Response to Reviewers for Seismica Manuscript

Characterizing High Rate GNSS Velocity Noise for Synthesizing a GNSS Strong Motion Learning Catalog

Timothy Dittmann, Jade Morton, Brendan Crowell , Diego Melgar, Jensen DeGrande, David Mencin

August 14, 2023

[Responses to reviewers are in blue.]

Reviewer A:

Review of Manuscript titled “Characterizing High Rate GNSS Velocity Noise for Synthesizing a GNSS Strong Motion Learning Catalog”

General Overview

The primary objective of this study is to test and develop models to overcome noise-related hurdles in data processing and interpretation of GNSS velocity measurements. Here, the authors have developed and used synthetic data catalogs (strong-motion data augmented with GNSS noise) to differentiate signal and noise. They also proceed to compare the extracted features with those obtained from real GNSS velocities to validate their model.

The study is interesting as it takes a step closer to better employing two types of data sets (inertial and GNSS) that have their own drawbacks (signal saturation, noise etc), for data-driven approaches in earthquake response/forecasting and deep learning.

The authors have also clearly listed the limitations of the current study and potential directions for further research.

The weakness of the manuscript lies mostly in its structure. While most of these corrections are minor, cumulatively they would make a significant difference in following the content. The figures need tweaking along with the captions. There are a couple of clarifications sought in the detailed comments below.

Major comments:

No major scientific comments other than a few clarifications/suggestions listed in the detailed comments below. Editorial corrections are provided for a smoother read.

Dear reviewer A:

Thank you for your thoughtful feedback. We appreciate the investment of your time you have made in this manuscript. We have addressed each of the **detailed and editorial comments** independently:

Line 14: change “ground motion observations: these signals..” → “ ground motion observations that ...”

Author response: change made as requested.

Line 17: Oxford comma – change “ … oscillators and complex ..” → “ … oscillators, and complex ...”

Author response: change made as requested.

Line 20: Break this sentence into two. “… accurate classifications; more complex ...” → “… accurate classifications. (Furthermore/ additionally) More complex...”

Author response: change made as requested.

Lines 30-31: Non-technical summary → “inertial-based instruments”, might require a description in a non-technical summary. Up to the authors and the editor, but could potentially refer to a simplified version Newton’s first law.

Author response: change made as requested.

Lines 72-73: Oxford comma – change “...catalog for training and the quality...” → “...catalog for training, and the quality...”. Please make changes in Lines 78-80, Lines 96, Line 122, Line 176, Line 231, line 358, lines 376-377, etc (wherever need be).

Author response: changes made as requested.

Line 177-179: Break the sentence into two. Change “...earthquake engineering; we use…” → “earthquake engineering. We use...”.

Author response: change made as requested.

Line 181-184: split the sentence.

Author response: change made as requested.

Lines 175-194: Avoid compound sentences separated by a semi-colon. While they make sense in some instances, they are often long-winded. They can easily be split into separate sentences to facilitate easier reading.

Author response: change made as requested.

Line 202: Specify the inteval between the 7 noise percentiles used for augmentation.

Author response: change made as requested.

Line 215: No section called “Open Research”. Did you mean “Data and code availability”?

Author response: Yes, change made as requested.

Lines 254-260: Best refer to Dittmann et al. (2022b) instead, as the lines describing the formulae are verbatim. Or re-word them.

Author response: removed equations and reworded as requested.

Line 269: Change “The Melgar et al ...” → “Melgar et al ...”

Author response: change made as requested.

Lines 271-273: Split the sentence.

Author response: change made as requested.

Lines 274 and 282: Change “(Shu et al., 2020)” → “Shu et al. (2020)”

Author response: change made as requested.

Line 285: Change “(Crowell et al., 2023)” → “Crowell et al. (2023)”

Author response: change made as requested.

Figure 4 caption: (a) change “NGAW” → “NGAW2”. (b) Split the sentence “The sum of these time series....waveform in magenta”, and insert the legend for the summed waveform (purple). Change “..4 largest amplitudes, the median ...” → “..4 largest amplitudes (solid magenta circles), the median ...”.

Author response: changes made as requested.

Assign subplot lable (c) for the PSD description.

Author response: Moved C-label up to top left corner for visibility.

Lines 297-301: Refer to Figure 6.

Author response: change made as requested.

Line 316: reference?

Author response: References added.

Figure 6 caption: change “...features; “time” are the time domain...” → “..features, “time” are the time domain...”. Indiacte ‘third row’ as (c) since that’s how it is labelled in the figure.

Author response: change made as requested.

Line 339: Change “figure 8” → “Figure 8”

Author response: change made as requested.

Lines 340-342 and figure 8: Please specify the reason(s) for choosing these percentiles for visualization instead of the other options.

Author response: change made as requested.

Figure 8 caption: Change “The bottom panels ...” → “The bottom panels (c-d)...”. Add “TPR – True Positive Rate”

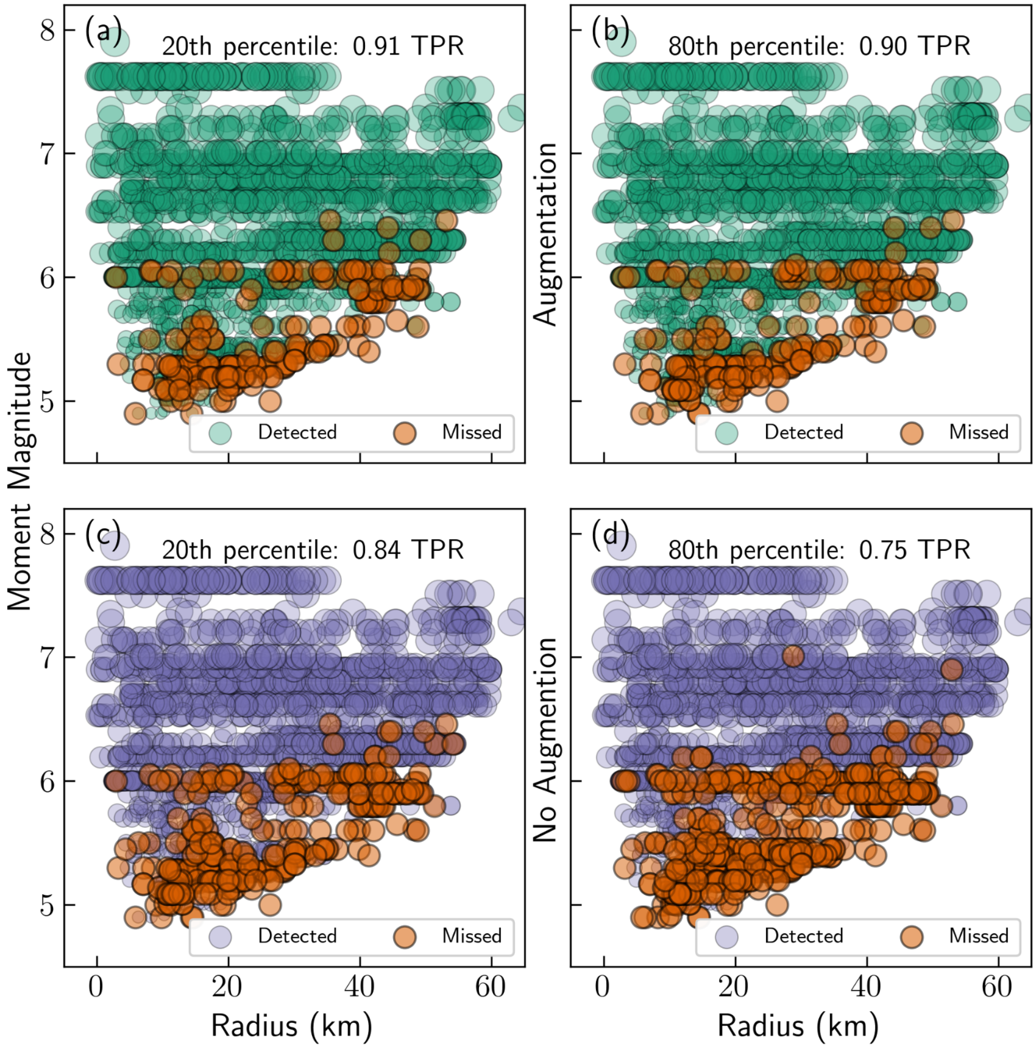
Author response: change made as requested.

Section 3.3 and Figure 8: Do you notice any trend in the detected vs missed events in terms of their magnitude? Or is it random? Is there a diffrence between augmented and non-augmented data in this regard?

Author response: Thank you for that feedback. SNR, in this context, is approximately a function of moment magnitude, hypocentral radius, and noise. We chose to illustrate detection as a function of SNR, but magnitude follows a similar pattern: missed detections are smaller events and/or further afield. Similarly, augmentation improves detectability at the boundaries of magnitude/radius OR SNR.

By viewing in the magnitude domain, it does become apparent that at the distances we are exploring (<60km), augmentation is beneficial for events <M6. Above ~M6.0 at these distances the signals are sufficiently strong to detect regardless of noise signatures presented in training. This aligns with our motivation for augmentation, in which we expand the sample training space to improve generalization, which in turn improves sensitivity for relatively weaker signals.

For the reviewers reference, here is the same figure but this time with Moment Magnitude instead of SNR as a function of radius:



Section 3.4: Refer to Figure 9 in this section.

Author response: change made as requested.

Figures:

Figures 1, 3, and 9 are not referred to in the main text at all. Please add them where necessary. Figure 2 and Table 1 are referred to much later in the text, and should be referred to sequentially for better structural flow of the manuscript.

Author response: change made as requested: Added ref for 3, 9, and earlier relevant references for Fig 2 and Table 1.

Figure 1: Specify x-axis time-frame.

Author response: The authors decided to remove this figure, as the information was redundant with Figure 3.

Figure 2: The colors used, while artistically pleasing, are not colorblind friendly. Therefore, it would be a good idea to use different outlines for the histogram (in a) and for the area curves (in b) (can retain color). Perhaps one solid and the other dashed? Also indicate the line-type with data-type in the figure or the figure caption. (same for Figure 6, 7, 8, 9)

Author response: Thank you for this feedback. When we selected these colors, we used the colorbrewer (https://colorbrewer2.org/) project "colorblind safe" set of colors for explicitly this reason. Given these feedback, we further evaluated this color selection using the following tool (https://davidmathlogic.com/colorblind/#%23D81B60-%231E88E5-%23FFC107-%23004D40), and can see how the 'teal' and 'purple', while different, could be difficult to distinguish. For this reason, we swapped in the 'orange' color in figs 1 and 8 (latest ms figure numbers). We also added a dashed/solid line for the cumulative plot. For figure 5 we chose a new 'colorblind friendly" color schema.

I hope you find this acceptable. If you or a colleague still has difficulty distinguishing, we hope you will let us know so we may further refine.

Figure 3: Add sub-plot labels (a, b, c) for easier reference, and edit caption accordingly.

Author response: change made as requested.

Figure 4: in (a) y-axis descriptor is cut off (maybe during PDF compilation). Please correct this. Would also suggest using different line thickness for overlapping waveforms in case the colors aren't distinguishable to the reader. Assign sub-plot label (c) for the PSD plot at the bottom

Author response: changes made as requested. Fixed formatting, moved up subplot label and varied thickness of lines/colors consistently across subfigures.

Figure 6: Indicate (b-1) and (b-2). Ideally use different line thicknesses for psd_t and t in (b) and (c). y-axis in (b) and (c) need proper descriptors within the subplots.

Author response: The authors decided to remove these subplots as it wasn't germain to the manuscript. Also, the color scheme changed to be "colorblind friendly." (colorbrewer2.org)

Reviewer B:

Review of “Characterizing High Rate GNSS Velocity Noise for Synthesizing a GNSS Strong Motion Learning Catalog” by Dittmann, Morton, Crowell, Diego Melgar , DeGrande, Mencin.

Friday, June 30, 2023

The following is a review of the paper “Characterizing High Rate GNSS Velocity Noise for Synthesizing a GNSS Strong Motion Learning Catalog” submitted by Dittmann et al., for publication in Seismica. In this work the authors seek to quantitively improve our understanding of the impact of GNSS noise characteristics for event detection by greatly expanding the data set of observations available via data augmentation. High-rate GNSS is a relatively new modality of measurement that is now being included in hazard monitoring scenarios. However, because the research community has <10 years of such measurements available the data set is relatively small especially for events of M>=7. This means that any technique developed using the observed data sets is limited, this is especially the case for machine learning (ML) methods were large data sets are required to train models and the lack of data prevents generalization of the models.

The authors research contribution in this work is that they develop a GNSS noise model from 247 5-Hz GNSS velocities recordings and transfer that onto the 80 year-long, NGAW2 string-motion data set. The authors create a labeled, pseudo-synthetic data set consisting of string-motion velocity recordings with an observed GNSS noise model imposed. To do this, they create a suite of percentile based GNSS velocity noise models from observed data and impose 7 of these on NGAW2 velocity waveforms. The study is restricted to events of M>=5 at radius of <=70 km and >=1 cm/s PGV. In doing this they expand the waveforms available for analysis from ~250 to ~14,000. The authors then evaluate the improved signal detection of this augmented data set using a random forest classifier. They demonstrate that a classifier trained using this augmented data set improves generalization of the ML model compared to an observed-only GNSS data set. It is of note that the real-GNSS velocities are included in the machine learning training thus artifacts unique to GNSS processing are retained. The authors also compare ambient analysis of real- and augmented ambient noise data set and report in significant change in the true negative rate.

A primary outcome of this work is a large set of pseudo-synthetic velocity recordings with realistic GNSS velocities imposed. The authors have made this labelled data set available for future work in an open repository. This, combined with the documentation of its generation provided in the paper, will be a high value data set to those interested in pursuing the use of GNSS velocities for monitoring purposes. With the growth of high-rate GNSS networks and the increasing realization of the need to include it in EEW and TEW networks this paper is very timely.

The paper is well written and clearly structured. The title is adequate and accurate as is the abstract. The methods are described in sufficient detail and the authors have provide links to the codes used and data sets analyzed. All references are listed in the reference section. The links supplied in the Data and Code Availability section are working links to data and code .The figures are good and improve the understanding of the paper. However, there is an issue that needs to be addressed in terms of referencing them.

My recommendation is that the work is significant, and the paper is well written. No changes are required expect for minor edits and comments that help with clarification are listed in the following section.

Dear Reviewer #2:

Thank you for your thoughtful feedback. We appreciate the investment of your time you have made in this manuscript. We have addressed each of the **comments** independently:

There is no reference to figures 1, 3 or 9 in the body of the text. These figures should be included in the paper but need to be referenced from the main text.

Author response: changes made as requested. (also requested by reviewer 1)

The first reference to figure 2 is on line 333 after the reference to figure 6.

Author response: change made as requested. Early reference made on line 187. (also requested by reviewer 1)

Figure 1 is missing x-axis label and tick marks. (I assumed this was a time axis).

Author response: The authors decided to remove this figure, as the information was redundant with Figure 3.

Figure 2a missing x-axis label “Number of events ?”

Author response: changes made as requested. Added “counts”.

Figure 3. Missing x-axis label (Seconds?).

Author response: changes made as requested.

Figure 6. Row b, the x-axis are labelled “Precision, Recall and F1” but should they be labelled SNR? The caption says the vales are expressed as a function of SNR.

Author response: The authors removed these subplots as it wasn’t germain to the manuscript.

The ground-truth data set has been selected based on several criteria , M>5, radius <10km, 1cm/s PGV. The authors should therefore make it clear that there results and find only apply to this range of observations. Its transferability could be reduced for events outside there ranges.

Author response: Thank you for this feedback. We have added text in the abstract and in the conclusions to make the scope of this manuscript's analysis and subsequent applicability range more clear.

Section 2.3 The strong motion observations have been passed through an acausal Butterworth filter and band passed filter. Have the GNSS velocities been treated similarly? Perhaps worth mentioning.

Author response: The GNSS velocities are not filtered. The strong motion observations are our “truth” upon which we superpose GNSS noise time series. Our objective is to have realistic pure signals (where “noise”, in its many possible forms, is removed) embedded in realistic GNSS noise.

If we were comparing detection performance of GNSS with strong motion instruments, then we would want to consider equivalent data handling. Filtering the GNSS may prove beneficial for other applications, but not our stand alone detection.

Line 317. Typo. The sentence begins, “Within the time domain …” , should this be “Within the frequency domain …”?

Author response: Changes made as requested.

Section 6.2. The recall scores are poor for all feature extraction strategies < 0.8 . This suggests a higher number of missed events. Could this pose a problem for implementation in monitoring projects were FNs have high consequence?

Author response: The authors found that detection decision thresholds are a balance of classifier sensitivity with classifier false alerting. For this analysis, we chose to balance sensitivity and false alerting by optimizing on F1 score, the harmonic mean of precision and recall. But, this underlying decision threshold and subsequent willingness to accept false alerts in order to avoid missed detections (or vice versa) are a tunable parameter for any downstream decision module. An excellent reference on optimizing these decisions from a cost perspective is Minson, et al. (2019).

Figure 7 and Section 3.2 paragraph 3. In figure 7 it seems the most important features are those in the time domain. This contrasts with the information shown in Fig 6 for F1 statistic and poor recall score. Could figure 7 be slightly misleading? If you were to calculate the area under the time domain portion of fig 6b and the area under the frequency domain section, would they be about equal?

Author response: Thank you for this insightful feedback. We agree with your thoughts, that it is the combination of information presented in Figure 5 and Figure 6 that explains the models behavior: when combined (psd_t), the time domain information has the highest single feature importances, but when grouped, the frequency domain features have the largest importance (the area under the curve, as you put it). For example, in figure 5(b), the highest single importances are the Max1-4, but the sum of the time domain features is 0.2, whereas the sum of the frequency domain features is 0.25. Each on their own performs reasonably well, but the ability to combine all information into higher dimension decision criteria that outperforms each stand-alone strategy justifies the overall ML approach.

In light of this feedback, the authors have added some text to make this more clear.