16th World Conference on Earthquake, 16WCEE 2017
Santiago Chile, January 9th to 13th 2017
Paper N° 4574
Registration Code: S-L1462283171

INFLUENCE OF THE SOURCE, SEISMIC VELOCITY, AND ATTENUATION MODELS ON THE VALIDATION OF GROUND MOTION SIMULATIONS

R. Taborda[1,2], N. Khoshnevis[2], S. Azizzadeh-Roodpish[1,2], M. M. Huda[2]

[1] Department of Civil Engineering, The University of Memphis, Memphis, TN, USA, ricardo.taborda@memphis.edu
[2] Center for Earthquake Research and Information, The University of Memphis, Memphis, TN, USA

Abstract

This study investigates the accuracy of deterministic, regional-scale ground motion simulations of moderate magnitude earthquakes, and the influence that the source, seismic velocity, and attenuation models have on synthetic results when compared to data. Thanks to recent advances in earthquake ground motion simulation algorithms and models, and to the continuous growth of high performance computing systems and applications, regional-scale simulations of earthquake ground motion using numerical approaches are now becoming commonplace in research environments. However, before these simulations can be transferred to earthquake engineering practice, much work needs to be done to confirm the robustness of models and simulation methods. This requires a continuous effort on simulation validation, which entails, among other things, the comparison of synthetics against recorded ground motions of past earthquakes. These comparisons provide confidence and help improve simulation procedures and models over time. To this end, we evaluate the accuracy of simulations using quantitative metrics that compare synthetics and data. We use metrics that have physical meaning to both seismologists and engineers. In previous efforts we have shown (separately) that validation results are significantly controlled by the choice of the velocity model. Here, we reinforce the influence of velocity models and further investigate the effects of attenuation parameters and models, as well as the influence of the source. We first focus our attention on the selection of the appropriate velocity model by performing a large set of simulations for multiple historical events in southern California, using the different community velocity models available for this region. After identifying the model that consistently yields best possible approximations, we concentrate in one recent event of moderate magnitude in the greater Los Angeles basin area for which there are significant number of high-quality data. Through validation of simulations for this event, we investigate the influence of relationships of the attenuation quality factor with respect to the shear wave velocity (Qs-Vs relationships), and the behavior of attenuation as a frequency dependent parameter. Here, intrinsic attenuation is represented using a viscoelastic model to mimic the effects of internal friction in the transmitting media. We also explore the relative influence of the source model. We analyze the results from point source and extended source models. Our results shed light on the relative weight of these factors with respect to each other, and how they influence validation results, and thus simulations as a whole.

Keywords: ground motion simulation; seismic velocity models; attenuation; source models; ground motion validation.

1. Introduction

Earthquake ground motion simulations are necessary to complement existing or substitute inexistent data from past events in regions susceptible to earthquake effects. Among the different methods available for ground motion simulation, physics-based (deterministic) approaches are now often employed in problems of regional seismology and earthquake engineering. These simulations use numerical methods such as the finite difference, finite element, or spectral element methods to resolve three-dimensional (3D) wave propagation problems in heterogeneous media and the power of parallel computers to reduce the time to solution, e.g. [1, 2]. Recent advances in this area facilitate the analysis of earthquake effects due to past and future (scenario) earthquakes, e.g. [3, 4]; help address complex engineering problems, e.g. [5, 6]; and provide a framework for reducing uncertainty in earthquake hazard estimation [7].

While much has been achieved in recent years in physics-based ground motion simulation—especially in reaching to higher maximum frequencies ($f_{max} \geq 1$ Hz), modeling rupture dynamics, including small-scale heterogeneities, and considering nonlinear effects, e.g. [8–10]—a long road lies ahead before these approaches can be fully implemented into engineering practice. To this end, constant effort is needed in verification and validation of ground motion synthetics. Here, we understand verification as the accuracy evaluation of a solution with respect to the intended conceptual model, and validation as the accuracy evaluation of a simulation with respect to observational data (i.e. seismic records). Recent efforts in verification have proven that current approaches in physics-based earthquake simulation are robust enough for solving complex wave propagation problems when compared to equivalent solutions, e.g. [11, 12]. Validation efforts, on the other hand, indicate that although very promising, much work is still necessary in terms of modeling accuracy [13, 14].

This paper addresses issues related to validation of ground motion simulation from the perspective of various modeling parameters. We are interested in investigating the effect that different velocity, source and attenuation models have on simulations, and in evaluating such effects quantitatively in terms of validation results when synthetics are compared to data from past recorded earthquakes. We conduct a series of simulations using a finite element software [15, 16] for regional scale earthquake ground motion simulation in 3D heterogeneous media under kinematic source excitation. In particular, we focus on the region of southern California as our natural laboratory and analyze synthetics and records from earthquakes in the vicinity of the greater Los Angeles region. We consider two seismic community velocity models, and borrow results from a previous related effort [17] to showcase that improvements in velocity models consistently lead to better validation results across a collection of events. We then focus our attention in one particular event, the M_w 5.4 2008 Chino Hills, California, earthquake. We simulate this event for a maximum frequency, $f_{max} = 2$ Hz using different minimum shear-wave velocities ($V_{Smin} = 200$ and 500 m/s), source models (point and finite fault), attenuation parameters (Q_S-V_S relationships), and attenuation viscoelastic models (frequency dependent and frequency independent Q). The analysis of results is mainly based on goodness-of-fit (GOF) metrics following the approach introduced by Anderson [18].

Our GOF results indicate that validation of ground motion simulations is significantly sensitive to the different modeling parameters, even for relatively minor changes and moderate magnitude events. It is expected, and it has been suggested by other [19], that much larger changes will result from considering more complex physics such as material nonlinearity. Therefore, it is important to first fully understand the implications of the varying modeling parameters in simpler problems, which is the main objective of this article and the subject of other ongoing efforts.

2. Events, Region of Interest, and Simulation Domain

We adopt the region of interest and simulation domain used in our previous validation studies for southern California [13, 14, 17]. This domain comprises a rectangular surface area of 180 km × 135 km, and the subsurface structure down to a depth of about 62 km. The region, shown in Fig. 1a, encloses the greater Los Angeles metropolitan area and other surrounding municipalities. The properties within the simulation domain include the main basins and geological structures in the region, namely: the Los Angeles, Chino and San

Bernardino basins; the San Fernando and Simi valleys; the San Gabriel, Santa Ana and Santa Monica mountains; and part of the Mojave desert, the Santa Clara river valley and the Ventura basin.

The material properties within the simulation domain (that is, the values of the seismic velocities, V_P and V_S, as well as the material's density, ρ) were extracted and stored into materialized (etree) models [20] using the Southern California Earthquake Center (SCEC) Unified Community Velocity Model (UCVM) software tools [21]. In particular, we consider here two seismic velocity models, CVM-S4 and CVM-S4.26.M01, which are models based on the CVM-S family [22] of community velocity models (CVMs) available for southern California and maintained and supported by SCEC. CVM-S4.26.M01 is the latest update to this family of models and was obtained after a series of iterations from a full-3D waveform tomographic inversion [23].

In the first part of our study [17], we considered 30 events of moderate magnitude ($3.6 \leq M_w \leq 5.4$) to evaluate the performance of the various CVMs available for this region (from the model families CVM-S and CVM-H) [22, 24]. Here, we replicate some of those results for the same 30 events—whose locations within the simulation domain are shown in Fig. 1b—and subsequently focus on one particular event to be used as benchmark. Such event is the 29 July 2008 M_w 5.4 Chino Hills, California, earthquake. Our interest in this particular event is the large dataset of observations available for validation (over 300 stations) and the possibility to draw conclusions based on previous validation efforts for this same earthquake [13, 14]. The location of the Chino Hills earthquake, which occurred at a depth of 14.7 km, is highlighted in Fig. 2b. Details about the hypocentral coordinates, magnitude, number of records available for validation, and focal mechanisms for the larger collection of events (A through AD in Fig. 2) are provided in [17].

For the particular case of the Chino Hills earthquake, here we considered two different source models. The first is a point source model represented by a double couple collocated at the hypocenter and with a slip rate function scaled to match the correct magnitude, as reported in [17]. The second, is a finite fault or extended source model used previously in [13, 14].

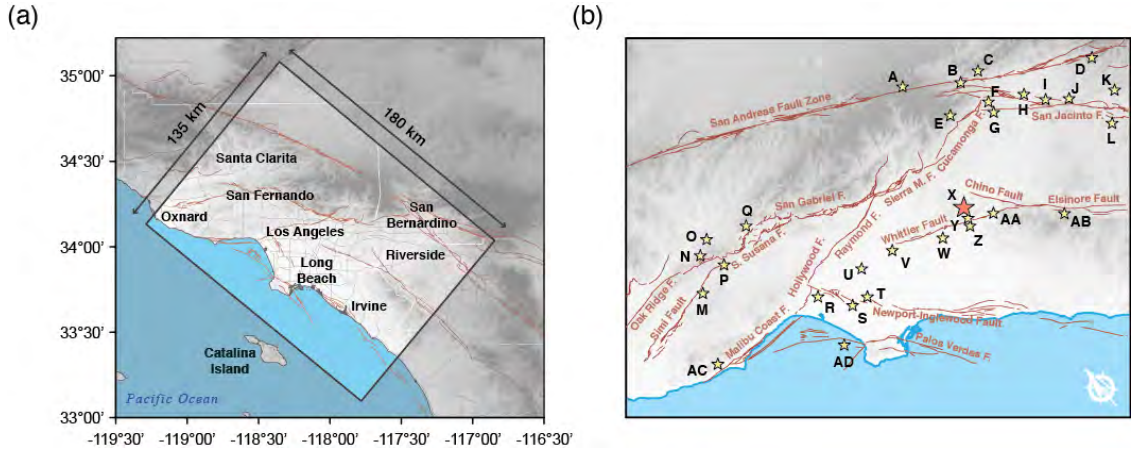

Fig. 1 – Region of interest, events, and simulation domain. (a) Horizontal surface projection of the simulation domain and location of the main cities within it. (b) Events used for evaluation of the velocity models (indicated with star symbols). Among all the earthquakes, the M_w 5.4 2008 Chino Hills earthquake, which is used here as a benchmark for the validation analysis, is highlighted with a larger, red stat (labeled as event X).

3. Simulation Method and Parameters

The simulations were carried out using Hercules, a finite-element parallel code developed at Carnegie Mellon University (CMU) [15, 16] and maintained and supported by researchers in CMU, the University of Memphis, EAFIT University, the National Autonomous University of Mexico (UNAM) and SCEC. Hercules implements a standard Galerkin method for discretizing the equations of elastodynamics in space, and advances explicitly in time to obtain the next-step state of nodal displacements. The time integration is approximated by first-order backward and second-order central differences for the velocity and acceleration fields, respectively [16].

Attenuation is introduced using a viscoelastic model using Maxwell and Voigt elements [25], which we call the BKT model. This model was originally formulated for frequency independent attenuation (Q) using 2 Maxwell elements and 1 Voigt element in parallel, but we have recently modified it to also account for frequency independent Q [26]. We call these models BKT2 (frequency independent, 2 Maxwell elements), BKT3 (frequency independent, 3 Maxwell elements) and BKT3F (frequency dependent, 3 Maxwell elements). Here, the attenuation quality factor is defined based on a target value Q_O, such that:

$$Q(f) = Q_O \quad \text{for } f < 1 \text{ Hz, and} \tag{1}$$

$$Q(f) = Q_O f^{\lambda} \quad \text{for } f \geq 1 \text{ Hz,} \tag{2}$$

where f is the frequency in Hz and λ is a constant that represents the exponential decay of attenuation with frequency. As mentioned before, Q_O corresponds to the desired (target) quality factor, which is obtained for either the *P*- or *S*-waves based on the relationships:

$$Q_S = \alpha V_S \text{, and} \tag{3}$$

$$Q_P = 2Q_S \text{,} \tag{4}$$

where α is a constant and V_S is given in km/s. The forms of Eqs. (1) through (4) is well-known in quantitative seismology and have been used before in numerous works, e.g. [3, 10, 27].

Table 1 – Simulation parameters.

Sim. ID	CVM-S		$V_{S\min}$		Pts. per wavelength		α in $Q_S = \alpha V_S$		λ in $Q(f) = Q_O f^{\lambda}$			Source		Magnitude		
	4	4.26	200	500	10	20	50	100	0 (a)	0 (b)	0.8 (b)	Point	Ext.	5.4	5.45	5.5
S1	•			•	•		•		•			•		•		
S2		•		•	•		•		•			•		•		
S3		•		•	•			•	•			•		•		
S4		•	•		•			•	•			•		•		
S5		•	•		•			•	•				•	•		
S6		•	•		•			•		•			•	•		
S7		•	•		•			•			•		•	•		
S8		•		•		•		•		•			•	•		
S9		•		•		•		•			•		•	•		
S10		•		•	•			•	•				•	•		
S11		•		•	•			•		•			•	•		
S12		•		•	•			•		•			•		•	
S13		•		•	•			•		•			•			•

(a) This corresponds to the attenuation model BKT2, which is frequency independent.
(b) This corresponds to the attenuation model BKT3, which can be frequency dependent if $\lambda \neq 0$.

In the first part of our study [17] we considered 4 different velocity models and 30 events, for a total of 120 simulations done for f_{max} = 1 Hz. In the second part of the study, which is the main focus of this article, we consider 13 additional simulations only for the 2008 Chino Hills earthquake, for which we focused on 2 velocity models (CVM-S4 and CVM-S4.26.M01) and consistently used f_{max} = 2 Hz and for all simulations. Table 1 summarizes the variable simulation parameters, which include the velocity model, the minimum shear wave velocity and number of points per wavelength used in for the generation of the finite element mesh, the attenuation parameters (explained above) and whether the source was represented by a single double couple (point source) or using the finite fault model employed in [13, 14] and obtained by [28]. We also did a couple of additional runs varying the magnitude of the event to see if there was any mismatch between the model and the inversion reported in [28].

4. Validation Method and Results

For validation we adopt the goodness-of-fit (GOF) method introduced by Anderson [18], with modifications by Taborda and Bielak [13]. This method compares synthetics against data using eleven individual parameters of physical meaning in seismology and engineering (Arias intensity integral, energy integral, Arias intensity value, total energy, peak acceleration, peak velocity, peak displacement, response spectrum, Fourier amplitude spectrum, cross correlation, and strong-phase duration) and ranks the similarity between any two signals with a final score in a scale ranging from 0 to 10. Previous calibration of the method has shown that scores above 8 indicate excellent fits, and scores below 4 are indicative of a poor fit, whereas values between 4 and 6, and between 6 and 8 are considered fair and good fits, respectively.

For each simulation we extracted synthetics at the same geographical locations of a collection of over 800 broadband and strong motion recording stations from different earthquake monitoring networks in the region. The three orthogonal components of synthetics were evaluated with respect to the corresponding recorded signals. Then, final scores were computed taking averages of all components and different frequency bands to produce GOF distribution maps for the simulation domain and to measure the improvements from one simulation to another, depending on the varying parameters.

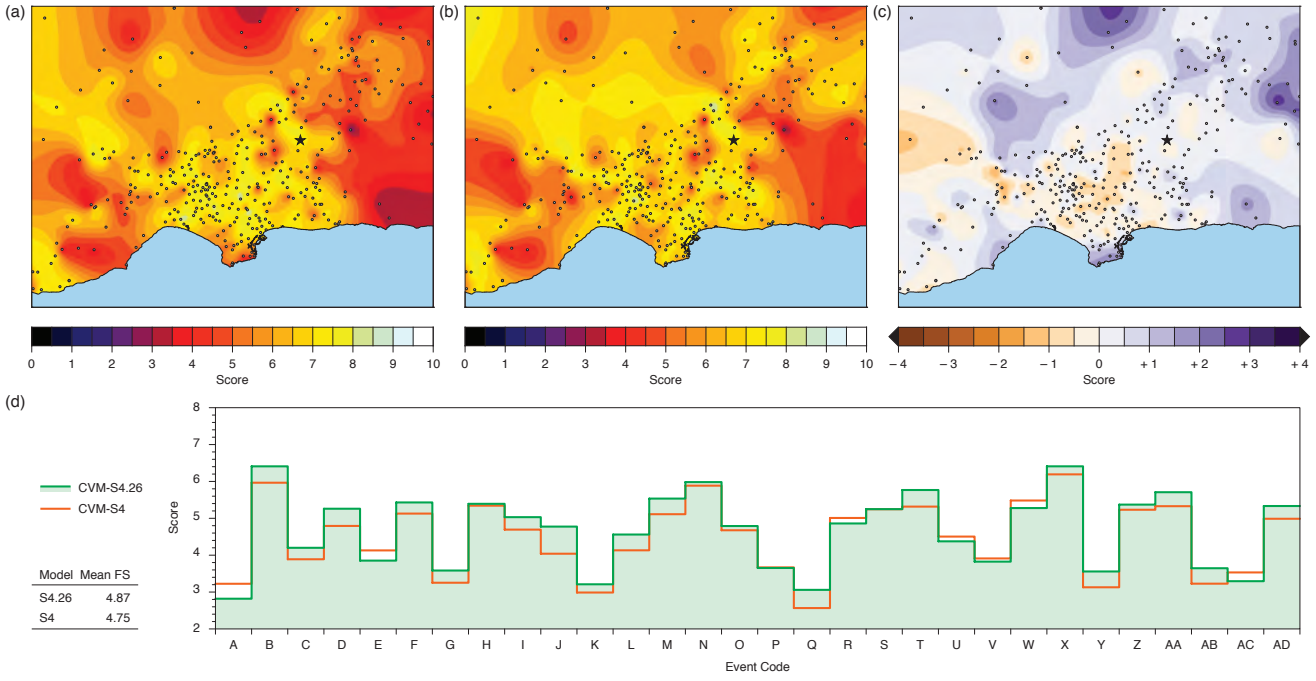

Fig. 2 – (a) Goodness-of-fit scores map obtained for the 2008 Chino Hills earthquake in the frequency range 0–1 Hz using the velocity model CVM-S4, i.e. simulation S1. (b) Same as before, but corresponding to results obtained using the model CVM-S4.26.M01, i.e. simulation S2. (c) Difference between the previous two maps, where positive values indicate improvement in S2 with respect to S1. (d) Summary chart of the final average GOF scores obtained for all events in Fig. 1b.

In the first part of our study, these comparisons allowed us to evaluate the effects of considering different velocity models [17]. Figs. 2a and 2b show the GOF maps obtained for the case of the 2008 Chino Hills earthquake for the two velocity models considered here, as presented in [17] and corresponding to frequencies between 0 and 1 Hz. In turn, Fig. 2c shows the difference between the two maps, where positive residuals indicate an improvement in the GOF scores. This exemplifies the benefit drawn from using the newer model CVM-S4.26.M01 over the older version CVM-S4. A summary of results (final GOF average score in the frequency band 0–1 Hz) for all the events in Fig. 1b is presented in Fig. 2d. This figure (modified from [17]) shows that for 23, out of the 30 events considered, the average GOF scores obtained with CVM-S4.26.M01 are better than those obtained using CVM-S4, which exemplifies the effect that using different velocity models can

have on the validation of simulations, and provides confidence in the latest version of the CVM-S family of velocity models for southern California.

We obtained similar results to those shown in Figs. 1a and 1b for all the simulations in Table 1, and several different combination of residual plots similar to that shown in Fig. 1c. We tried different combinations of the simulations in order to draw conclusions about the effect that the different simulation parameters have on the validation results. However, before we present those results, it is appropriate draw a parallel between the significance of the GOF scores and the simulation results in terms of what they mean when visually comparing synthetics and data. As an example we chose one station (CE.24984) located near the intersection between W Exposition Blvd. and S 7th Ave. in Los Angeles. Fig. 3 shows the comparisons between the synthetics of most of the simulations in Table 1 and the recorded particle velocity during the 2008 Chino Hills earthquake for this particular station. The record, of course, is common to all pairs, and the number next to each set of seismograms correspond to the GOF score obtained for that comparison.

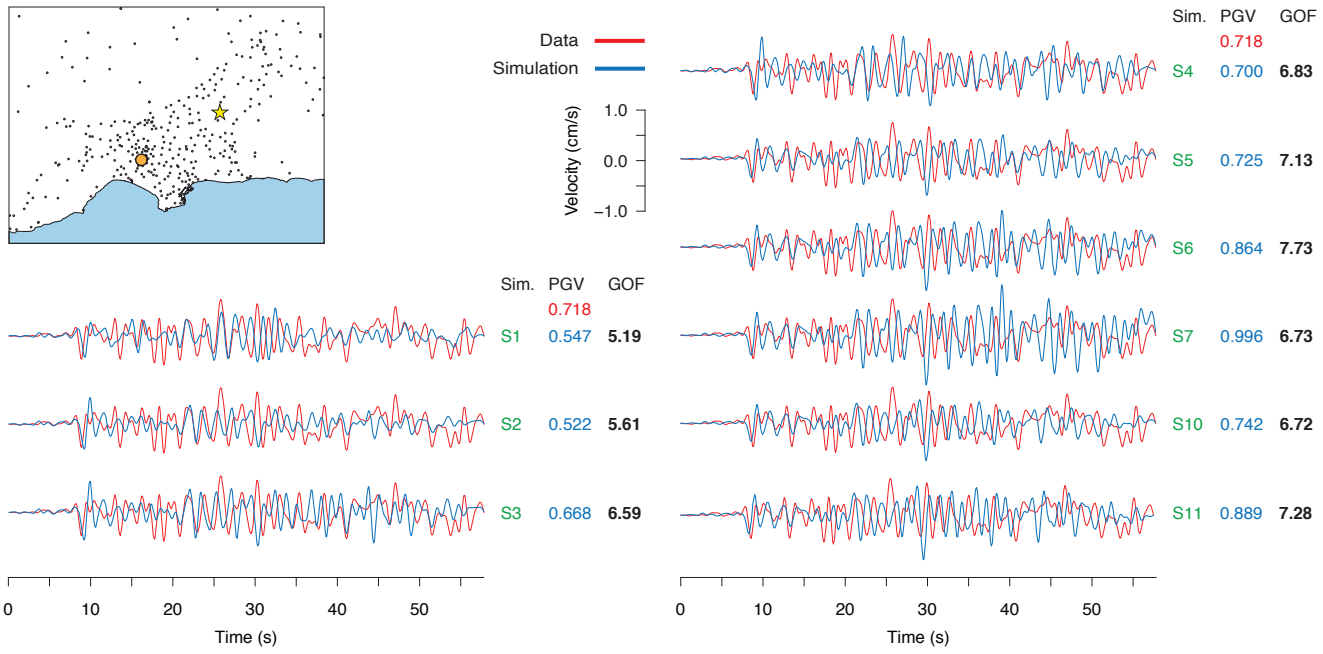

Fig. 3 – Comparison of synthetics for a subset of the simulations described in Table 1, for east-west component of motion at station CE.24984. The location of the station with respect to the earthquake epicenter and other stations is shown in the inset map on the top-left corner. The record at this station, which is common for all pairs, is shown in red, whereas the synthetics for each simulation are shown in blue. The simulation IDs are shown in green to the right of each pair, followed in blue by the peak ground velocity value (PGV) of the synthetics. At the top of the PGV values, that corresponding to the record is shown in red. The numbers in bold, to the right are the GOF scores obtained for each comparison.

As it can be seen from Fig. 3 for the station chosen here, the GOF scores closely adjust to the level of similarity between the synthetics and the record. The largest score of all, 7.73, corresponds in this case to simulation S6, which was obtained with the attenuation relationship $Q_S = 100V_S$ for frequency independent Q, using the viscoelastic model BKT3 and the velocity model CVM-S4.26. The GOF maps shown in Fig. 1 and other figures presented in the following section are composed combining the results for the three components at each station and then drawing a smoothed contour surface for the simulation domain. Fig. 4 shows the GOF maps obtained for most of the simulations described in Table 1. This figure clearly shows that the validation results (GOF maps) vary depending on each simulation, as similarly also implicit from Fig. 3. In other words, we show here how validation results exhibit a significant level of dependency on the modeling and simulation parameters, as is to be expected. In the following section we look at the differences between the various GOF maps in greater detail.

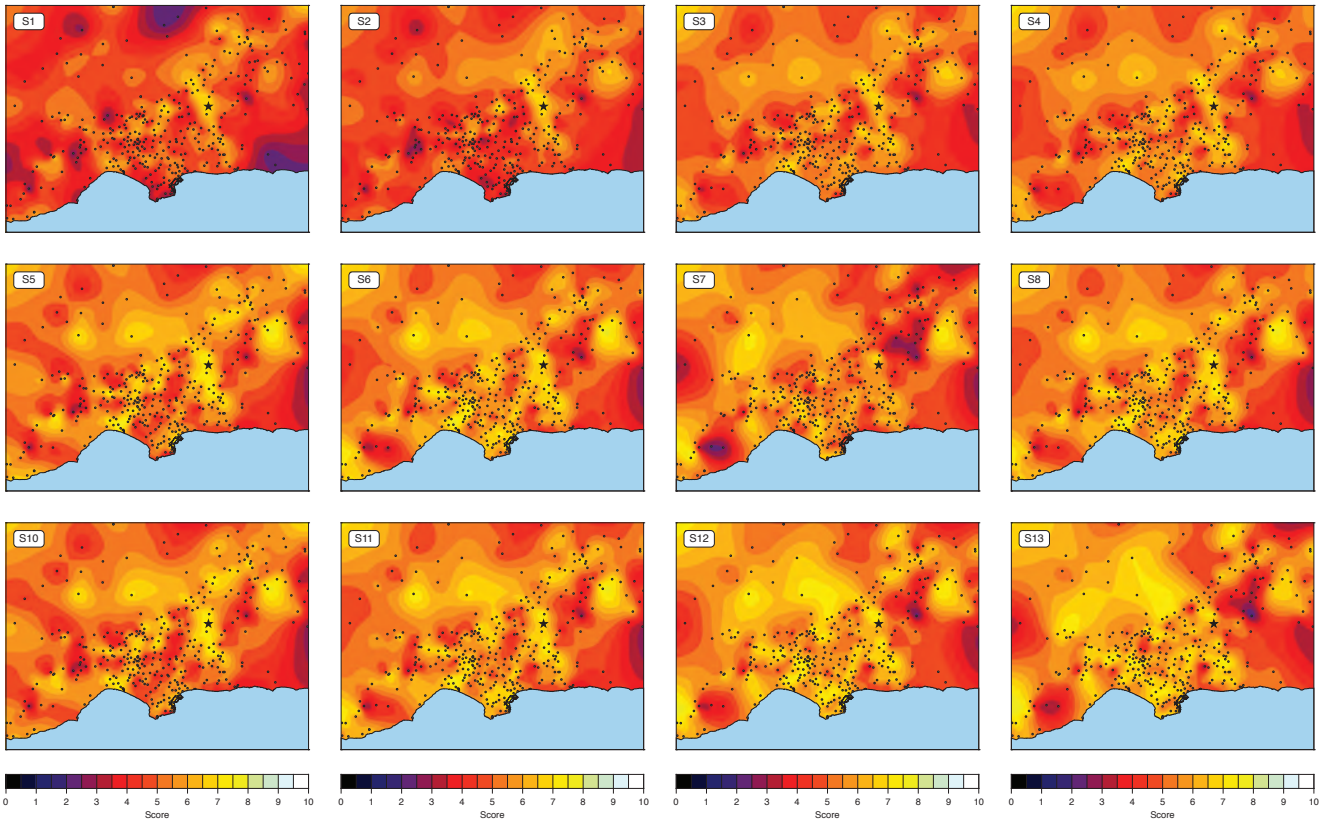

Fig. 4 – GOF maps obtained for the simulations described in Table 1 (S9 was omitted to facilitate the figure's arrangement).

5. Analysis and Discussion of Results

It is clear from Figs. 2–4 that the different models and parameters chosen for the simulations described in Table 1 affect the outcome of the validation. We are interested in disaggregating such effects to understand which weight more on the validation outcome. As mentioned before, in the first part of our work [17] we confirmed that the velocity model CVM-S4.26.M01 yields better results than CVM-S4. Here we introduce additional levels of complexity. We start by revisiting those results for the case of the 2008 Chino Hills earthquake and by exploring the effect of considering the soft deposit units in the velocity model.

Fig. 5 shows the residuals of GOF maps for S2 – S1 (read S2 minus S1), and S3 – S2. The former (left plot on Fig. 5) compares CVM-S4.26.M01 with respect to CVM-S4, while the latter (right plot in Fig. 5) compares the use of $V_{S\min}$ = 200 m/s with respect to the results obtained when using $V_{S\min}$ = 500 m/s. We observe that, as it was shown previously in Fig. 2, the model CVM-S4.26.M01 tends to perform better than CVM-S4, especially outside the Los Angeles basin and in rock sites along the San Gabriel and Santa Ana mountains, and the Mojave desert. These changes are, however, rather marginal. In the order of 0.5 to 1 GOF points. Regarding the consideration of low velocity profiles ($V_S \leq$ 500 m/s), we note that the simulation with $V_{S\min}$ = 200 m/s yield only marginally better GOF scores than those obtained with $V_{S\min}$ = 500 m/s. It is possible that the small effect noticed here is a consequence of the simulations being limited to $f_{\max}$ = 2 Hz. In other words, it is likely that low velocity profiles might contribute to better validation results for larger maximum frequencies.

Fig. 6 shows the residuals for the cases with different attenuation parameters. First, we focus on the empirical selection of Q values according to Eq. (3). The residuals S3 – S2 compare the use of α = 100 with respect to α = 50. Here, a larger value of α means less attenuation in the material as a linear function of stiffness (i.e. the larger the V_S, the larger the Q, then the lesser the attenuation). The result indicates that the validation results improve with the larger value of α by up to 2 and 3 GOF points, especially within the Los Angeles and

Simi basins, the San Fernando valley and the Oxnard plain, but also in parts of the San Gabriel mountains and Mojave desert.

In turn, the residual S6 – S5 in Fig. 6 shows that the viscoelastic model BKT3 performs somewhat better than the model BKT2 in most areas but not in parts of the San Gabriel and Santa Monica mountains, and in or around the San Bernardino basin area or close to the epicenter. In a parallel study [26] we show that the model BKT3 is better than the model BKT2 when it comes to approximating the attenuation effects to a constant, frequency-independent target Q_O. In the past this has been a well-accepted premise for simulations under 1 Hz, but it is also well-known that above 1 Hz, Q becomes frequency dependent. The model BKT3F adopts a frequency depend $Q(f)$ model following the relationships shown in Eqs. (1) and (2). The residuals S7 – S6 in Fig. 6 show the effect of using BKT3F with respect to BKT3. This comparison seems to suggest that the use of the frequency dependent model is counterproductive. However, it is also likely that the combination of $\alpha = 100$ with $Q(f)$ over relaxed the attenuation and that a combination of $\alpha = 50$ and $Q(f)$ may lead to better results. Other studies have already shown the benefit of using a frequency dependent attenuation model [27] and it has been suggested that Q may also be depth dependent. In addition, it is also known that other factors such as the presence of small-scale heterogeneities can affect the results [10], thus further research on the effects of the attenuation parameters will be required.

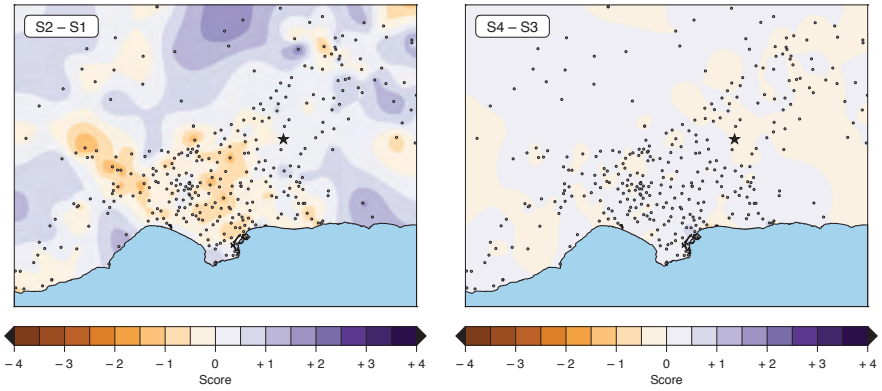

Fig. 5 – GOF residual maps obtained between simulations S1, S2, S3 and S4. Positive values indicate improvement in the validation results by the given number of points in the GOF scale, and negative values otherwise. The residuals on the left refer to using CVM-S4 versus CVM-S4.26.M01 (S2 – S1), and the ones on the right refer to using V_{Smin} values of 200 m/s instead of 500 m/s (S4 – S3).

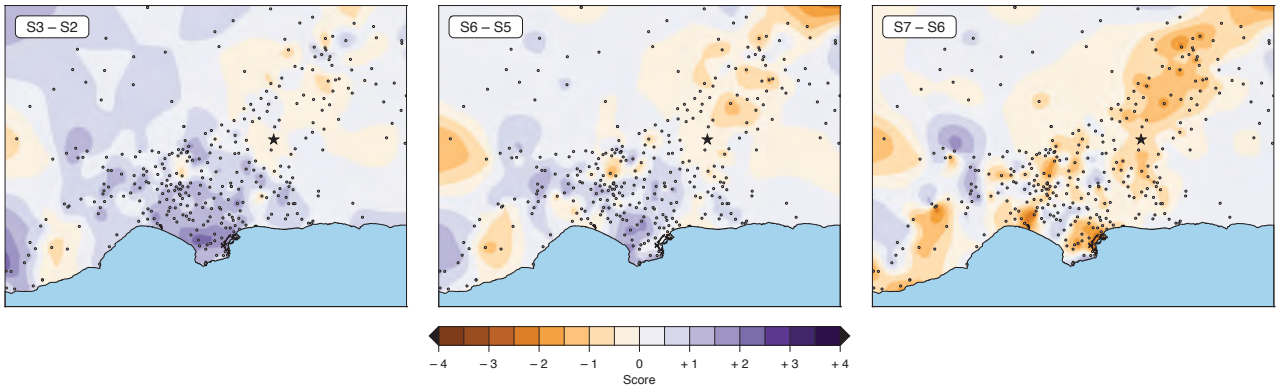

Fig. 6 – Similar to Fig. 5 but now showing GOF residual maps obtained between simulations S2 and S3, and S5, S6 and S7. Left: effect of using $Q_S = 50V_S$ versus $100V_S$ (S3 – S2). Center: effect of using the models BKT2 versus BKT3 under frequency independent attenuation (S6 – S5). Right: effect of using frequency dependent attenuation with an exponential slope of $\lambda = 0.8$ with respect to the case of frequency independent Q (S7 – S6).

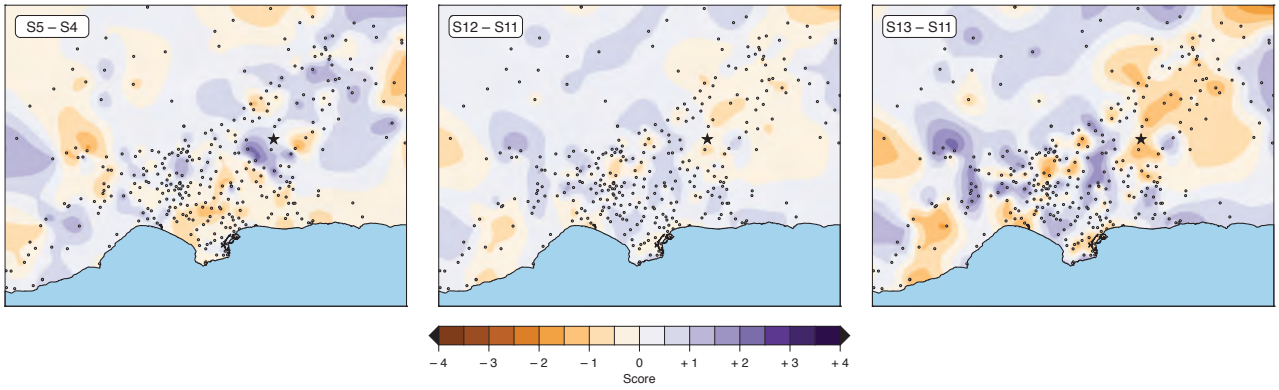

Fig. 7 – Similar to Fig. 6 but now showing GOF residual maps obtained between simulations S4 and S5, and S11, S12 and S13. Left: effect of using an extended source model instead of a point source (S5 – S4). Center: effect of scaling the magnitude of the earthquake to M_w 5.45 instead of 5.4 (S12 – S11). Right: effect of scaling the source to magnitude M_w 5.5 instead of 5.4 (S13 – S11).

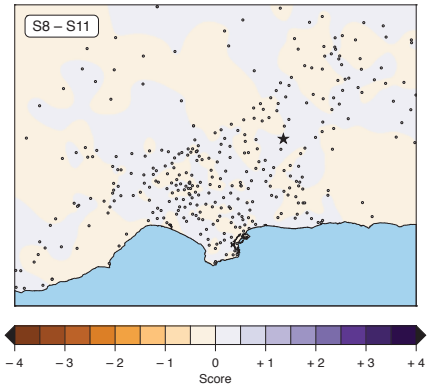

Fig. 8 – Similar to Fig. 7 but now showing GOF residual maps obtained S8 – S11. These residuals show the effect of using double the number of points per wavelength in the generation of the finite element mesh.

In Fig. 7 we analyze the effects of the source model. The residuals in S5 – S4 correspond to the effect of using a finite fault model over a point source model. Here the results are somewhat mixed and without a definite pattern. Perhaps the most notable effect is near the source, where we see an improvement in the GOF values, especially west of the epicenter. This is to be expected given that as epicentral distance increases, the source will be increasingly seen as a point by the farther stations, yet we observe that even for some far stations the extended source model has an effect. We also considered the possibility of a certain level of uncertainty in the source magnitude and ran simulations S12 and S13, for which the slips in the extended source model were scales to produce magnitudes M_w 5.45 and 5.50, respectively, as opposed to the official M_w 5.4 of the 2008 Chino Hills earthquake. The outcome of the residuals S12 – S11 and S13 – S11 suggest that for this particular case, the validations may be benefited from consider a slightly large magnitude source model than the official moment tensor solution. These residuals show significant improvements of the order of 1 and 2 GOF points in good portions of the Los Angeles basin, the Long Beach and Santa Ana mountains area, Simi and San Fernando valleys and the San Gabriel mountains and the Mojave desert.

Last, we consider a matter of numerical accuracy. Most simulations reported in Table 1 were done using finite-element meshes tailored to satisfy a minimum of 10 point per wavelength in space. We, however, conducted a couple of simulations (S8 and S9) with meshes tailored to satisfy a minimum of 20 points per wavelength. Fig. 8 shows the residuals of one of these 20-points runs with respect to its corresponding equivalent 10-points run. This is the residual S8 – S11. It can be seen that these residuals in Fig. 8 are not as clearly determinant as others shown before, ant they suggest that the numerical accuracy obtained with 10 points

is as good as that obtained with 20 points, at least when it comes down to validation. While we know the simulation with 20 points is certainly more accurate, in our experience, 10 points per wavelength has usually performed well with the type of tri-linear cubic 8-node finite elements implemented in Hercules when it comes to heterogeneous domains. On the other hand, more homogeneous domains—where inaccuracies have more "space to grow"—seem to require larger number of points per wavelength, but this is an aspect that requires careful evaluation for each problem as it may be sensitive to the geometry of the model and the frequency content of the excitation. We did not run simulations with varying time-discretization parameters. The combined influence of both the space and time discretization parameters is a subject that may also require further analysis but was out of the scope of this article.

6. Concluding Remarks

Upon performing a collection of simulations with varying modeling and simulation parameters for the case of the 29 July 2008 M_w 5.4 Chino Hills, California, earthquake, we have found that the results obtained from quantitative validation analysis using well-known goodness-of-fit (GOF) metrics are significantly sensitive to the parameters used to prepare the simulations. Additional (preliminary) statistical analysis ran on the results shown in this article but not included here for brevity indicate that the GOF scores within any given simulation tend to exhibit standard deviations of about ±1 GOF points. This means that under any given conditions, the GOF scores may oscillate within a range of about 2-points width. Considering the GOF scale used here categorizes the quality of the fit from poor to excellent in bins of 2-points width (< 4, 4 to 6, 6 to 8, and 8 to 10), knowing this variability helps to constrain the level of epistemic uncertainty on the parameters used for (physics-based) ground simulation. Future efforts from our research group are intended to better characterize this uncertainty in validation results by performing similar analysis to that shown here for multiple events.

7. Acknowledgements

This work was supported by the United States Geological Survey (USGS) through the award G14AP00034, and by the Southern California Earthquake Center (SCEC) through the award 14032. Additional support was provided through the U.S. National Science Foundation (NSF) awards ACI-1148493, and ACI-1450451. This research is also part of the Blue Waters sustained-petascale computing project, which is supported by NSF (awards OCI-0725070 and ACI-1238993) and the state of Illinois. Blue Waters is a joint effort of the University of Illinois at Urbana-Champaign and the National Center for Supercomputing Applications (NCSA). Computational support was possible through PRAC allocations supported by NSF awards ACI-0832698 and ACI-1440085. SCEC is funded by NSF Cooperative Agreement EAR-1033462 and USGS Cooperative Agreement G12AC20038. The SCEC contribution number for this paper is 6252.

8. References

[1] Graves RW (1996): Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences. *Bulletin of the Seismological Society of America*, **86** (4): 1091–1106.

[2] Bao H, Bielak J, Ghattas O, Kallivokas LF, O'Hallaron DR, Shewchuk JR, Xu J (1998): Large-scale simulation of elastic wave propagation in heterogeneous media on parallel computers. *Computational Methods in Applied Mechanics and Engineering*, **152** (1-2): 85–102.

[3] Aagaard BT, Brocher TM, Dolenc D, Dreger D, Graves RW, Harmsen S, Hartzell S, Larsen S, McCandless K, Nilsson S, Petersson NA, Rodgers A, Sjogreen B, Zoback ML (2008): Ground-motion modeling of the 1906 San Francisco earthquake, Part II: Ground-motion estimates for the 1906 earthquake and scenario events. *Bulletin of the Seismological Society of America*, **98** (2): 1012–1046.

[4] Olsen KB, Day SM, Dalaguer LA, Mayhew J, Cui Y, Zhu J, Cruz-Atienza VM, Roten D, Maechling P, Jordan TH, Okaya D, Chourasia A (2009): ShakeOut-D: Ground motion estimates using an ensemble of large earthquakes on the southern San Andreas fault with spontaneous rupture propagation. *Geophysical Research Letters*, **36** L04303.

[5] Kim K, Elgamal A, Petropoulos G, Askan A, Bielak J, Fenves GL (2014): Seismic response of a large-scale highway interchange system. *Earthquake Geotechnical Engineering Design*, Maugeri M, Soccodato C (Eds.), Geotechnical, Geological and Earthquake Engineering, **28**: 223-240.

[6] Isbiliroglu Y, Taborda R, Bielak J (2015): Coupled soil-structure interaction effects of building clusters during earthquakes. *Earthquake Spectra* **31** (1): 463–500.

[7] 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*, **168** (3-4): 367–381.

[8] Taborda R, Bielak J, Restrepo D (2012): Earthquake ground motion simulation including nonlinear soil effects under idealized conditions with application to two case studies. *Seismological Research Letters*, **83** (6): 1047–1060.

[9] Shi Z, Day SM (2013): Rupture dynamics and ground motion from 3-D rough-fault simulations. *Journal of Geophysical Research*, **118** (3): 1122–1141.

[10] Withers KB, Olsen KB, Day SM (2013): Deterministic high-frequency ground motion using dynamic rupture along rough faults, small-scale media heterogeneities, and frequency-dependent attenuation. *Annual Meeting of the Southern California Earthquake Center*, Palm Springs, USA.

[11] Bielak J, Graves RW, Olsen KB, Taborda R, Ramírez-Guzmán L, Day SM, Ely GP, Roten D, Jordan TH, Maechling PJ, Urbanic J, Cui Y, Juve G (2010): The ShakeOut earthquake scenario: Verification of three simulation sets. *Geophysical Journal International*, **180** (1): 375–404.

[12] Chaljub E, Moczo P, Tsuno S, Bard P-Y, Kristek J, Kaser M, Stupazzini M, Kristekova M (2010): Quantitative comparison of four numerical predictions of 3D ground motion in the Grenoble valley, France. *Bulletin of the Seismological Society of America*, **100** (4): 1427–1455.

[13] Taborda R, Bielak J (2013): Ground-motion simulation and validation of the 2008 Chino Hills, California, earthquake. *Bulletin of the Seismological Society of America*, 103 (1): 131–156.

[14] Taborda R, Bielak J (2014): Ground-motion simulation and validation of the 2008 Chino Hills, California, earthquake using different velocity models, *Bulletin of the Seismological Society of America*, **104** (4): 1876–1898.

[15] Tu T, Yu H, Ramírez-Guzmán L, Bielak J, Ghattas O, Ma K-L, O'Hallaron DR (2006): From mesh generation to scientific visualization: An end-to-end approach to parallel supercomputing. *SC'06: ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis*, IEEE Computer Society, 15 pp.

[16] Taborda R, López J, Karaoglu H, Urbanic J, Bielak J. (2010): Speeding up finite element wave propagation for large-scale earthquake simulations. *Technical Report CMU-PDL-10-109*, Parallel Data Lab, Carnegie Mellon University, Pittsburgh, USA.

[17] Taborda R, Azizzadeh-Roodpish S, Khoshnevis N, Cheng K (2016): Evaluation of the southern California seismic velocity models through simulation of recorded events. *Geophysical Journal International*, **205** (3): 1342–1364.

[18] Anderson J (2004): Quantitative measure of the goodness-of-fit of synthetic seismograms. *13th World Conference on Earthquake Engineering*, Vancouver, British Columbia, Canada.

[19] Roten D, Olsen KB, Day SM, Cui Y, Fäh D (2014): Expected seismic shaking in Los Angeles reduced by San Andreas fault zone plasticity. *Geophysical Research Letters*, **41** (8): 2769–2777.

[20] Schlosser SW, Ryan MP, Taborda R, López J, O'Hallaron D, Bielak J (2008): Materialized community ground models for large-scale earthquake simulation. *SC'08: ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis*, IEEE Computer Society, 11 pp.

[21] Small P, Maechling P, Jordan T, Ely G, Taborda R (2011): SCEC UCVM—Unified California velocity model, *AGU Fall Meeting*, Abstract No. S21B-2200, San Francisco, USA.

[22] Kohler MD, Magistrale H, Clayton RW (2003): Mantle heterogeneities and the SCEC reference three-dimensional seismic velocity model version 3. *Bulletin of the Seismological Society of America*, **93** (2): 757–774.

[23] Lee E-J, Chen P, Jordan TH, Maechling PJ, Denolle M, Beroza GC (2014): Full-3-D tomography for crustal structure in southern California based on the scattering-integral and the adjoint-wavefield methods. *Journal of Geophysical Research*, **119** (8): 6421–6451.

[24] Plesch A, Tape C, Graves R, Shaw J, Small P, Ely G (2011): Updates for the CVM-H including new representations of the offshore Santa Maria and San Bernardino basin and a new Moho surface. *Annual Meeting of the Southern California Earthquake Center*, Palm Springs, USA.

[25] Bielak J, Karaoglu H, Taborda R (2011): Memory-efficient displacement-based internal friction for wave propagation simulation. *Geophysics*, **76** (6): T131–T145.

[26] Taborda R, Huda MM, Khoshnevis N, Bielak J (2017). An adaptive memory-efficient displacement-based internal friction model for wave propagation simulation. *Geophysics*, in preparation.

[27] Withers KB, Olsen KB, Day SM (2015): Memory-efficient simulation of frequency-Dependent *Q*. *Bulletin of the Seismological Society of America*, **105** (6): 3129–3142.

[28] Shao G, Ji C, Hauksson E (2012): Rupture process and energy budget of the 29 July 2008 M_w 5.4 Chino Hills, California, earthquake. *Journal of Geophysical Research*, **117**: B07307.