

2020 SCEC Annual Report

Mining CyberShake for Earthquake Scenarios with Correlation Structure

Report for SCEC Award 20164

PI:

Dr. Kim B. Olsen

Institution:

Department of Geological Sciences, San
Diego State University, San Diego, CA
92182-1020

Publications and Reports:

None

Summary. We have implemented a process for enhanced mining and enhancement of the vast SCEC CyberShake (CS) database, in order to extract individual scenarios and prepare them for use in societal applications such as earthquake loss estimation and structural engineering analysis. The enhancement includes support for disaggregation for an area rather than a single site, as well as high-resolution scenario spectral acceleration maps for selected periods, including spatial and inter-frequency correlation. The final product enhances the applicability of the CS database in areas including loss estimation for the insurance industry.

Interpolated Scenario Maps. SCEC's CS platform was designed for calculation of seismic hazard curves at stations with a relative coarse spacing (kilometers to tens of kilometers). The CS Platform is set up to interpolate intensity measures at the hazard sites by means of Ground Motion Prediction Equations, extending the coarse resolution of the CS locations to much finer resolution in seismic hazard maps (see). However, the (100,000+) individual scenarios only exist in CS in the coarse (kilometer scale) resolution, which is not very useful for exercises such as damage and loss calculation on a much finer spatial scale. Two-component synthetic seismograms, as well as pre-computed spectral accelerations (SAs) at selected periods, are stored at all the hazard sites. Here, we have developed and streamlined a process to extract fine-grained SA scenario maps for scenario-based analysis, using the same principles as currently employed to generate CS hazard maps. The user will be able to choose between scenarios stored in the CS database, for example, based on disaggregation analysis.

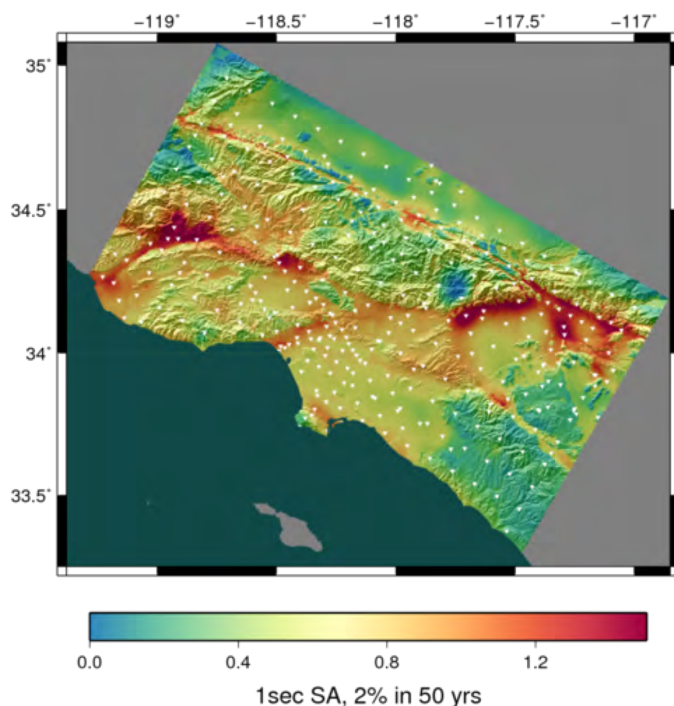


Figure 1. CyberShake Hazard map. The white triangles depict location of hazard sites. CyberShake was originally set up to generate smooth, interpolated hazards maps, while individual scenarios in the database are only defined at the hazard sites.

The PI has been working with SCEC developer Kevin Milner, who has provided support for generating smooth, interpolated scenario maps of spectral accelerations at selected periods (see **Fig. 2**). While smooth scenario maps can easily be generated using GMPEs, they are not based on physics-based calculations, and can be very different (see example in **Fig. 2**).

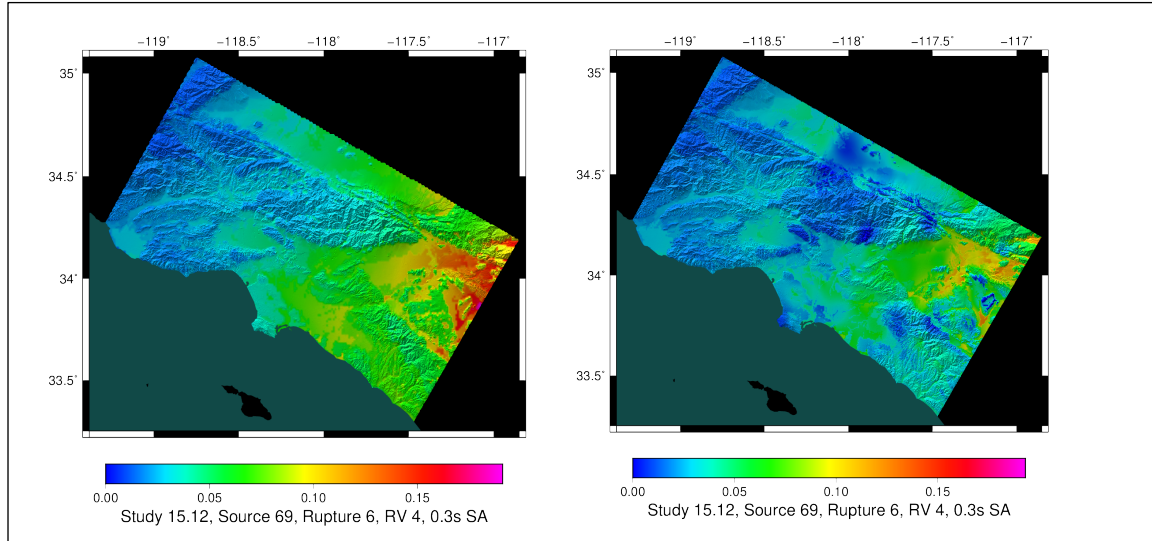


Figure 2. Illustration of interpolated spectral acceleration scenario maps for CyberShake, using Source variation 69, rupture ID 6, rupture variation 4, period 0.3 s (broadband). (left) Scenario results for GMPE values only, and (right) scenario results for CyberShake 15.12. Note that the two maps are very similar toward the north (left), where no CS hazard sites exist. On the other hand, the interpolated CS scenario is considerably different from the GMPE results in the SE (right) area, where the CS results are obtained from 3D modeling.

Disaggregation for Regions. Before this project, CS provides support to disaggregate the database to extract rupture scenarios that contribute the most to peak ground motions (SAs) at selected sites. In addition, this disaggregation procedure does not allow for more advanced queries, such as largest average SAs in a region, which would be desirable for loss calculation.

To enable the new commands, please download the file **disagg_op.py.tgz** from: <https://kbolsen.sdsu.edu/research.html> and untar/zip. This package includes the python script **disagg_op.py** plus files needed for input.

Java and pip are required to run. Use the following line to check whether pip is installed: `python -m pip --version`. If not installed, install using: `curl https://bootstrap.pypa.io/get-pip.py -o get-pip.py; python get-pip.py`. Python packages can be installed using: `pip install -r requirements.txt`.

To run the script, call the python script with arguments, for example:

```
python disagg_op.py --site_name=LADT --k 2 --radius 10 --period 1 -o loc_-118.26_3.05_1
```

This line will gather sites within 10 km away from site “LADT”, and perform disaggregation based on SA-1s. The outputs are stored in the directory “loc_-118.26_3.05_1”. The top k events are selected to plot their SA maps with spatial correlation.

It is also possible to specify a certain longitude/latitude, instead of site name. Multiple periods should be separated by commas. The resolution of the SA maps is user determined (here 0.1 degree).

`python disagg_op.py --lon -119 --lat 34 --k 2 --radius 10 --period 1,2 --spacing 0.1 -o /tmp`
The following command will find the sites of interest (but will not plot the SA maps):

`python disagg_op.py --site_name=LADT --k 1 --radius 10 --period 1 -o --plot_sa_map 0 loc_-118.26_3.05_1`

This line will print help messages: `python disagg_op.py -h`. The screen output should look similar to **Fig. 3**:

```
[18:16 05/06]disagg >>> python disagg_op.py --site_name=LADT --k 2 --period 1 -o loc_-118.26_3.05_1

Nearby sites will be written to file: loc_-118.26_34.05_1/sites_location.txt

Location = (-118.26, 34.05)
Periods = [1]; Radius = 10 km

Output to directory:
loc_-118.26_34.05_1

Nearby sites:
LADT, P14, P9, USC, P10, s347, s348, s389, P11, P19, P8, s366, P12, s346, P18, s349, s391, s388, P7

The 2 most severe scenarios are:

source_desc      source_id  erf_id  rup_id  ...  vel_model_id  Contribution  ExceedRate
Puente Hills (LA)      244      36      8      ...      5      1203.59      0.001257
Elysian Park (Upper)   158      36      8      ...      5      714.87      0.000757

[2 rows x 7 columns]
java -Xmx2G -cp opensha-cybershake-all.jar org.opensha.sha.cybershake.maps.CyberShakeScenarioShakeMapGenerator --study STUDY_15_12 --period 1 --source-id 244 --rupture-id 8 --rupture-var-id 6 --gmpe NGAWest_2014_AVG_NOIDRISS --spatial-corr-fields 20 --spacing 0.1 -o loc_-118.26_34.05_1
```

Figure 3. Example output from Python script for regional disaggregation.

Note, that (1) the script has only been tested on the OLCF Andes platform; (2) Java dependency may be a problem on some MacOS systems; (3) connection to the USC database can be unstable due to their firewall configurations; and (4) the entire procedure may take tens of minutes or longer.

A few files are generated in the output directory, unless specified in the arguments. The site locations are included in “**sites_location.txt**”, and shown in the map “**sites_location.pdf**” and “**sites_location.png**”:

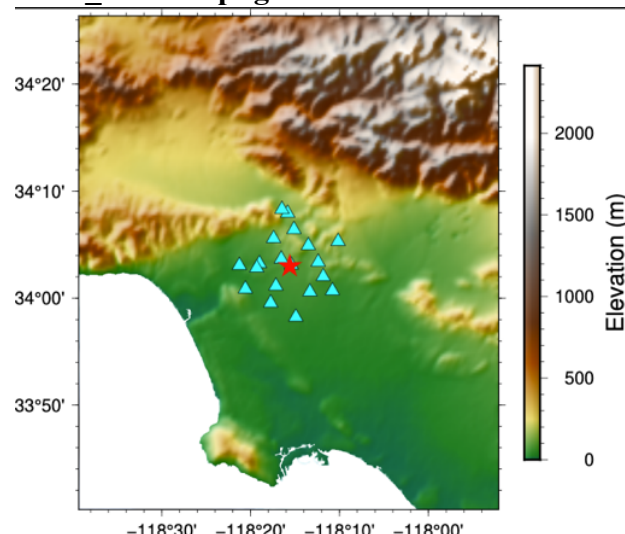


Figure 4. Output file “disagg.png” or “disagg.pdf”. The red star is the target location, and cyan triangles are the nearby stations within 10 km for areal disaggregation.

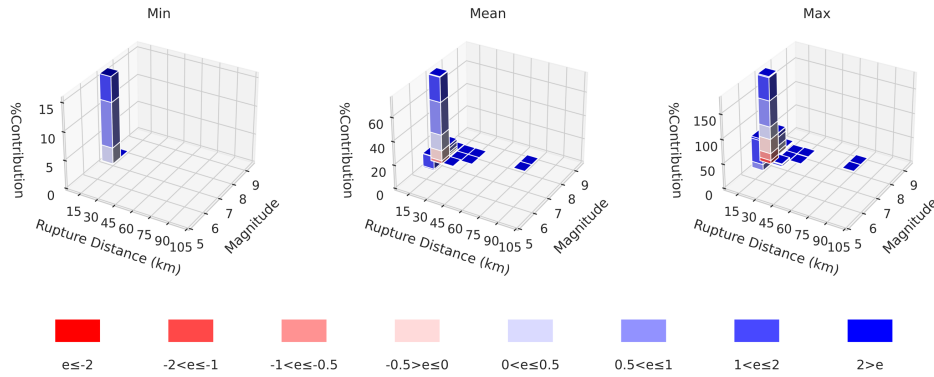


Figure 5. Example of output for areal disaggregation.

A full list of contributing scenarios is recorded in “**scenarios.csv**”, sorted by their hazard contribution in a descending order (see Table 1).

	source_desc	source_id	erf_id	rup_id	rup_var_id	vel_model_id	Contribution	ExceedRate
59	Puente Hills (LA)	244	36	8	6	5	1.2e+03	0.00126
29	Elysian Park (Upper)	158	36	8	6	5	715	0.000757
57	Puente Hills	242	36	8	6	5	430	0.00045
42	Hollywood	184	36	8	6	5	416	0.000423
62	Raymond	247	36	8	6	5	322	0.000328
210	Santa Monica, alt 2	263	36	8	6	5	106	0.000107
216	Verdugo	283	36	8	6	5	89.1	9.39e-05
48	Newport Inglewood Connected alt 2	219	36	8	6	5	76.7	7.77e-05
52	Newport-Inglewood, alt 2	216	36	8	6	5	66.4	6.74e-05
50	Newport-Inglewood, alt 1	215	36	8	6	5	53.3	5.41e-05
207	Santa Monica Connected alt 2	265	36	8	6	5	50.1	5.07e-05
60	Puente Hills (LA)	244	36	8	9	5	45.5	4.58e-05
46	Newport Inglewood Connected alt 1	218	36	8	6	5	40.5	4.12e-05
61	Puente Hills (Santa Fe Springs)	245	36	8	6	5	24.4	2.62e-05
58	Puente Hills	242	36	8	9	5	11.7	1.18e-05
47	Newport Inglewood Connected alt 1	218	36	8	9	5	10.1	1.01e-05
49	Newport Inglewood Connected alt 2	219	36	8	9	5	9.08	9.14e-06
51	Newport-Inglewood, alt 1	215	36	8	9	5	8.83	8.89e-06
53	Newport-Inglewood, alt 2	216	36	8	9	5	5.92	5.96e-06
44	Malibu Coast, alt 1	207	36	8	6	5	3.46	3.49e-06
208	Santa Monica Connected alt 2	265	36	8	9	5	2.31	2.33e-06
209	Santa Monica, alt 1	262	36	8	6	5	1.88	1.89e-06
43	Hollywood	184	36	8	9	5	1.59	1.6e-06
56	Palos Verdes Connected	232	36	8	9	5	1.47	1.48e-06

Table 1. List of contributing scenarios is recorded in sorted by their hazard contribution.

In addition, interpolated SA maps and GMPE base maps are also plotted. The SA values are stored in “*.txt” files, including:

“cs_shakemap_src_*_rup_*_rv_*_*s_cs_amps.txt”: SA values at CyberShake hazard sites
“cs_shakemap_src_*_rup_*_rv_*_*s_interpolated.txt”: SA values at interpolated grids
“cs_shakemap_src_*_rup_*_rv_*_*s_gmpe_amps.txt”: GMPE SA values at interpolated grids
“cs_shakemap_src_*_rup_*_rv_*_*s_rand*.txt”: Correlated SA values (interpolated)

Ground motion correlation. Earthquake loss calculations based on ground motion estimates depend on realistic correlation structure in the results (e.g., Burks and Baker, 2014; Weatherill et al., 2015; Stafford, 2017; Bayless and Abrahamson, 2019). While the coarse-grained, deterministic (low-frequency, $< \sim 1$ Hz) part of the CS database includes spatial correlation effects

from the finite-difference wave propagation, this is not the case for the interpolated set of sites where the intensity measures are derived, in part, from GMPEs. In addition, when broadband CS synthetics are generated, typically with a stochastic approach, the spectral accelerations depending on the higher frequencies are depleted in inter-frequency and spatial correlation. We have augmented the interpolated scenario spectral accelerations with realistic inter-frequency and spatial correlation, using the SCEC-funded approaches by Wang et al. (2019) and Wang et al. (2021).

An example of the results is shown in:

http://opensha.usc.edu/ftp/kmilner/cybershake/shake_map_gen/spatial_corr_example/.

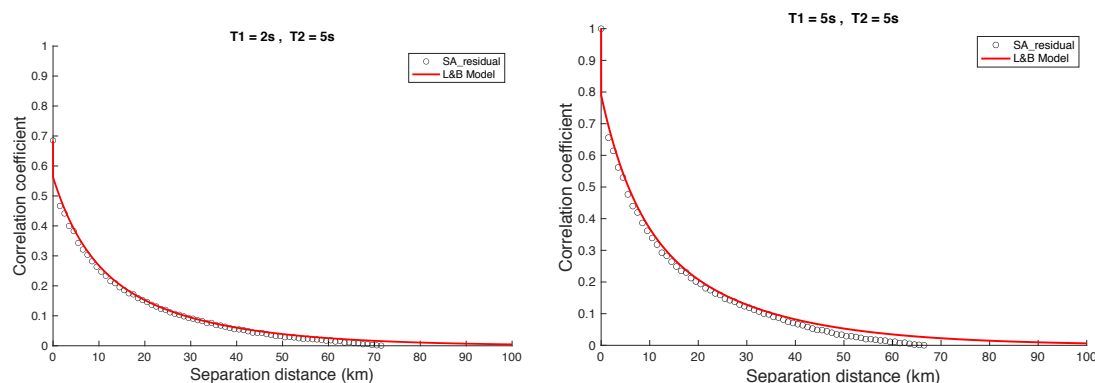


Figure 6. Example inter-frequency and spatial correlation results for (left) SA-3s versus SA-5s (right), SA-5s versus SA-5s in terms of SA residual (circles) calculated on the CS platform. The red line depicts the model by Loth and Baker.

Note, when “spatial-” relevant parameters are provided in the command line arguments for the disaggregation python script described above, maps with spatial correlations included will also be generated (see example in **Fig. 7**).

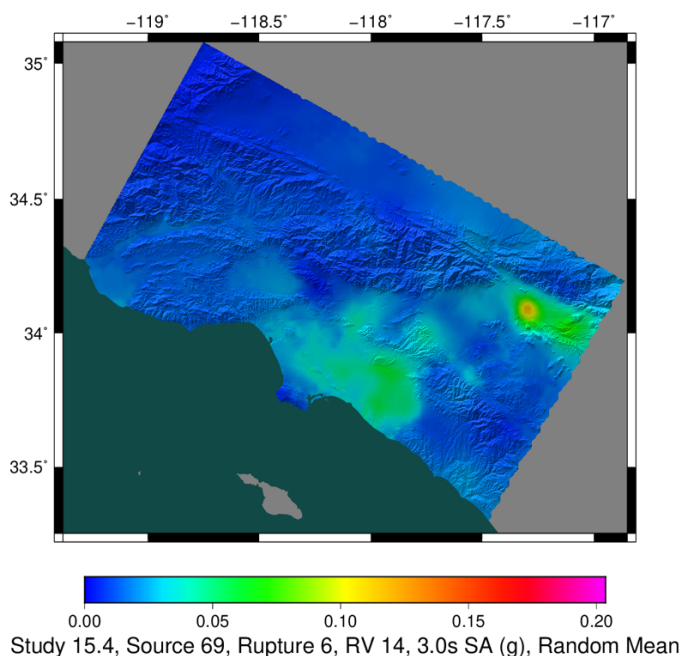


Figure 7. Example interpolated SA map (SA-1s), study 15.4, source 69, rupture 6, rupture variation 14, SA period 3.0 s, include inter-frequency and spatial correlation.

References

- Bayless J., and N.A. Abrahamson (2019). An empirical model for the interfrequency correlation of epsilon for Fourier amplitude spectra, *Bull. Seismol. Soc. Am.* **109**, no. 3, 1058–1070.
- Burks, L. S., and J. W. Baker (2014). Validation of ground-motion simulations through simple proxies for the response of engineered systems, *Bull. Seismol. Soc. Am.* 104, no. 4, 1930–1946.
- Stafford, P. J. (2017). Interfrequency correlations among Fourier spectral ordinates and implications for stochastic ground-motion simulation, *Bull. Seismol. Soc. Am.* 107, no. 6, 2774–2791.
- Weatherill, G. A., V. Silva, H. Crowley, and P. Bazzurro (2015). Exploring the impact of spatial correlations and uncertainties for portfolio analysis in probabilistic seismic loss estimation, *Bull. Earthq. Eng.* 13, no. 4, 957–981.
- Wang, N., K.B. Olsen, and S.M. Day (2021). A Frequency-Dependent Ground-Motion Spatial Correlation Model of Within-Event Residuals for Fourier Amplitude Spectra, *Earthquake Spectra*, DOI: 10.1177/8755293020981995.
- Wang, N., R. Takedatsu, K.B. Olsen, and S.M. Day (2019). Broadband Ground-Motion Simulation with Inter-Frequency Correlations, *Bull. Seis. Soc. Am.* 109, 2437–2446, doi: 10.1785/0120190136.