Reviews of Crowell et al., "Validation of Peak Ground Velocities Recorded on Very-high rate GNSS Against NGA-West2 Ground Motion Models"

Reviewer #1

Review of "Validation of Peak Ground Velocities Recorded on Very-high rate GNSS Against NGA-West2 Ground Motion Models" by Crowell et al.

This paper demonstrates the use of very-high-rate GNSS derived velocities in ground motion models. The authors use the GNSS PGV in established NGA-West2 GMMs and in a model derived specifically for this dataset. The manuscript demonstrates the utility of GNSS PGV in the ShakeMap product, expanding upon the lead author's 2021 SRL paper. The manuscript also makes available a large dataset of GNSS-derived velocities for a wide range of earthquake magnitudes, which will likely be of interest to the seismology/geodesy community for future work. I find the manuscript to be very clearly presented and organised, with no concerns about the scientific content. Rather, I include minor corrections, clarifications and suggestions below. With minor exception (see below), I think this paper is publishable as is. Well done.

Line 20: Define PAGER

Figure 2: Can you clarify whether the red histograms are all the events or everything except Kaikoura?

Line 196: This says the Q-Q plot for CY14 is good out to 3 quantiles, but the plot appears to not quite reach 3 quantiles (having the 1 and 3 labels on the axes might help). Is saying 'out to 3' really justifiable, or is it more like 2.5? Please clarify/correct. (same in line 236)

Section 3.2: Why station ARQT? Why was this station assumed to be representative of all the stations in the dataset?

Figure 6: While not required, it would be useful to include the locations of DYFI (circles, perhaps, as they are shown in published USGS models). DYFI often makes up for gaps in station data, but in looking at the USGS event page it appears this earthquake had no DYFI reports in the near-source area. This is important context.

Reviewer #2

The paper by Crowell et al. investigates the performance of applying high-rate GNSS to strong motion monitoring. The authors validate the GNSS-derived velocities with ground motion models and macroseismic intensity observations. This paper is more a performance assessment of existing rather than introducing new processing methods. Nonetheless, I find this study as an interesting and worth-publishing contribution to the subject. However, I also believe the manuscript requires revision before publishing. The authors are asked to address my itemised comments.

- Line 37-38. This is true, but we assume that such GNSS solution outages are correctly eliminated, thus, should not imply the solution.

- Line 40. I cannot fully agree with that statement. It depends on the positioning model (absolute vs. relative) and the baseline length in the case of the latter model. Moreover, currently, site-specific errors are also recognised as an essential source of accuracy decline.

- Line 117. Are you sure that tropospheric delay may induce high-frequency noise in the displacement time series? I doubt. This error is not changing rapidly with time, which is in contrast to ionospheric delay and multipath.

- Line 203. What about the time autocorrelation of GNSS phase observations? This is an important issue in the high-rate GNSS data processing that implies the noise of the solution. This issue was, however, missed by the authors.

- Please avoid acknowledging several papers in the same sentence. Please show the advances of each paper explicitly. Line 31-32.

Reviewer #3

This is a nice paper. The authors convincingly show that PGV measuremementns extracted from high-sample rate (5Hz+) are (a) consistent with published PGV ground motion models and (b) useful for downstream products such as ShakeMap.

I have several minor comments in the attached PDF, which I will not reproduce here but would be nice to see addressed. I think they could strengthen and increase the clarity of the work.

My only main concern is the choice of PGV GMM as applied to subduction zone events. NGA-West2 is really not appropriate here, especially not for M7.9+ events. Generally, in our GMM community, it is widely accepted that one should not extrapolate the GMM to a tectonic environment it was not developed for and outside of its magnitude/distance ranges. Just because you obtained mean ~ 0 residuals and those had a lognormal distribution, it does not mean it is appropriate to carry out this extrapolation.

This issues is further exacerbated because there is a good alternative. The NGA-Sub model is now published in Parker (2022) and references therein. It is also available for use in OpenQuake. I encourage the authors to explore using that instead.

Refs:

Parker, G. A., Stewart, J. P., Boore, D. M., Atkinson, G. M., & Hassani, B. (2020). NGA-Subduction global ground-motion models with regional adjustment factors. PEER Report 2020/03. Berkeley, CA: Pacific Earthquake Engineering Research Center, University of California. University of California, Berkeley.

We thank the three reviewers for their helpful comments and kind words. We have outlined our changes/responses in red below. For our annotated manuscript, we used the LATEX added and replaced functions, which defaults to blue colored text for changes.

Reviewer #1

Review of "Validation of Peak Ground Velocities Recorded on Very-high rate GNSS Against NGA-West2 Ground Motion Models" by Crowell et al.

This paper demonstrates the use of very-high-rate GNSS derived velocities in ground motion models. The authors use the GNSS PGV in established NGA-West2 GMMs and in a model derived specifically for this dataset. The manuscript demonstrates the utility of GNSS PGV in the ShakeMap product, expanding upon the lead author's 2021 SRL paper. The manuscript also makes available a large dataset of GNSS-derived velocities for a wide range of earthquake magnitudes, which will likely be of interest to the seismology/geodesy community for future work. I find the manuscript to be very clearly presented and organised, with no concerns about the scientific content. Rather, I include minor corrections, clarifications and suggestions below. With minor exception (see below), I think this paper is publishable as is. Well done.

Line 20: Define PAGER

Added in the definition here

Figure 2: Can you clarify whether the red histograms are all the events or everything except Kaikoura?

The red histogram is all events here. I simply ran two plot functions in GMT, the first histogram with all data, the second (grey) with just the Kaikoura data. We added some clarification to the Figure caption to reflect this.

Line 196: This says the Q-Q plot for CY14 is good out to 3 quantiles, but the plot appears to not quite reach 3 quantiles (having the 1 and 3 labels on the axes might help). Is saying 'out to 3' really justifiable, or is it more like 2.5? Please clarify/correct. (same in line 236)

You are correct here, its more like 2.5. There is not enough data to reach fully out to 3 quantiles. We have changed the text accordingly.

Section 3.2: Why station ARQT? Why was this station assumed to be representative of all the stations in the dataset?

ARQT was chosen here because it is one of a few sites that has 20 Hz RINEX data available, so we could perform the downsampling prior to running through SNIVEL. What we were really going for in this section was trying to show the relative noise level between different sample *rates for the same period of time. We intend to perform a much more thorough noise analysis in the future for an ensemble of stations using the probabilistic power spectrum method.*

Figure 6: While not required, it would be useful to include the locations of DYFI (circles, perhaps, as they are shown in published USGS models). DYFI often makes up for gaps in station data, but in looking at the USGS event page it appears this earthquake had no DYFI reports in the near-source area. This is important context.

Thank you for this suggestion, we added the DYFI locations to panels A-C and modified the caption.

Reviewer #2

The paper by Crowell et al. investigates the performance of applying high-rate GNSS to strong motion monitoring. The authors validate the GNSS-derived velocities with ground motion models and macroseismic intensity observations. This paper is more a performance assessment of existing rather than introducing new processing methods. Nonetheless, I find this study as an interesting and worth-publishing contribution to the subject. However, I also believe the manuscript requires revision before publishing. The authors are asked to address my itemised comments.

- Line 37-38. This is true, but we assume that such GNSS solution outages are correctly eliminated, thus, should not imply the solution.

In real-time positioning, solution outages lead to incredibly poor performance for displacement estimation. Ambiguities in general are not carried through large outages, and thus need to be re-estimated.

- Line 40. I cannot fully agree with that statement. It depends on the positioning model (absolute vs. relative) and the baseline length in the case of the latter model. Moreover, currently, site-specific errors are also recognised as an essential source of accuracy decline.

I am slightly confused by this comment, the sentence is comparing GNSS displacements to inertial seismic observations (i.e. strong-motion accelerometers or broadband seismometers). There is no positioning model that is going to be less noisy than those instruments (in general, 2-3 orders of magnitude less noisy), regardless of site specific terms. True, under very high shaking, seismic inertial observations will suffer due to rotations and tilts, which makes them less accurate, however, the background noise level is significantly less.

- Line 117. Are you sure that tropospheric delay may induce high-frequency noise in the displacement time series? I doubt. This error is not changing rapidly with time, which is in contrast to ionospheric delay and multipath.

For the hydrostatic delay (removed through Niell's correction already), it is not generating high-frequency noise, however, the wet delay can change quite rapidly due to the heterogeneous nature of weather patterns. We have added in the word 'wet' before troposphere to signify this difference.

- Line 203. What about the time autocorrelation of GNSS phase observations? This is an important issue in the high-rate GNSS data processing that implies the noise of the solution. This issue was, however, missed by the authors.

This is an excellent point, and something we have not formally addressed to date with regards to estimation of velocities. We have added an additional statement here:

also, there is a strong autocorrelation of noise between the components and a regional correlation of noise due to the similar constellation geometries

- Please avoid acknowledging several papers in the same sentence. Please show the advances of each paper explicitly. Line 31-32.

While we agree that showing the advances of each paper explicitly is a good idea, we think in this instance it distracts from the message of the paper, which is that GNSS velocities can measure ground velocity well. The value of displacements to EEW and TEW are highlighted in those papers, and it allows us to go to the 'caveat' quickly that displacement estimation still has issues in real-time.

Reviewer #3

This is a nice paper. The authors convincingly show that PGV measurememtns extracted from high-sample rate (5Hz+) are (a) consistent with published PGV ground motion models and (b) useful for downstream products such as ShakeMap.

I have several minor comments in the attached PDF, which I will not reproduce here but would be nice to see addressed. I think they could strengthen and increase the clarity of the work.

Our specific responses to the PDF comments are shown at the end of this document.

My only main concern is the choice of PGV GMM as applied to subduction zone events. NGA-West2 is really not appropriate here, especially not for M7.9+ events. Generally, in our GMM community, it is widely accepted that one should not extrapolate the GMM to a tectonic environment it was not developed for and outside of its magnitude/distance ranges. Just because you obtained mean ~ 0 residuals and those had a lognormal distribution, it does not mean it is appropriate to carry out this extrapolation

This issues is further exacerbated because there is a good alternative. The NGA-Sub model is now published in Parker (2022) and references therein. It is also available for use in OpenQuake. I encourage the authors to explore using that instead.

Refs:

Parker, G. A., Stewart, J. P., Boore, D. M., Atkinson, G. M., & Hassani, B. (2020). NGA-Subduction global ground-motion models with regional adjustment factors. PEER Report 2020/03. Berkeley, CA: Pacific Earthquake Engineering Research Center, University of California. University of California, Berkeley.

This is an excellent point and one that we considered prior to submission. We felt that using the NGA-West2 GMMs, which went through an extensive validation exercise and many of which form the basis for the ShakeMaps in the western US, was an initial good test case to show that the PGV values we are obtaining are good, within the bounds of published models.

Based on this comment, we have looked at the statistics of our dataset closer. 34% of our observations fall in the M < 7 range and 39.7% are between 7 <= M < 8. Likewise, 61% of our observations are from non-subduction zone events. We went back and reexamined the statistics of the subduction events and non-subduction events and found far lower standard deviations for all 3 GMMs using only the subduction events over the non-subduction events, even when excluding Kaikoura completely. While we do not rerun all of our analyses by adding in the Parker GMM here, we do add the following paragraph (and edits to the Table of statistics):

"We also computed statistics for two additional event subsets: subduction zone events and non -subduction events (minus Kaikoura). None of the three GMMs were not directly developed for the subduction environment, so it is important to understand any systematic biases that may arise due to the tectonic environment. When discounting Kaikoura, there are 268 observations of subduction earthquakes and 293 observations from primarily strike-slip faults. When using R_p, the median residuals for the non-subduction events were all negative, indicating the GMMs are underestimating ground motions. For the subduction events, the standard deviations were all lower than the non-subduction events, and the median residuals were better for BSSA14 and CB14. This result is somewhat paradoxical in that the GMMs we are comparing against in this study were developed primarily for upper-crustal faults, so a better fit for those events would be expected, however, the shallower events will have far more variability with regards to source terms, distance measurements, and directivity such that a greater variability in the ground motion residuals would be observed. For subduction events, the source distances are generally greater so much of the GMM complexity can be averaged out. We would expect even better fits by using subduction zone specific GMMs (e.g., \cite{parker22})."

PDF Comments

Line 17: PGD and spectral displacement - *We have added both of these in here.*

Line 26: I'm not sure this is strictly true. There's plenty of PGVs in the NGAWest2 and NGGA-Sub datasets. Whether they are the "true" PGVs I guess could be debated but because PGV is an "intermediate"period metric there's no reason to think they are wrong.
In other words the onus is on you to prove they are unreliable and you've not done that so far so this statement seems like an overreach.

We agree that this is probably slightly confusing here. There is a larger issue that integration of acceleration to velocity and displacement is problematic due to sensor rotations and tilts. We have shown in previous works that correction schemes (i.e. baseline corrections) lead to both different amplitudes and phase (see Melgar et al., 2013 https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/jgrb.50135), at least in the displacement arena, but there is no reason to think the velocity is not impacted as well. We have rephrased the sentence a bit to:

"There is also a question of whether or not PGV observations for large ground motions recorded with inertial sensors are recording the true ground motions due to sensor rotations and tilts (Boore et al., 2002; Clinton, 2004)."

Line 78: ShakeMaps are often refined days after the fact. Is it worthwhile to study how much the PGV estimates change between broadcast-> rapid-> final orbits?

This is an excellent comment and certainly a topic of future research. Through initial tests, rapid and final orbits do not yield much reduction in noise since the velocity of GPS satellites is on the order of 4 km/s and the position errors of broadcast to final orbits goes from 10s of cms to a few cms, so the relative precision between the two orbit solutions is negligible. This is the prime advantage of using the variometric approach.

Figure 1 caption: Is Rp meant to be the Thompson & Baltay distance? If so you've not defined it yet and you've not said what exponent value you are using...

Might be worthwhile explaining why use this distance metric since it is by no means the common choice in GMMs, more common are Rrupt or Rjb.

Yes, the caption in Figure 1 is Rp from Thompson and Baltay. We have defined it in the figure, our p value is -2 which is the optimal for PGV in T&B.

We do use Rrup throughout as well, but we think that Rp is a more honest representation of the fault distance (due to playing with these distances for the geodetic working group of ShakeAlert). Rrup treats any part of the fault surface as ground motion generating, which works fine for simple ruptures, but long and complex ruptures, it leads to huge over-estimation in ground motion. Rjb is similar to Rrup in this aspect, but for any subduction event, leads to very unrealistic predictions

Section 2.1: Am I reading this section correctly to mean that you only use 5Hz+ data? Is 1Hz then not good enough because of aliasing? Would it be beneficial to outright say that and have some example plot that shows one or several PGVs as a function of sampling rate to drive that point home?

Smalley (2009) did a little bit of this but I think it was just from theory and not from actual data…

In this study, we are only using the 5+ Hz dataset because of the issue of aliasing for smaller magnitude earthquakes (e.g. the corner frequency argument in the second paragraph). Thank you for pointing out the Smalley paper (I read it at the time a decade ago, but had forgotten it). In that paper, they are looking at aliasing on the displacement streams, mostly from theory, but also with a few examples. In PGD, the coseismic contribution (zeroth frequency) becomes more predominant above M7 and at lower magnitudes, the primary source of PGD is in the surface waves. PGV would have slightly different frequencies here, but same general concept.

We are starting to look more in depth at the 1 Hz observations and the impact of aliasing, but that is a topic the 2nd author is working through for another manuscript.

Section 2.1: I think a map is necessary with the events somewhere.

We have added a map of events. There is also a spreadsheet of events and observations with the dataset.

Line 109: This doesn't make sense to me. Why???

The filtering done here is very light to remove some of the high frequency background noise in the observations. We have found that this does not change the peak values very much, if at all, especially for the larger ground motions. For the smaller ground motions, it is easier to visually spot the seismic waves in the time series when applying the light filtering. With seismic observations in ShakeMap, it is often standard procedure to employ filtering on the waveforms prior to inclusion in ShakeMaps (e.g. https://pubs.geoscienceworld.org/ssa/bssa/article/89/1/311/120463/Continuous-monitoring-of-ground-motion-parameters).

Line 111: Feels arbitrary. Would it not be better to apply some SNR cutoff and only use sites above that level?

While we agree that some more automated technique would be elegant, there weren't that many records to manually inspect. When I compiled the dataset for Goldberg et al., I visually inspected all of the waveforms, which made a huge difference compared to the Ruhl et al. dataset (which uses SNR). What we found was many observations were excluded in Ruhl on the lower end, but many observations that were clearly noise were also included. Visual inspection is fairly easy to do here since the frequency content of shaking is very different from *the colored noise structure of GNSS. The recent paper by Dittmann in JGR provides an alternative approach to this as well by using a random forest classifier.*

Section 2.3:This section is missing a few things.

Are the finite faults "pruned"? As in is slip below a certain peak removed? This is common practice and done for GNSS in Goldberg et al 2020 for example.

Goldberg also used Rp in their study. When we do use Rrup, we do not perform any pruning of the slip model. This is something that we have considered in the past (i.e., use only fault patches above 10% of the max), but we found the USGS models don't include an egregious excess fault areas some models do.

I hate to be the ground motion person but NGA West2 is only really valid for California. There's no map so I'm a little confused about the earthquakes in the dataset but for subduction zone events (e.g. Tohoku/Tokachi) it's not valid. Also big events in this paper are outside the valid M range fro the GMMs and that's a big no-no, to extrapolate past the distance/magnitude ranges of a given GMM. I suggest using NGA-Sub, this is published now in Parker (2020,2021) and available on OpenQuake

See the first comment for this review, where we address this issue.

How are the basin depth terms Z1.0 or Z2.5 defined?

Basin terms are defined using the formulations within the specific GMMs. In CY14, this is EQ 1 (Z1.0). In CB14, its equation 33. BSSA14 does not include this term. We have added a sentence in here stating that we are using the basin term equations from the GMMs.

Line 164: The USGS-NEIC FF is very different from the one in the Hammling Science paper. Worthwhile to sue that one instead?

We agree that there are certainly other FF models to use that would yield better results, but if we modified our procedure for Kaikoura, we would have to for all of the other events, which is a terrible rabbit hole to go down. Plus, the formats of these FF models are not standardized, so going with the USGS one makes it easy.

Table 1: Is it possible to compare the residuals to what is published in the BSSA14 or other NGA-West papers? i.e. are you doing better, worse, or the same?

We do mention the standard deviations on line 162 for the 3 NGA west GMMs. In short, on par with our GMM, slightly worse using them directly, excluding Kaikoura roughly on par. The main caveat is that the reported GMM uncertainties contain many low amplitude observations, which are far easier to fit and there are many thousands more observations.

Section 4: Would it be useful to do mixed effects here to separate the residuals into source, site, and path?

This is a topic that we would love to address in the future, however, there is not much overlap between specific stations (only a few stations record different events), so we don't know how accurate any mixed-effects regression would be here. One potential solution would be to combine seismic PGV observations for lower magnitude events to get better station-event overlap, which is something we are starting to look at (and given that this entire dataset is open, we would hope others in the ground motion community would use our observations).