Geophys. J. Int. (0000) **000**, 1–4

Toward a Framework for Ground Motion Simulation Validation using Attenuation Relationships. Part1: Calibration Between NGA-West2 Predictions, Physics-Based Synthetics, and Data

Ricardo Taborda,[1,2]⋆ Shima Azizzadeh-Roodpish,[1,2]
and Naeem Khoshnevis,[2]

[1] *Department of Civil Engineering, University of Memphis, Memphis, TN* 38152, *USA*
[2] *Center for Earthquake Research and Information, University of Memphis, Memphis, TN* 38152, *USA*

To be submitted by March, 2016

1 INTRODUCTION

Increasing advances in ground motion simulation emboldens researches to seek for reliable methods of simulation validation. Validation is an effort to check the correctness of the formulation and the accuracy with respect to observations. The concept of simulation has found its way from systems science (e.g. Sargent, 2005; Oberkampf et al., 2004) to earthquake simulation community (e.g., Bielak et al.,2010; Chaljub et al., 2010). Up untill now, validaton of ground motion simulations is done by comparing synthetic seismograms against observations from past earthquakes (Olsen and Mayhew, 2010; Taborda and Bielak, 2013 and 2014, Taborda et al., 2016) predominantly using quantitative methods known as goodness-of-fit (GOF) or misfit criteria (**?**); Olsen and Mayhew, 2010; Kristekova et al., 2009). These validations are concentrating comparisions to the availability of data, which for most of the past events is limited to a reduced number of locations.

However, the simulations are resulved for the whole computational domain, not just the specific stations, and this provide us with a complete surface datsets which can be used to build simulation specific attenuation curves comparable to well-established attenuation relationships.The goal of this project is to shape a new simulation validition framework based on comparision with attenuation relationships. To this end, we start with the available recorded data for southern california region from past earthquakes and synthetics from four different velocity model to define and calibrate a goodness of fit criteria using their corrisponding attenuation relationships. Later we will extend the findings to the whole simulation domain and will perform the evaluation between NGA-West2 prediction and our physic-based synthetics.

2 DATA AND SYNTHETIC

Considered Data for conducting this study contains thirty moderate-magnitude earthquakes ($3.5 > M_w > 5.5$) occurred between 1998 and 2014. The latest version of ground motion prediction equations (GMPEs) database includes small-to-moderate earthquakes in California which will be helpful in calibration of compariosion criteria. Selected events are spread throughout a simulation domain with a volume projection of 180 km × 135 km × 62 km, which covers the entire Los Angeles metropolitan area and most of the significant geologic structures around. Fig. 1 shows the simlation domain in which we labled the events with sequential letter code from A to Z and AA to AD and Table 1 provides detailed information for those selected events.The unprocessed recorded data from Southern California Earthquake Data Center (SCEDC) and Center for Engineering Strong Motion Data (CESMD) are downloaded for each of the earthquakes to take advantage of the numerous time series available from different data centers. We obtained records for moree than 800 stations, Some of which were in common between different networks and some were not usable. Records from SDEDC and CESMD were processed and selected for each event. we performed gain and baseline corrections, and applied a high-pass filter at 0.05 Hz before integrating to obtain velocities and displacements. The final number of chosen records for validation for any of the selected earthquake can be found in Table 1.

The simulations of the selected events are all done for the four major velocity models currently available for the region of southern California: CVM-S4, CVM-S4.26, CVM-H and CVM-H+GTL. 3D simulations with point source model are done at maximum frequency of 1 Hz and minumum shear wave velocity of 200 m/s for each earthquake and velocity model combination using Hercules, a parallel 3D finite element computer application for solving forward anelastic wave propagation problems which has been thoroughly tested and verified in multiple supercomputers. All the runs are done on Blue Waters at the National Center for Supercomputing Applications.

Using a developed python code, we collect the results of the Peak Ground Acceleration (PGA), Peak Ground Velocity (PGV), and Peak Ground Displacement (PGD) in three directions (EW, NS, Up), for all the events in all available stations for both synthetics and data. This information will be used later in validation method.

3 EVALUATION METHOD

Attenuation relationship relates the distance from source to the ground motion features such as PGA, PGV and PGD. Plotting this data for different stations in the logarithmic scale for different cases suggest that attenuation relationship can initially be approximated by a single line. So we will use linear regression (straight line-first order polynomial degree regression model) for both data and synthetics. b is the slope of the regression line, ($y = a + bx + \epsilon$, being ϵ the error rodisturbance term associated to that model and a the

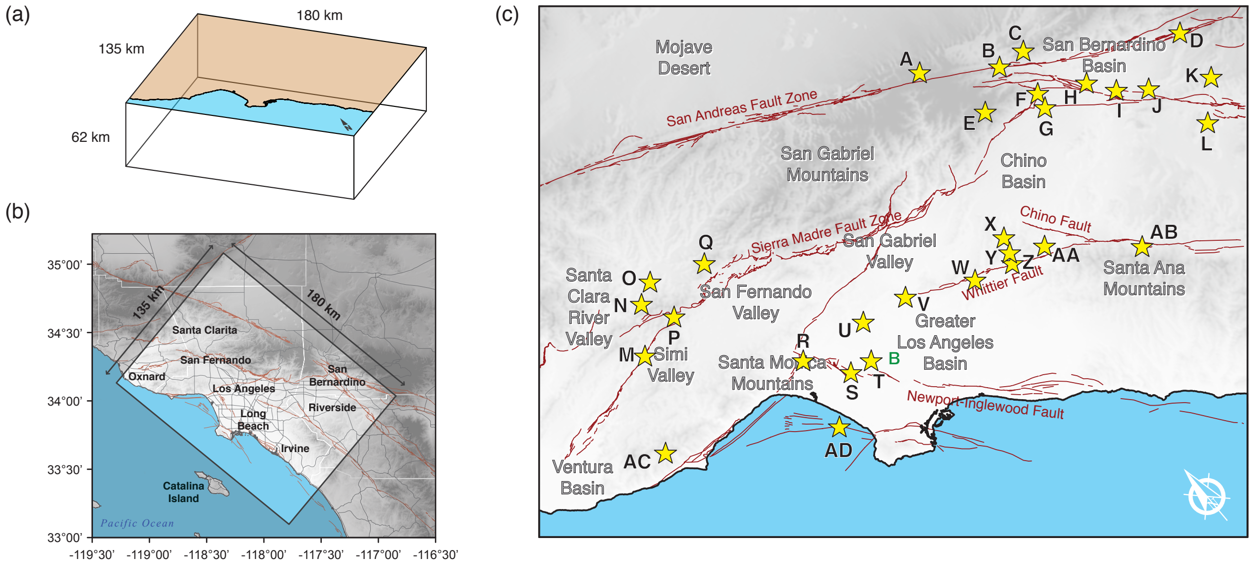

Figure 1. Region of interest and simulation domain. (a) 3D view of the simulation domain. (b) Geographical location and surface projection of the simulation domain, along with the names of the main cities surrounding the Los Angeles metropolitan are. (c) Major geologic structures including basins, valleys and mountains, along with the main quaternary faults in the region. The color background represents the surface V_{S30} values included in the CVM-H+GTL model, with a topography shading effect.

Table 1. Selected events with their ID **?**) and description of source location, magnitude, focal mechanism (strike, dip, and rake angles) **?**), and date and time (in coordinated universal time, or UTC).

Code	Earthquake name	Event ID	M_w	Coordinates (lon., lat.)	Depth (km)	Strike/Dip/Rake	Date (yyyy/mm/dd)	UTC Time (hh:mm:ss)	Num. Stns.
A	Wrightwood	9064568	4.40	–117.6480, 34.3740	8.99	285/57/86	1998/08/20	23:49:58.198	17
B	NW of Devore	10972299	3.79	–117.4642, 34.2655	10.91	98/58/68	2001/07/19	20:42:36.470	52
C	NNE of Devore	14494128	3.72	–117.3838, 34.2587	7.18	344/69/-33	2009/08/01	12:55:55.317	77
D	Yucaipa	14155260	4.88	–117.0113, 34.0580	11.61	75/59/55	2005/06/16	20:53:26.225	172
E	N of Rancho Cucamonga	10216101	3.60	–117.5762, 34.2058	4.92	54/69/16	2006/11/04	19:43:44.376	55
F	2002 Fontana	13692644	3.74	–117.4288, 34.1613	6.54	233/72/-28	2002/07/25	00:43:14.872	55
G	2005 Fontana	14116972	4.42	–117.4387, 34.1250	4.15	222/88/-25	2005/01/06	14:35:27.593	83
H	San Bernardino	10370141	4.45	–117.3042, 34.1073	14.22	87/70/28	2009/01/09	03:49:46.051	159
I	N of Loma Linda	9140050	4.37	–117.2525, 34.0500	15.36	270/90/-6	2000/02/21	13:49:43.017	38
J	Redlands	10541957	4.10	–117.1797, 34.0045	8.53	33/46/-68	2010/02/13	21:39:06.349	97
K	2010 Beaumont	10530013	4.28	–117.0232, 33.9322	13.93	234/89/9	2010/01/16	12:03:25.345	76
L	2006 Beaumont	14239184	3.90	–117.1122, 33.8560	11.53	45/31/-25	2006/07/10	02:54:43.809	66
M	Simi Valley	14000376	3.59	–118.7530, 34.2722	13.81	234/62/60	2003/10/29	23:44:48.206	54
N	WSW of Valencia	9753489	3.90	–118.6678, 34.3705	14.21	83/62/57	2002/01/29	06:00:39.140	52
O	N of Pico Canyon	9096972	3.98	–118.6090, 34.3980	11.53	287/55/54	1999/07/22	09:57:23.502	26
P	Chatsworth	14312160	4.66	–118.6195, 34.2995	7.58	82/27/51	2007/08/09	07:58:48.888	109
Q	Newhall	15237281	3.86	–118.4580, 34.3508	3.59	236/58/33	2012/10/28	15:24:23.172	120
R	Beverly Hills	9703873	4.24	–118.3885, 34.0590	7.90	262/81/4	2001/09/09	23:59:17.695	130
S	Inglewood Area	10410337	4.70	–118.3357, 33.9377	13.86	243/60/25	2009/05/18	03:39:36.126	213
T	NW of Compton	9716853	3.98	–118.2702, 33.9290	21.13	116/68/71	2001/10/28	16:27:45.388	55
U	Downtown Los Angeles	9093975	3.77	–118.2180, 34.0100	9.53	125/49/79	1999/06/29	12:55:00.371	25
V	Whittier Narrows	14601172	4.44	–118.0817, 33.9923	18.85	282/36/73	2010/03/16	11:04:00.026	180
W	La Habra	15481673	5.10	–117.9300, 33.9220	5.00	239/70/38	2014/03/29	04:09:42.970	311
X	Chino Hills	14383980	5.39	–117.7613, 33.9530	14.70	47/51/32	2008/07/29	18:42:15.960	335
Y	2002 Yorba Linda	9818433	4.75	–117.7758, 33.9173	12.92	34/84/-10	2002/09/03	07:08:51.675	67
Z	2009 Yorba Linda	10399889	3.98	–117.7892, 33.8940	4.23	208/65/26	2009/04/24	03:27:49.840	91
AA	ESE of Yorba Linda	9644101	3.64	–117.6882, 33.8777	3.59	56/65/37	2001/04/13	11:50:11.916	53
AB	Lake Elsinore	10275733	4.73	–117.4770, 33.7322	12.60	65/59/58	2007/09/02	17:29:14.827	116
AC	Westlake Village	10403777	4.42	–118.8825, 34.0667	14.17	254/73/30	2009/05/02	01:11:13.084	94
AD	Hermosa Beach	14738436	3.69	–118.4578, 33.8572	11.23	57/41/54	2010/06/07	23:59:27.165	93

intercept), which can be calculated by equation (1). Later, if necessary, we can expand this concept to a bilinear approximation for a better fit.

$$\mathrm{b} = \left(\frac{\sum\limits_{i=1}^{n}((x_i - \bar{x})(y_i - \bar{y}))}{\sum\limits_{i=1}^{n}(x_i - \bar{x})^2} \right) . \tag{1}$$

For camparing two lines, there are two components to be considered. The distance of two lines (we call it "Amplitude score") and their slope (we call it "Rate score").The evaluation process proposed here is based on defining a goodness-of-fit criteria for these two aspects.

For comparing the difference between the amount of peak ground characteristic of data and synthetics at each station, the goodness-of-fit (GOF) criteria (peak acceleration (C5), peak velocity (C6), peak displacement (C7)) proposed by **?**) is used. For these GOF scale, Each parameter is mapped onto a numerical scale ranging from 0 to 10, with 0 for the worsth and 10 for the best match between two signals. Any of the C5, C6 and C7 score can be calculated by equation (2).

$$\mathrm{S(p_1, p_2)} = 10exp\left(-(\frac{p_1 - p_2}{min(p_1, p_2)})^2\right) . \tag{2}$$

There are many benefits in using this function for calculating score. The function is monotonically decreasing as the difference between the parameters increase. It is symmetrical and because of that the score is similar regrdless of the order of bigger or smaller values. Multiplying the factor of 10 puts the score into a comfortable range of value for giving a score between 0 and 10. Small differences are not penalized too severely since it is an exponential function. We calculate this GOf at each station and find their average and assign that amount as the "amplitude score".

To define a GOF score for the differences in the slope of two lines, There are some statistical approach available. We suggest to use the Student's t-test based on the standard error of regression models according to an article by Andre and Estevez-Perez (2014). The advantages of using t-test is that this parameter not only considers the different slopes, but also is affected by the number of available data and synthetics and the standard deviation of the ground motion parameters. Generally, when there is a goofd fit of slope, the t has lower amount. Lower number of stations and higher stdv will increse the value of t. This will serve as a good factor for defining a GOF function for slope comparision. For that, certain amount of judgment will be needed to define the tren and to choose the necessary levels which should not be considered useful fits for engineering applications.

Most t-test statistics can be formulated as $t_{experimetal} = \frac{(\hat{\theta} - \theta)}{SE}$, being θ a population parameter, $\hat{\theta}$ an estimator of θ and SE the standard error of the estimator (or equivalently, an estimation of the standard deviation of the estimator). The test statistic is equation (3).

$$\mathrm{t} = \left(\frac{b_1 - b_2}{s_{b_1 - b_2}} \right) . \tag{3}$$

where

$$\mathrm{s_{b_1 - b_2}} = \left(s_{Res}\sqrt{\frac{1}{s_{x_1}^2(n_1 - 1)} + \frac{1}{s_{x_2}^2(n_2 - 1)}} \right) . \tag{4}$$

being n_1 and n_2 the sample sizes for each sample datta and s_{x_1} and s_{x_2} standard deviations. s_{Res} is a unique estimator which is a weighted average of two variances (alos known as s_{pool} since by that we can pool the estimates of the error variances, weightening each by their degrees of freedom) and can be calculated from equation (5).

$$\mathrm{s_{Res}^2} = \left(\frac{(n_1 - 2)s_{y,x_1}^2 + (n_2 - 2)s_{y,x_2}^2}{(n_1 - 2) + (n_2 - 2)} \right) . \tag{5}$$

in that, s_{y,x_1} and s_{y,x_2} are the residual variance (often known as squared standard error of the regression), which estimates the variance of the regression or variance of the model from the experimental data. Using above euation we can find the rate score of each pair of regression lines.

There can be different approaches in the way of combination of amplitude and rate score. We can use simple average or also give weight to each score and calculate a weighted average. In this report, we simply considered the total average as simple average of two calculated scores.

4 RESULTS

Here, We have the result of quantitative measures to assess the level of agreement between the data and the simulation, using amplitude and slope parameters. A python program is developed to produce plots of all the available pairs of data and synthetics attenuation relationship comparision, and also to compare all four velocity models together.Fig. 2 shows results of PGV in total (average of 3 directions) for all 4 models as an example. As it can be seen, the four models have close regression lines in most of the cases.

For the first set of trials the t prameter was calculated. Then, considering the variations in the t value ranges, an initial score is proposed corresponding to each t. Current selected ranges are: 10 if $0 \leq t \leq 0.1$, 9 if $0.1 < t \leq 1$, 8 if $1 < t \leq 10$, 7 if $10 < t \leq 20$, 6 if $20 < t \leq 100$ and 5 if $100 < t \leq 500$. Later it will be an acceptable function of t for calculating score instead of just assigning the score for different t ranges. This will make the GOF more accurate and applicable.

In order to better see the changes in amplitude score, rate score, and total score using a python program plots for the trend of changes in all 30 events are produced. Here some of them are discussed with the aim of comparing four different velocity models. Fig. 3 shows trend of amplitude score for PGD, PGV and PGA for all 4 models. In all of them the same trend is visible. PGD has better scores, then PGV, then PGA. Among all cvms426-223 has better overall result.

Fig. 4 shows trend of rate score for PGD, PGV and PGA for all 4 models. PGD still has better scores then PGV, then PGA. Bur we can see different trends here.

Fig. 5 shows trend of total average score for PGD, PGV and PGA for all 4 models. Again, the trend is similar, the score is better in PGD, then PGV, then PGA, and CVMS426-223 shows a better overall results.

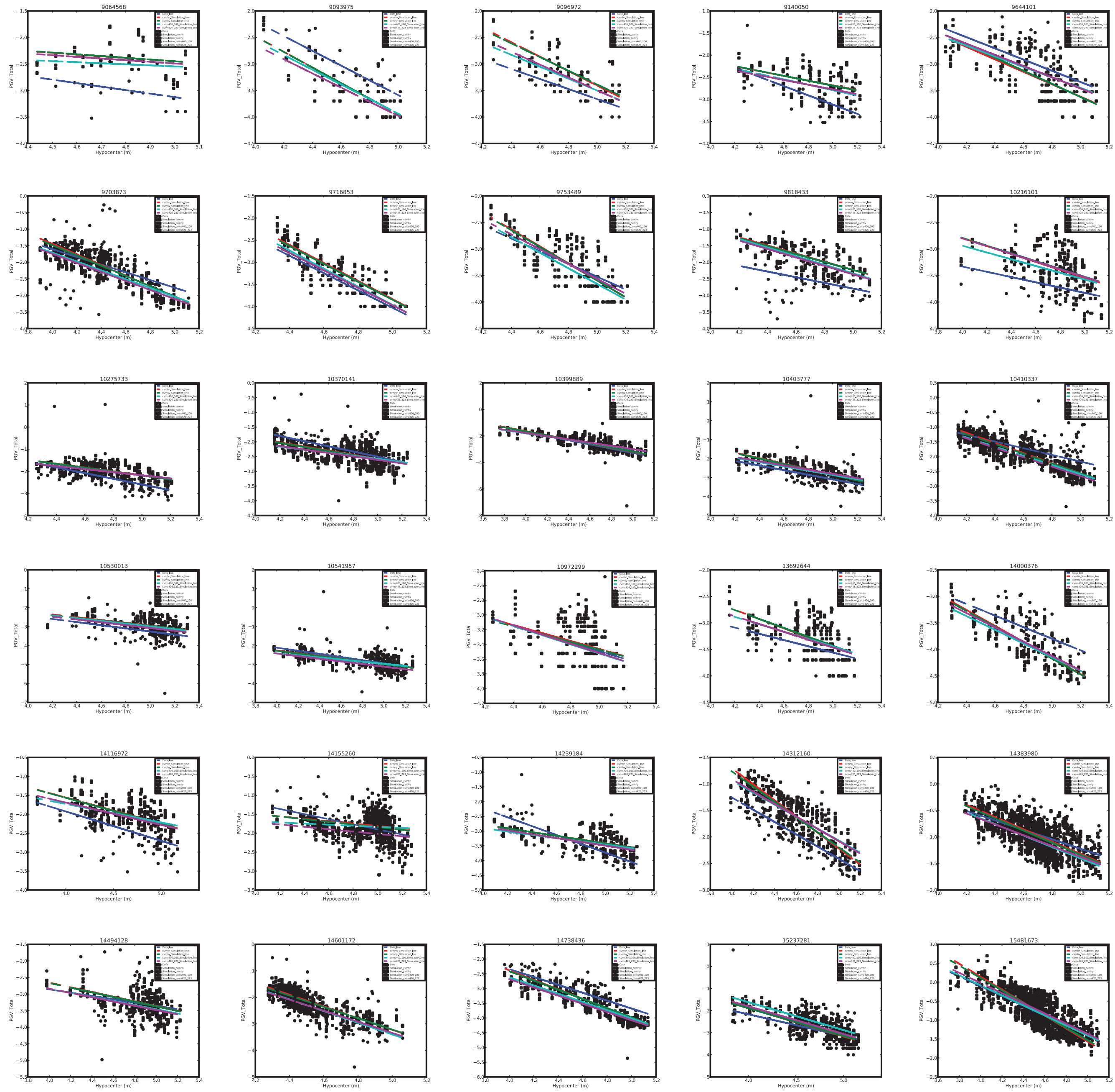

Figure 2. PGV for all the 30 events and all four models are presented.

5 FURTHER WORKS

To improve the framework, here are some suggestions for the next steps. As it mentioned before, we can try to define the score directly according to the t value. So some different function will be tested to find an acceptable funcion. Combination of amplitude and rate score can be done by simple average or by assigning weigh to each of the scores corrisponding to its importance. So that is something to work on. Huge amount of data is available from the simulation domain area and we want to do the validation using all of those data. Similar procedure can be use with some modifications.The comparisopns will be the same, except that it might be neceessary to select two different approximated lines for two different segments of the data and synthetics to get a more realistic approximation and more accurate results. In that case, the scores will be calculated for each segment separetely and then they will be combined with a simple or weighted average. In addition, with having many synthetic points, the amplitude score, instead of for each station, can be calculated in selected steps for the approximated line. The next step would be to check the results of synthetics with NGA predictions.

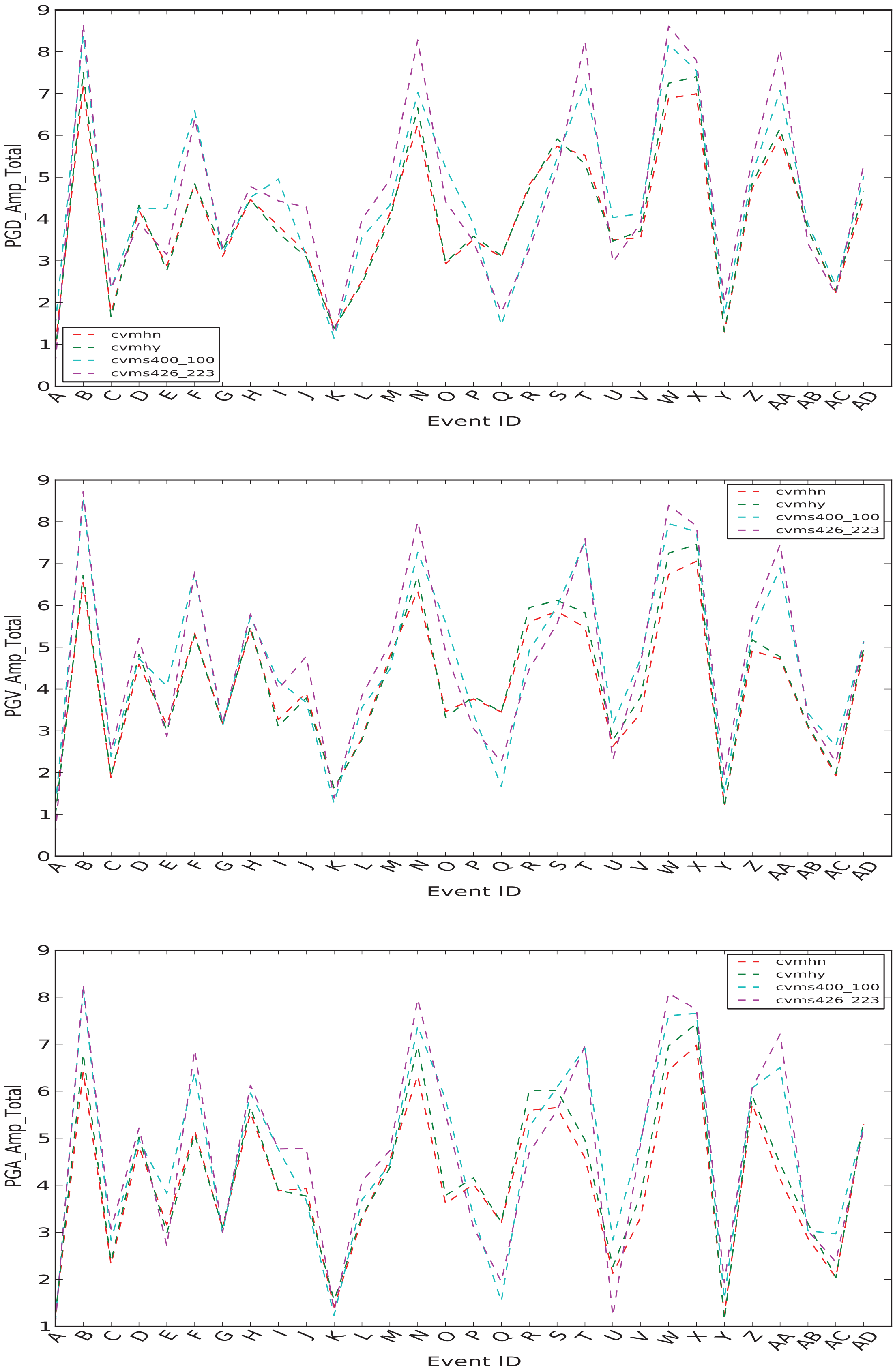

Figure 3. PGD, PGV and PGA amplitude score for all the 30 events and all four models are presented.

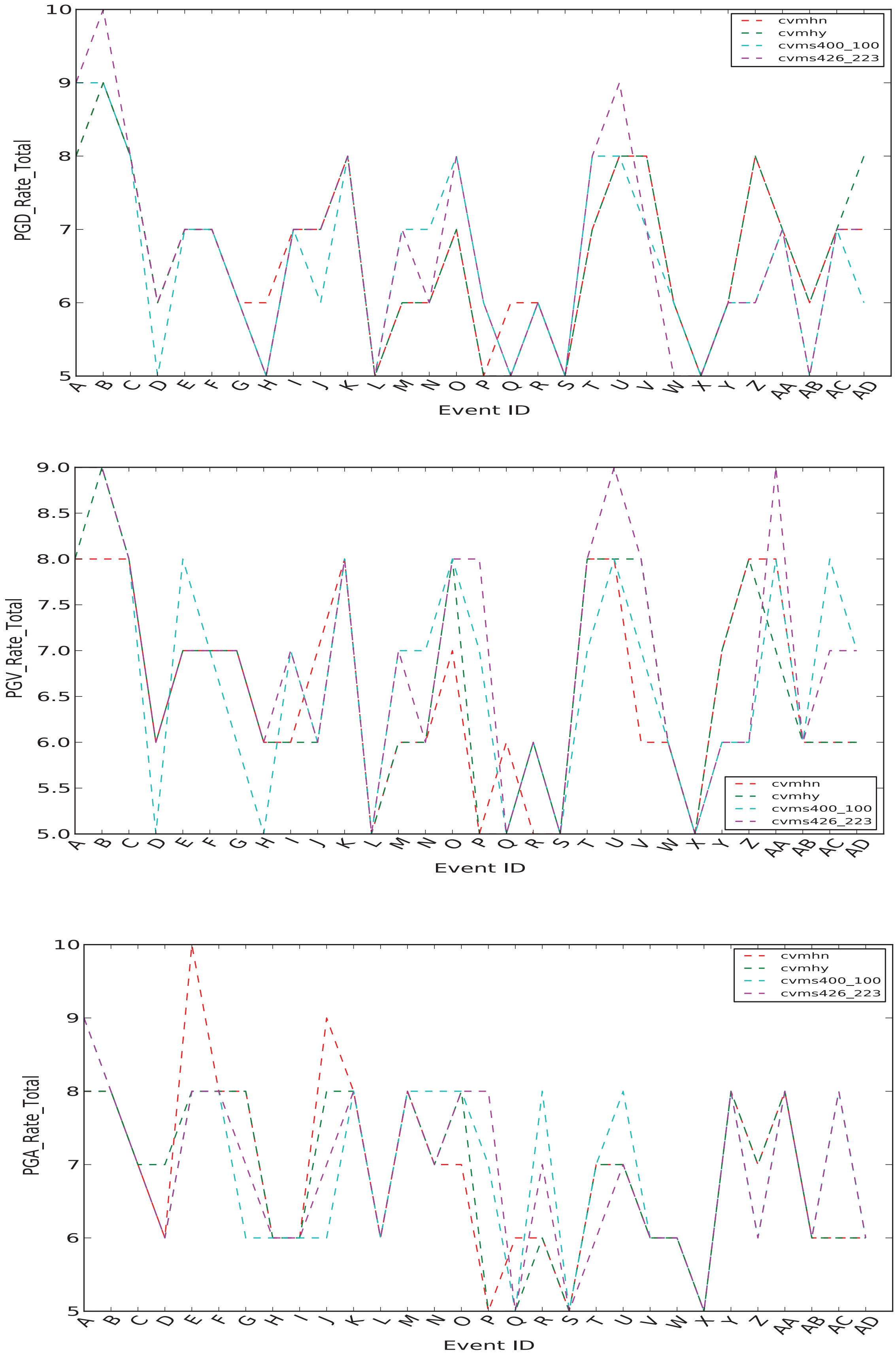

Figure 4. PGD, PGV and PGA rate score for all the 30 events and all four models are presented.

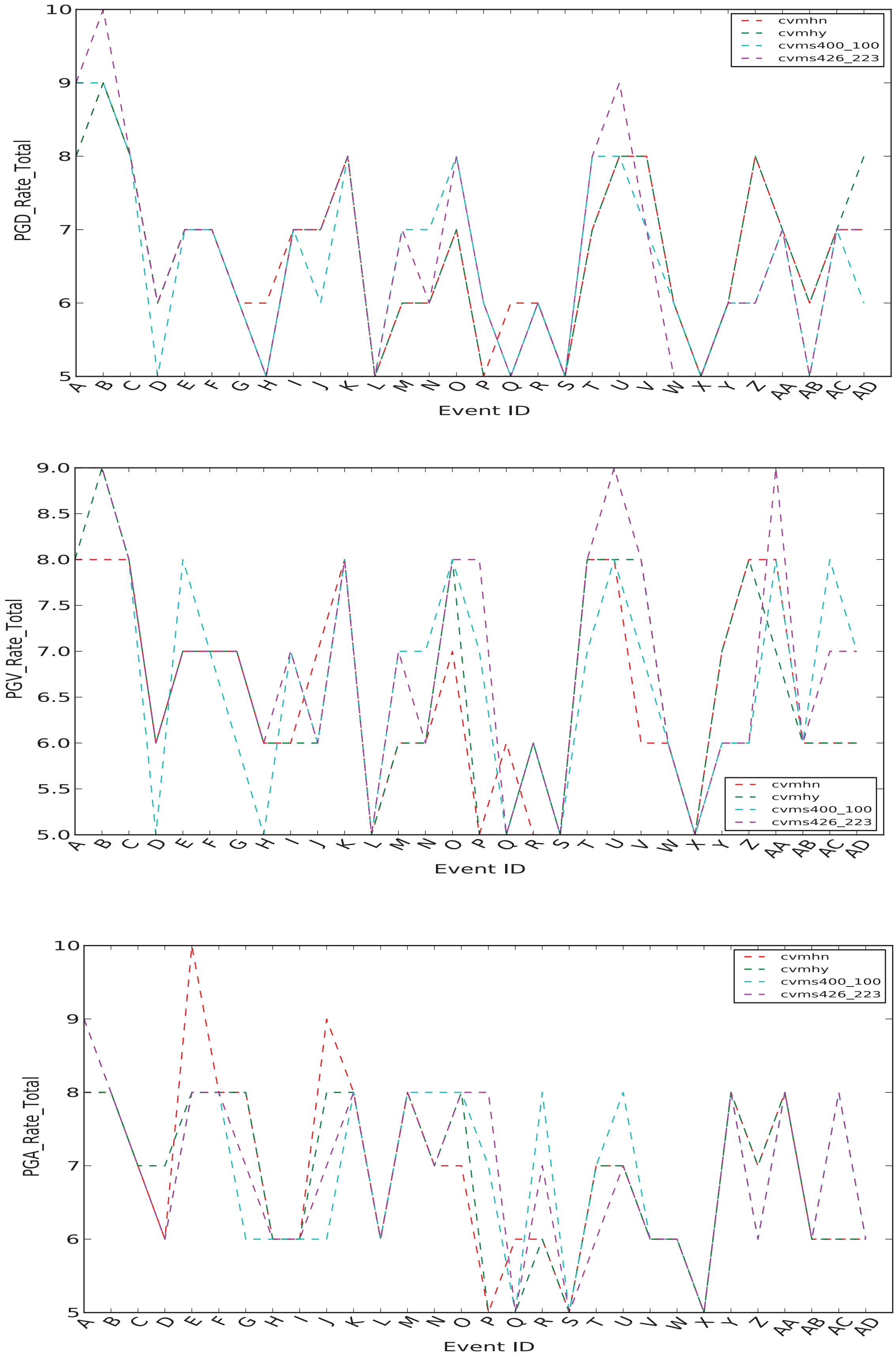

Figure 5. PGD, PGV and PGA total average score for all the 30 events and all four models are presented.