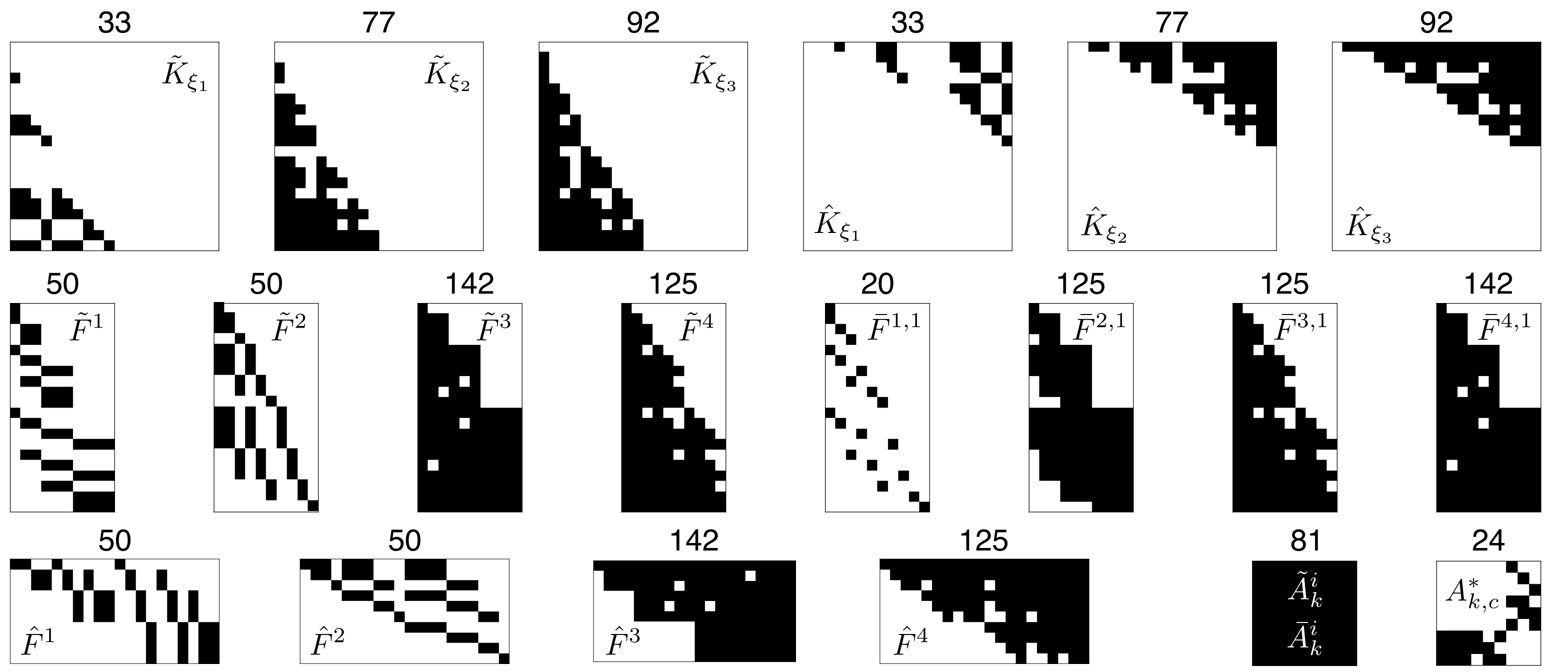


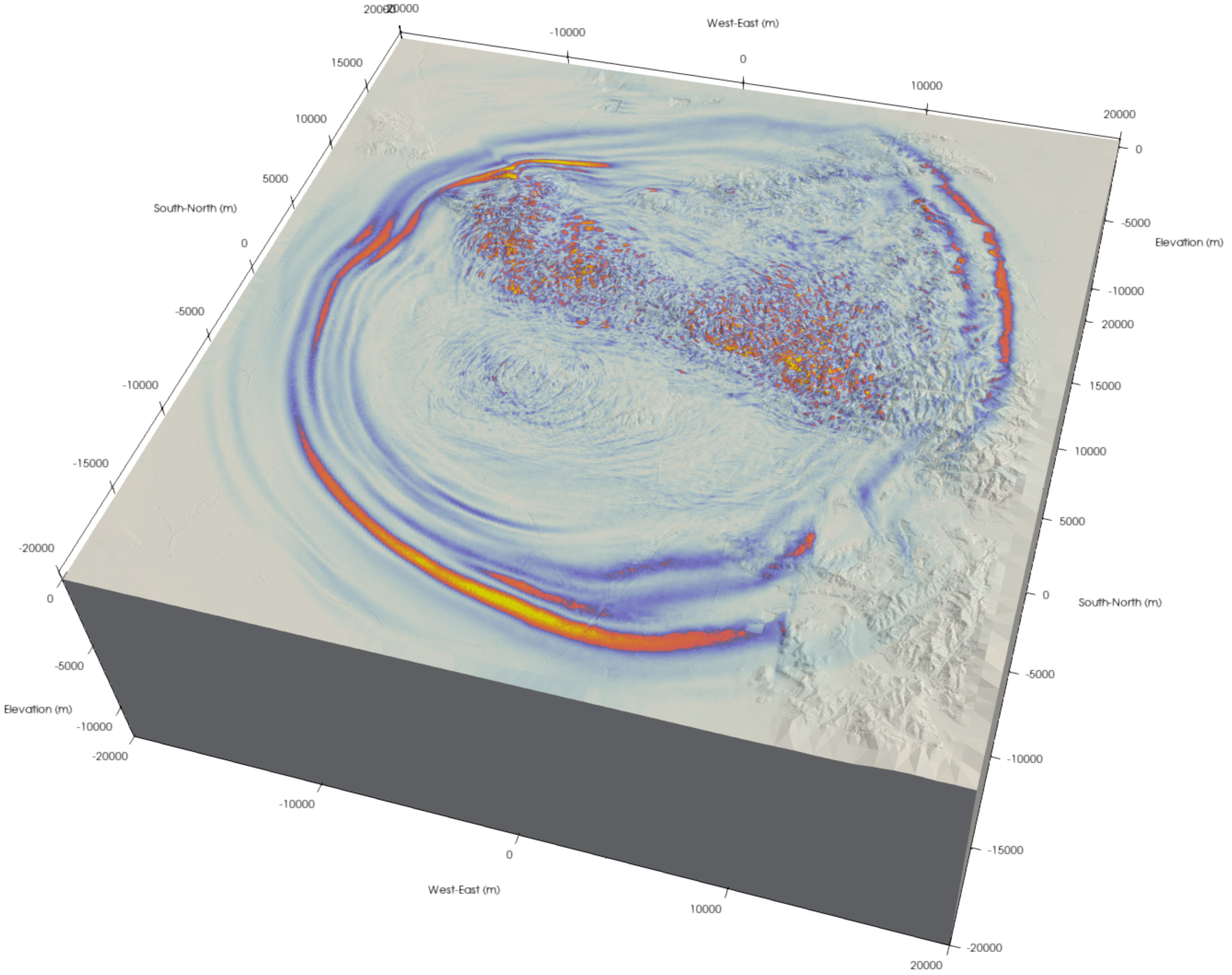
Illustration of all involves sparsity patterns for a fourth order ADER-DG discretization in EDGE. The numbers on top give the non-zero entries in the sparse matrices.

Feeding the Beast:
From Field Data to Synthetics

Input Data

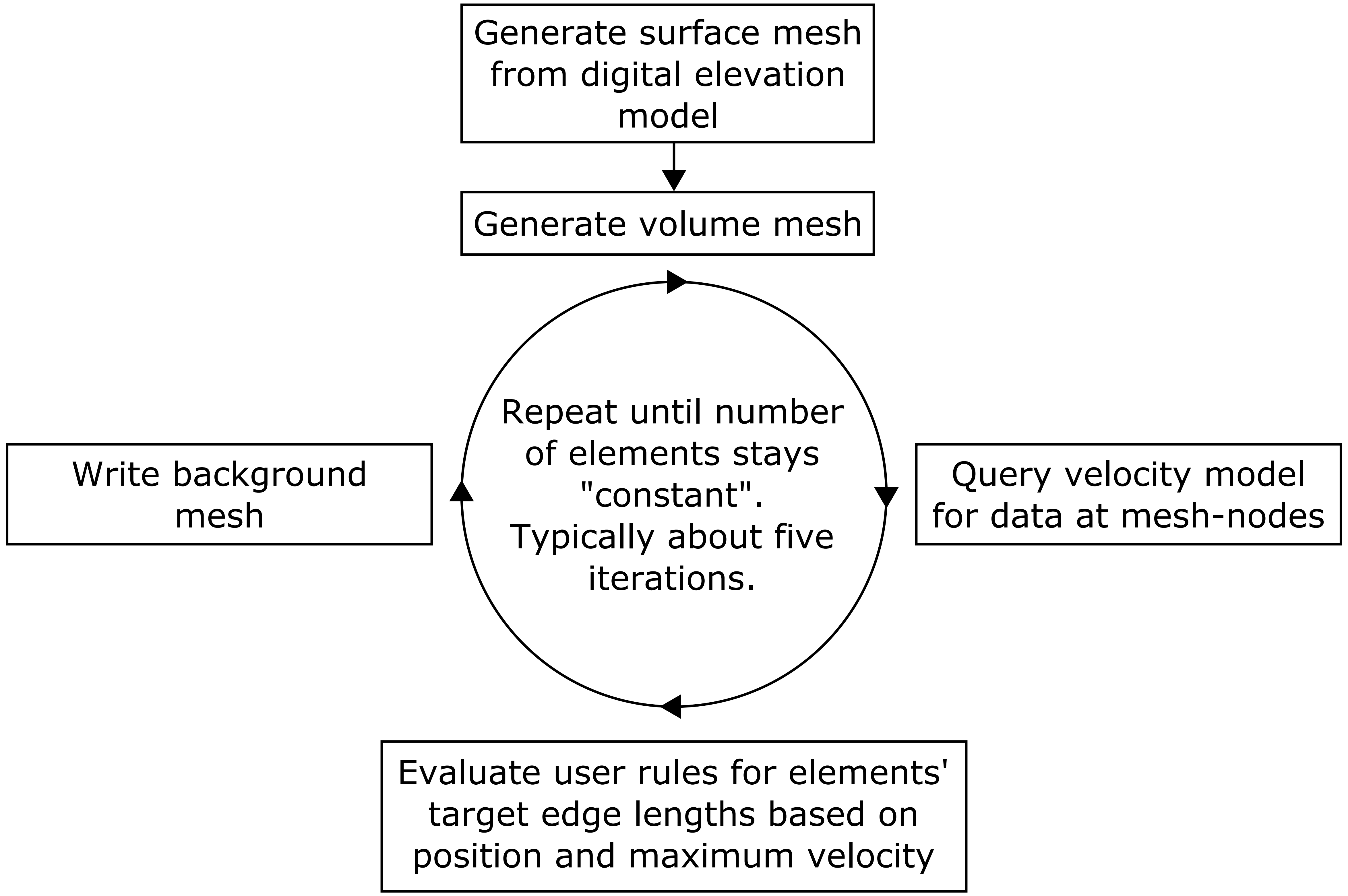
How do we get the data into our simulations?

- We have to combine:
  - Velocity model (pre-processed)
  - Kinematic source(s)
  - Topography
- But: Our solver expects a consistent view
- Constraint: We don’t like wasting resources

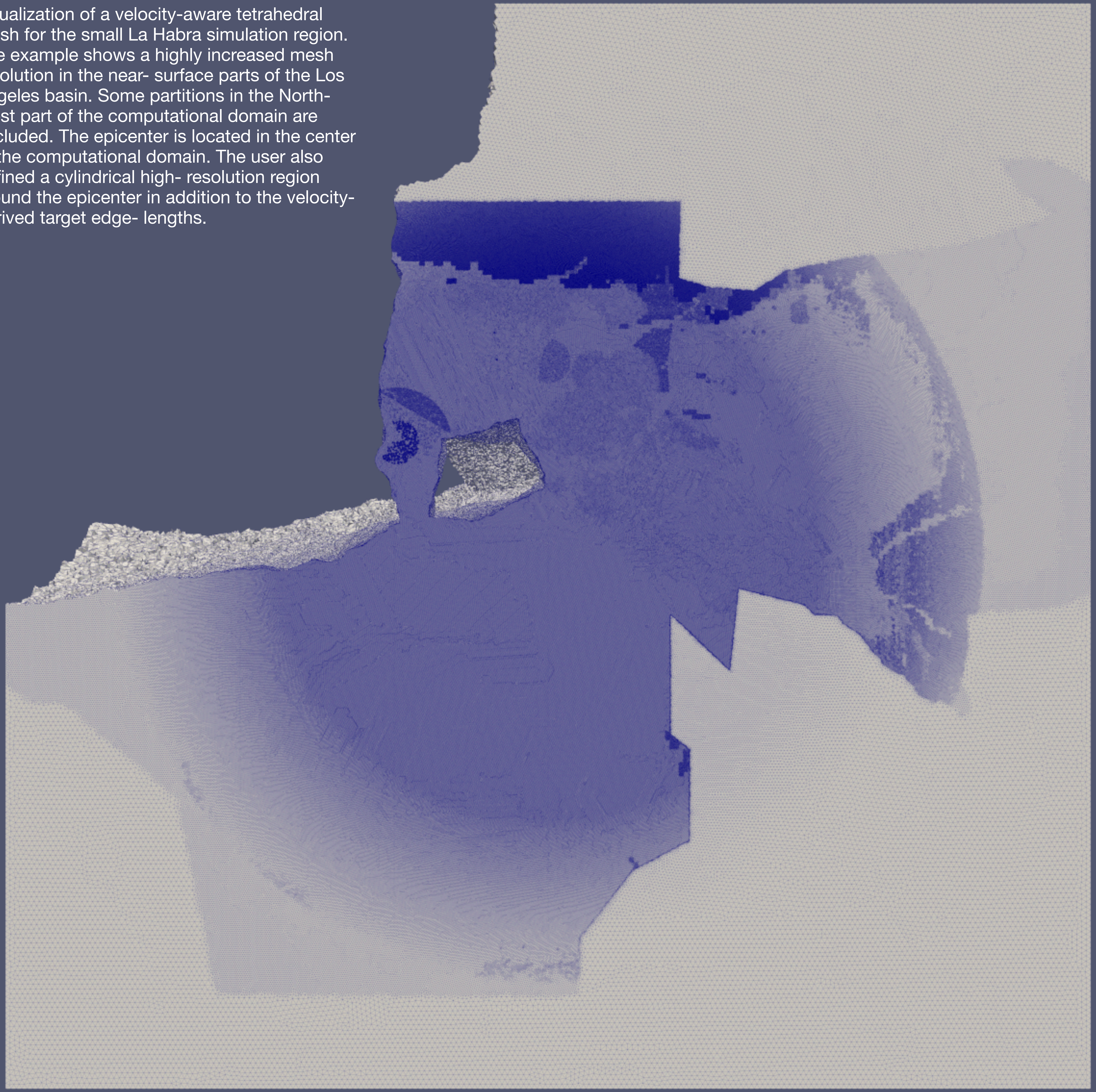


Fueling Our Solver

Velocity-Aware Meshing



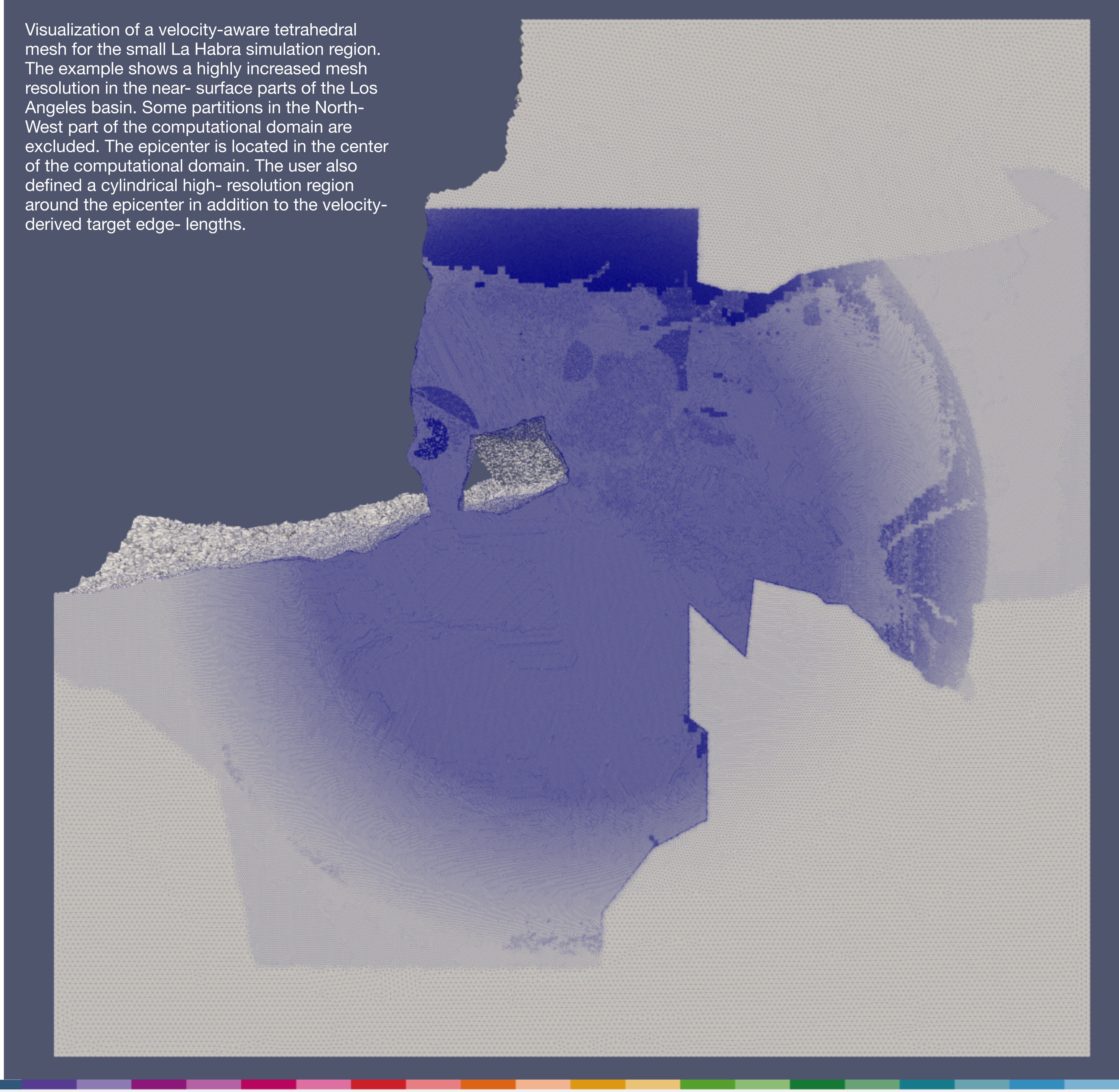
Visualization of a velocity-aware tetrahedral mesh for the small La Habra simulation region. The example shows a highly increased mesh resolution in the near- surface parts of the Los Angeles basin. Some partitions in the North-West part of the computational domain are excluded. The epicenter is located in the center of the computational domain. The user also defined a cylindrical high- resolution region around the epicenter in addition to the velocity-derived target edge- lengths.



Fueling Our Solver

Massaging the Mesh

- Generate problem-aware tetrahedral mesh
- Derive local time stepping groups
- Derive weights for elements and faces representing computation and communication
- Partition the mesh using the weights
- Reorder the mesh based on the partitioning
- Write partitioned mesh to disk
- Assemble second file per partition which contains data required by the core solver, e.g., local communication structure or per-element seismic velocities

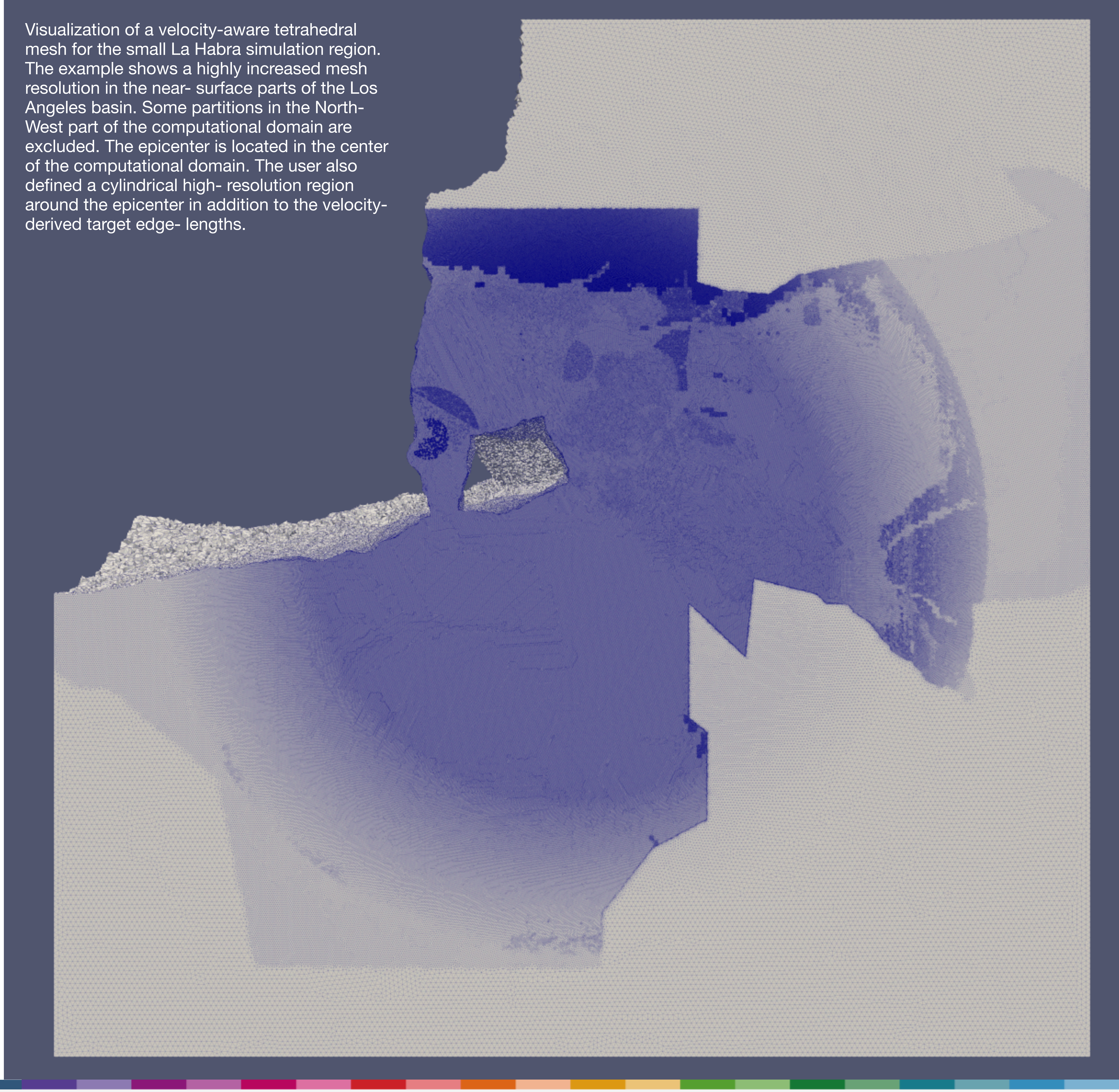


FRIEDRICH-SCHILLER-UNIVERSITÄT JENA

Working with Data

Some Thoughts

- Efficient and automated pre-processing is key!
- Core solver running at scale should be kept as simple as possible
- End-to-end use of open source software for the entire production workflow makes software accessible



FRIEDRICH-SCHILLER-UNIVERSITÄT JENA

Summary
Wrapping Up

- EDGE relies on unstructured meshes for high flexibility
- The solver stands and falls with its preprocessing capabilities
- We presented our automation of this process using the tools EDGEcut and EDGE-V
- Everything EDGE is available from: https://dial3343.org
- The slides of this presentation are available from: http://short.dial3343.org/gs21sli