Supervised Machine Learning of High Rate GNSS Velocities for Earthquake Strong Motion Signals.

Timothy Dittmann[1], Yunxiang Liu[2], Y. Jade Morton[3], and David Mencin[3]

[1]UNAVCO Inc.
[2]University of Colorado, Boulder
[3]University of Colorado Boulder

November 26, 2022

Abstract

High rate Global Navigation Satellite System (GNSS) deformation time series capture a broad spectrum of earthquake strong motion signals for rapid contributions to hazard warnings and assessment, but experience regular sporadic noise that can be difficult to distinguish from true seismic signals. Previous studies developed methods for automatically detecting these signals but most rely on various external inputs to differentiate true signal from noise. In this study we generated a dataset of high rate GNSS time differenced carrier phase (TDCP) velocity time series concurrent in space and time with expected seismic surface waves from known seismic events. TDCP velocity processing has increased sensitivity relative to traditional geodetic displacement processing without requiring sophisticated corrections. We trained, validated and tested a random forest machine learning classifier. We find our supervised random forest classifier outperforms the existing detection methods in stand-alone mode by combining frequency and time domain features into decision criteria. We optimized the classifier on a balance of sensitivity and false alerting. Within a 100km epicentral radius, the classifier automatically detects 86% of events greater than MW5.0 and 98% of events greater than MW6.0. The classifier model has typical detection latencies seconds behind S-wave arrivals when run in real-time mode on “unseen” events. We conclude the performance of this model provides sufficient confidence to enable these valuable ground motion measurements to run in stand-alone mode for development of edge processing, geodetic infrastructure monitoring and inclusion in operational ground motion observations and models.

Supervised Machine Learning of High Rate GNSS Velocities for Earthquake Strong Motion Signals.

T. Dittmann[1,2], Y. Liu[1], Y. Morton[1], D. Mencin[2]

[1]University of Colorado, Boulder
[2]UNAVCO Inc.

Key Points:

- We assembled a labeled dataset of 5Hz GNSS velocity time series from 77 earthquakes over nearly 20 years.
- We trained a supervised random forest classifier for detecting seismic motion that outperforms existing detection methods.
- Improved detection enables lightweight, high rate GNSS velocity processing to be included in operational ground motion observations.

Corresponding author: Timothy Dittmann, stdi2687@colorado.edu

Abstract

High rate Global Navigation Satellite System (GNSS) deformation time series capture a broad spectrum of earthquake strong motion signals for rapid contributions to hazard warnings and assessment, but experience regular sporadic noise that can be difficult to distinguish from true seismic signals. Previous studies developed methods for automatically detecting these signals but most rely on various external inputs to differentiate true signal from noise. In this study we generated a dataset of high rate GNSS time differenced carrier phase (TDCP) velocity time series concurrent in space and time with expected seismic surface waves from known seismic events. TDCP velocity processing has increased sensitivity relative to traditional geodetic displacement processing without requiring sophisticated corrections. We trained, validated and tested a random forest machine learning classifier.

We find our supervised random forest classifier outperforms the existing detection methods in stand-alone mode by combining frequency and time domain features into decision criteria. We optimized the classifier on a balance of sensitivity and false alerting. Within a 100km epicentral radius, the classifier automatically detects 86% of events greater than M_W 5.0 and 98% of events greater than M_W 6.0. The classifier model has typical detection latencies seconds behind S-wave arrivals when run in real-time mode on "unseen" events. We conclude the performance of this model provides sufficient confidence to enable these valuable ground motion measurements to run in stand-alone mode for development of edge processing, geodetic infrastructure monitoring and inclusion in operational ground motion observations and models.

Plain Language Summary

Continuously operating, high sample rate Global Navigation Satellite System (GNSS) sensors that experience ground shaking from an earthquake can provide valuable data regarding the nature of the ground motion. If this data is streamed in real-time, these observations can complement existing traditional seismic infrastructure measurements that are used for earthquake early warning or rapid ground motion assessments. However, the data from these sensors can be noisy and have non-earthquake artifacts that are difficult to tell apart from true seismic signals. In this work we used a nearly 20 year archive of high sample rate GNSS velocities occurring during known seismic events to train, validate and test a machine learning model for earthquake detection. This machine learning approach is taken from existing algorithms used for a wide variety of challenging classification problems where a label can be applied to a sample. We demonstrate that this data-driven method, without any external information, is more likely to detect these signals with less false alarms when compared to existing methods. The added confidence this algorithm provides will allow these valuable measurements to be included in operational seismic assessment and warning decision criteria.

1 Introduction

Real-time measurements of medium to great earthquake ground motions are vital to rapid hazard assessment and earthquake early warning (EEW) systems. Higher rate ($\geq$ 1Hz) continuous GNSS measurements capture dynamic motions and permanent displacements of propagating strong-motion waveforms from such events (Nikolaidis et al., 2001; Larson et al., 2003). These geodetic strong motion measurements (Larson, 2009) will rarely clip nor require double integration that leads to magnitude saturation in the near-field of larger, destructive earthquakes common to inertial velocity sensors (Bock et al., 2004; B. W. Crowell et al., 2013; Colombelli et al., 2013). Furthermore, additional material low-latency observations densify existing ground motion measurements. These observations are particularly valuable when damaging seismic events occur in sparsely instrumented regions (Grapenthin et al., 2017) or when networks or infrastructure fails.

However, geodetic deformation timeseries are noisier than traditional inertial sensors (Melgar et al., 2020). This makes separating signal from noise challenging: signal amplitudes from the largest, most costly events can be difficult to distinguish from non geophysical events, such as filter reconvergence or signal loss of lock. Medium magnitude events, often difficult to detect above the geodetic noise floor, can be destructive or tsunamigenic. The ability to make accurate, low-latency distinction between true signals and noise in stand-alone mode, without external sensors or information, minimizes points of failure and decision latency and maximizes integral network decision inputs and potential edge processing capabilities.

Current approaches to detect motion use variations of time domain thresholds to flatten the decision to a function of signal amplitude. Several existing approaches make use of low-pass filters similar to traditional STA/LTA seismological phase picking (Allen & Ziv, 2011; Ohta et al., 2012; Minson et al., 2014; Kawamoto et al., 2016; Goldberg & Bock, 2017) that extract static offsets for finite fault inversion but filter valuable dynamics information. Recent interest in peak dynamic signals (Melgar et al., 2015; Ruhl et al., 2019; Fang et al., 2020; B. W. Crowell, 2021) prompted use of unfiltered timeseries to capture peak signals for magnitude scaling laws and ground motion intensity measurements. These epoch-wise threshold detection methods (B. W. Crowell et al., 2009; Psimoulis et al., 2018; Hohensinn & Geiger, 2018; Hodgkinson et al., 2020; Dittmann et al., 2022) use instantaneous measurements to estimate motion onset, but have limited "real-world" testing and mitigate high false alert rates by spatially correlating detections with nearby stations or windowing in time from seismic triggers. These processes reduce the utility of these measurements for rapid decision criteria.

In this work, we evaluate whether existing GNSS hardware can: more reliably detect motion signals that are 1) constellating near the ambient temporal noise floor 2) with minimal false alerting 3) in a stand-alone mode and 4) with no specific fault or network geometry. We trained a machine learning classifier on a supervised dataset of GNSS velocity time series concurrent in space and time with known seismic source signals. We assembled, processed and labeled a dataset of 1701 earthquake-station high rate (5Hz) time series pairs. We optimized the classifier on this dataset with applied domain knowledge to feature selection and feature engineering. We present the superior performance of this classifier relative to existing methods within this motivational context. We offer advantages and implications of deploying this processing and trained model at scale for network wide monitoring, with particular emphasis on the improved sensitivity and integrity of stand-alone GNSS event detection without external inputs.

2 Methods

2.1 Signals of Interest

We define our detection domain as a binary *motion* or *no motion* state classification. A critical component of developing a robust classification model is a substantial dataset from which to train, validate and test the model. For optimal results, this dataset requires broad spectrum noise and signal samples such that the model can "learn" and generalize our classification and distinguish signal from noise. We assembled a catalog of 1701 station-event pairs from 77 events by cross referencing available 5Hz GNSS observational data in the UNAVCO geodetic archive with Advanced National Seismic System Comprehensive Earthquake Catalog (ComCat) of earthquakes greater than $M_W 4.5$. While 1Hz data is more readily available, this sample rate is insufficient for capturing certain event spectra (Joyner, August 1984; Smalley, 2009), such as $\sim M_W 6.0$ events in the nearfield. For larger magnitude events it's likely that sampling closer to 10 Hz is nec-

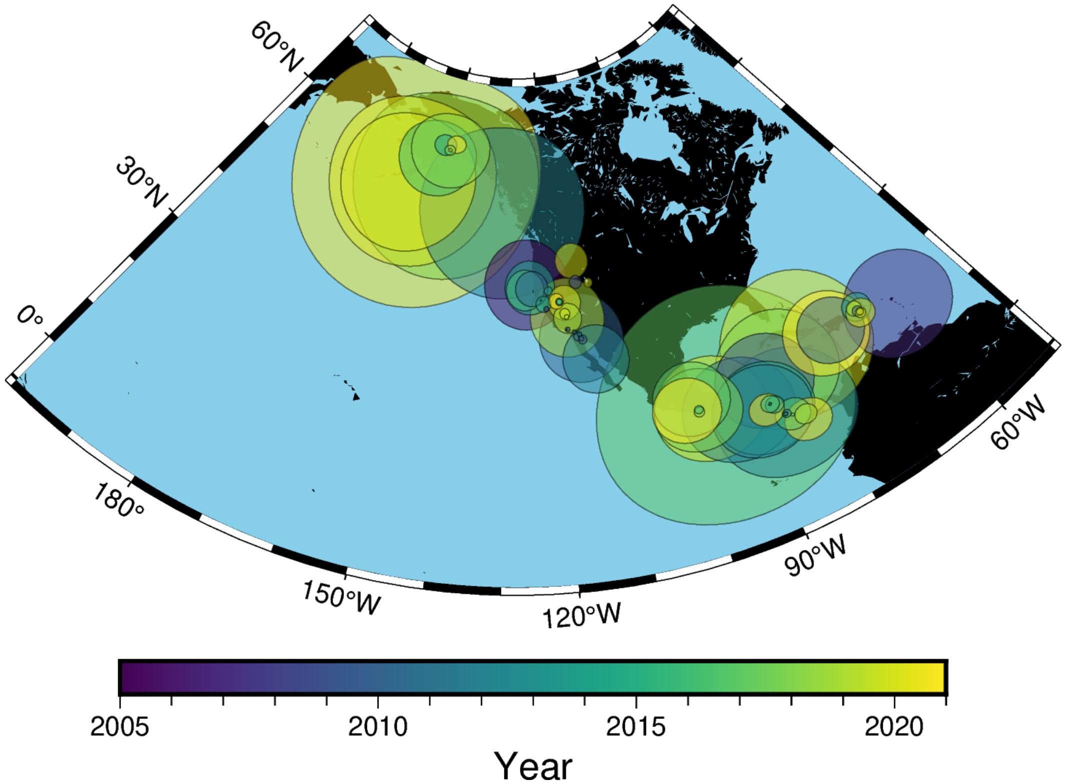


Figure 1. Catalog of events and radii used for this work. The number of stations used in each event is a function of the radii depicted here and the ground station network density.

essary to avoid aliasing (Shu et al., 2018), but we balance this design parameter with the need for sufficiently large available datasets for training. We assigned a conservative radius of detection for each event using ambient noise estimation from Dittmann et al. (2022). For each station-event pair within this spatial footprint, a time series window began 2 minutes prior to earthquake origin time (OT), and extends out in time as a function of radius (Figure 1). We conservatively buffered the radius and time window to ensure the existing model does not limit this result.

Current use of GNSS-derived seismic ground motion for operational EEW (Murray et al., 2018) use precise point positioning (PPP) derived topocentric coordinates to capture dynamic waveforms or static offsets relative to a stations *a priori* position. Instead, we align synchronous carrier phase epoch-wise changes, predicted satellite orbital velocity and line-of-sight geometry to accumulate coherent energy with respect to the shared receiver clock drift rate and directional velocities in a local reference frame. Variations of this geodetic processing method, known as time differenced carrier phase (TDCP) (van Graas & Soloviev, 2004) or variometric velocities, can record co-seismic velocity waveforms (Grapenthin et al., 2018; Hohensinn & Geiger, 2018; B. W. Crowell, 2021) as well as integrated over time into seismic displacement waveforms (Colosimo et al., 2011; Branzanti et al., 2013; Fratarcangeli et al., 2018). We processed these 5hz measurements with the open-source SNIVEL package (B. W. Crowell, 2021) using broadcast ephemeris and narrow lane phase combinations. We chose TDCP over PPP because it is more sensitive to motion (Fang et al., 2020; Dittmann et al., 2022), and it is "lightweight" in that it does not require sophisticated corrections and is computationally inexpensive. From a machine learning perspective, this could be considered a first step in our feature en-

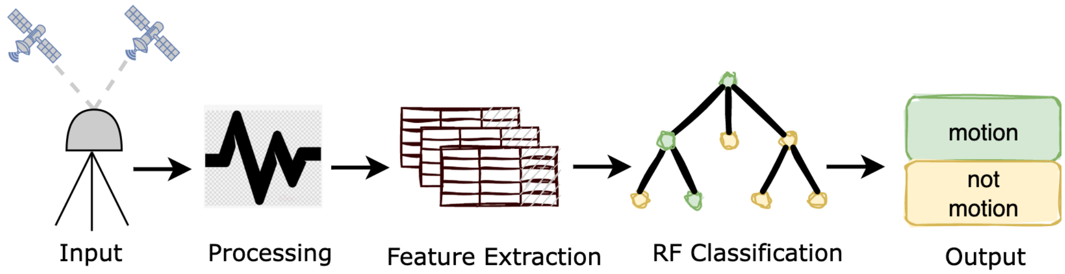


Figure 2. Schematic of our classification workflow: Inputs were 5Hz GPS phase measurements and broadcast ephemeris, which are processed using narrow lane combinations using SNIVEL. Target labeling combined with Feature extraction were used for training a supervised random forest classification model to predict motion classification on testing subsets.

gineering, or applying domain knowledge to extracting features that are correlated with motion in observed carrier phase measurements.

2.2 Feature Engineering Pipeline

Data-driven supervised machine learning models are widely used in computer vision and natural language processing due to their superior accuracy for challenging classification, regression and clustering problems. Earth scientists have adopted many of these models for geoscience research (Kong et al., 2019). Recent catalogs of historic seismic data training sets (eg. Stanford Earthquake Data Set (Mousavi et al., 2019), INSTANCE (Michelini et al., 2021)) have contributed to benchmarking improvements of earthquake detection, phase picking, localization, and magnitude estimation (eg. Meier et al. (2019); Mousavi et al. (2020); Kong et al. (2019). These extensive labeled data sets enable sophisticated data-driven classifiers and deep learning models using inertial seismic data. Several geodetic applications of machine learning algorithms have demonstrated promising results with respect to seismic processes. Crocetti et al. (2021) used a random forest classifier for antenna offset detection, including due to earthquake offsets, from low-rate, 24-hour position solutions. Habboub et al. (2020) applied a neural network to coordinate time series anomaly detection applicable to specific regional datasets well above the noise floor. Dybing et al. (2021) used neural networks for earthquake detection and Lin et al. (2021) employed deep learning used for rapid event magnitude estimation; both of these studies used extensive synthetic displacement waveforms derived from real-world fault geometries and real-world PPP noise models.

In our study, we used a random forest algorithm for our classifier (Breiman, 2001) of GNSS velocities. Random forest is an ensemble of decision trees; a single decision tree is a classifier where input features are split along thresholds to separate source, or root, data from end node classifications, or leaves. An ensemble or forest of trees each vote on the feature decision criteria to select the optimal decisions towards minimizing correlated noise. Due to the infrequent nature of larger magnitude earthquakes, the event classes are naturally imbalanced but by pre-selecting specific time series of events, we have reduced this imbalance for training (Table 1) and testing. Random forest hyperparameters were selected using a grid search over the number of decision trees used, the maximum decision splits within a tree, and imbalance classification weighting strategies.

SNIVEL TDCP processing generates 5 Hz time series of the three topocentric velocity components and the clock drift rates. From these event-station pair time series of velocities, we generated feature sets to label for our supervised classification (Figure

Table 1. Distribution of classification sample labels used in training/testing datasets by component and label.

	East (n=46,778)	North (n=46,778)	Up (n=46,778)
No	84%	85%	97%
Yes	14%	13%	2%
(Maybe*)	2%	2%	<1%

*Maybe's excluded from training/testing

2). Our feature samples consisted of three directional components of 30 second windows overlapping every 10 seconds; within these windows we included the four maximum component norm window values, window median, window median absolute deviation and window lower frequency power spectra as features. These features and windowing allowed our model to incorporate signal and noise amplitude in the time domain, akin to the traditional threshold approach, as well as power spectra in the frequency domain. Labels were assigned through visual inspection as *no* or 0 for no motion, *yes* or 1 for motion, and *maybe* for windows that we are not able to distinguish between yes or no and excluded from testing and training. Each directional component was labeled independently. This resulted in 140,334 labels for the approximately 30 time samples for 1701 station event pairs of three component velocity time series. We evaluated two feature extraction models. Feature set #1 was a combined array of all 3 directional components with a single label at each window. The horizontally concatenated components resulted in $3\times m$ features and n samples, where m is the number of features per component ($m = 36$ in our pipeline) and n is the number of window samples. If any component was labeled "1" for motion, the feature set #1 sample label was "1" for motion. If a *maybe* label was present without *yes* motions on the other concurrent components, the window was excluded from training/testing. Feature set 2 included a target vector for each component but excluded the noisier vertical signals. These vertically concatenated components resulted in m features and $2\times n$ samples. In this extraction case any *maybe* labels were excluded from training and testing.

We employed a nested cross validation approach for unbiased testing of our dataset. We initialized 10 different folds of randomly splitting the 77 events into 90% training and 10% testing. By splitting on events we avoided "leakage" of information from our training into our testing, including correlation of seismic waveforms from any given event observed across a network. By cross validating over 10 folds we minimized biasing our result by the relatively small testing subsets of events, and can quantify the ability of our classification model to generalize for future events. Each event was observed by a different number of stations depending on network density and sensing radius, and each station-event pair had differing number of time samples; consequently the feature vectors of training and testing were not precisely 90/10 split in samples. In each fold, we held the test set aside as "unseen", and tuned our model using K-fold cross validation (Bishop & Nasrabadi, 2007) on the remaining training set (Figure 3). We implemented 5 inner folds in our K-fold cross validation to find the best hyperparameters. This cross validation approach allowed us to minimize overfitting the training dataset and evaluate the performance of our model on unseen data as though it were running such a classifier on yet-to-occur events.

The traditional "accuracy" metric, or the ratio of the correctly classified labels relative to the total number of labels, of our classification will be less sensitive regardless of optimization choices due to the infrequent events of our imbalanced classification. Instead, we optimized on metrics that reflect accurately classifying the infrequent events. Precision, or positive predictive value, is equal to the number of true positives (TP) over

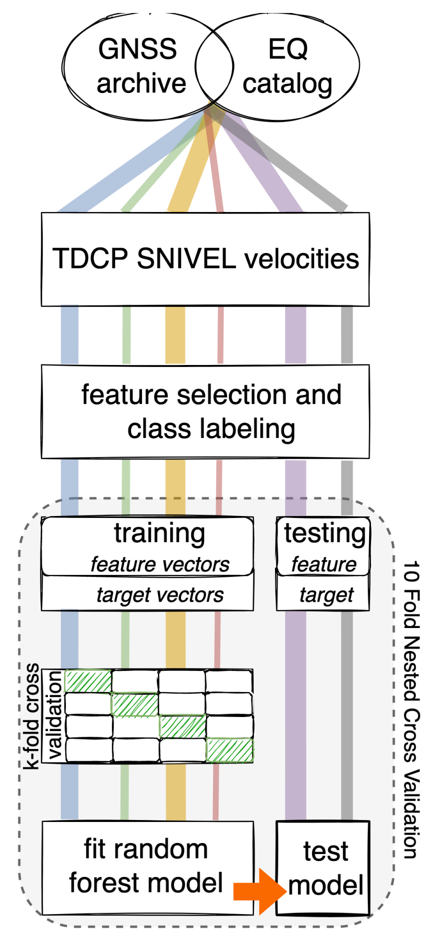


Figure 3. Schematic of a single fold random forest pipeline. For evaluation, we ran 10 folds of train/test splits of the 77 seismic events and report the mean and standard deviation of the test metrics to evaluate how well our features and models generalize across different testing sets.

Table 2. 10 fold nested cross validation results comparing Feature Set 1 is where all 3 components are combined for each window, and Feature Set 2 is where each horizontal component is tested independently.

		Feature Set #1	Feature Set #2
Precision	mean	0.79	0.73
	stdev	0.21	0.21
Recall	mean	0.71	0.67
	stdev	0.10	0.13
F1	mean	**0.73**	0.68
	stdev	0.14	0.14

the sum of TP and false negatives (FN).

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

Recall, or sensitivity, is the number of TP over the sum of TP and false positives (FP).

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

F1 is the harmonic mean of precision and recall:

$$F_1 = 2 \times \frac{precision \times recall}{precision + recall} \tag{3}$$

Here, positive denotes motion and vice versa.

Precision and recall are approximately inversely related and each is a function of our random forest decision threshold. Quantifying missed detections and false alert rates is imperative for the effectiveness of any EEW system (Minson et al., 2019). We optimized hyperparameters on F1 scores, a balance of precision and recall, but this parameter is a knob available to tune depending on societal missed detection of false alerting tolerances of a future operational system.

3 Results and Discussion

We evaluate the two optimal feature selection strategies and a range of random forest hyperparameters via a grid search. Given the F1 scores of our 10 fold nested cross validation approach (Table 2), our optimal model used feature set #1, with all available spatial components with a single target label to accumulate as much signal as possible towards our binary classification. Each train/test fold selected different optimal hyperparameter combinations for testing via cross validation, but the majority used 500 or 1000 decision trees, 100 decision splits and no class weighting. We used a decision threshold of 0.5 for this feature engineering approach (Figure 4) to optimize F1, a balance of precision and recall. Our mean and one standard deviation nested cross validation F1 score of 0.73±0.14 indicates our ability to successfully train a model using random forest. The variance in our results as a justifies our nested cross validation approach to quantify the variability in results as a function of the testing set; presumably some variability will resolve with expanded target catalogs.

3.1 Feature Importance

A benefit of random forest is that individual feature importance is readily extracted from the trained model. When evaluating feature set 1, we find several aspects of the

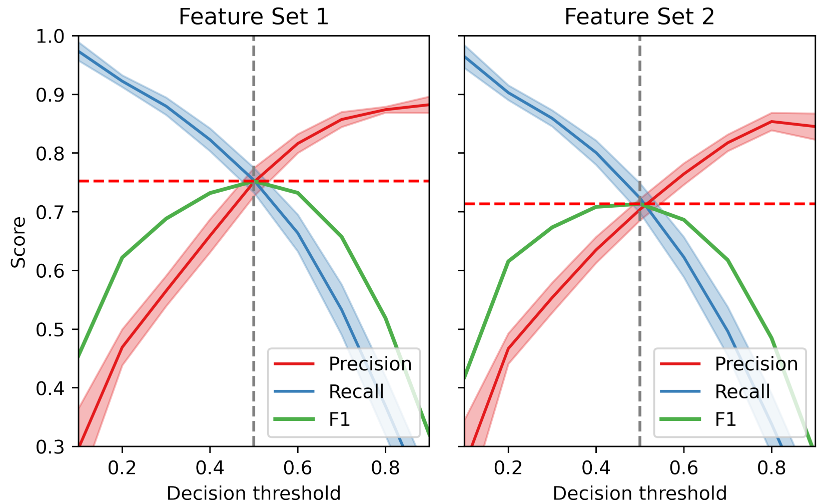


Figure 4. Mean precision, recall and F1 as a function of decision thresholds for the 10 fold nested cross validation evaluation. The shaded regions are the standard deviations across the 10 folds as a function of threshold. The dashed vertical lines are the maximum F1 decision threshold, with the dashed horizontal lines being the corresponding maximum F1 score.

feature importances that align with our domain knowledge and therefore contribute to the explainability of our trained model. The horizontal velocity components dominate the contribution to the model (Figure 5a). GNSS ambient noise on the vertical component is much higher than that of the horizontal components and vertical seismic signal amplitudes are diminished relative to horizontal motion along horizontal strike-slip fault mechanics that are common in the spatial region of this study. These less frequent signals amidst a higher relative noise floor were harder to detect and thus contributed less to the empirical classification model. Within a horizontal component, the lower frequency spectral features had the most influence (Figure 5b). The most important frequency bins were between 15-6 second periods, aligned with the prevalent frequencies of seismic surface waves. Our 5Hz sampling, as compared to lower rates, boosted the detectability around the noise floor, and avoided corner frequency aliasing of certain magnitudes. The time domain features contributed to the model, albeit much less than the lower frequency spectral content. After initial evaluation, we removed higher frequency power spectra from our features; these are logically "noise" in our system and were not contributing to classification. Altogether, these feature importances illustrate a key attribute of such a machine learning approach: combining features in an explainable way into an effective decision process.

3.2 Comparison with Existing Methods

A critical performance indicator is evaluating how our classification model performs over a range of test events relative to existing threshold approaches. Logic was applied to map existing continuous epoch-wise time domain threshold detection to our 30 second overlapping window target labels. For a threshold method comparison similar to the approach of Hodgkinson et al. (2020) and Dittmann et al. (2022), we estimated the noise threshold in the 2 minute window prior to seismic origin time. Hodgkinson et al. (2020) characterized the stand-alone sensitivity of detection using ambient noise antecedent to an event as a Gaussian heuristic threshold. Dittmann et al. (2022) approximated the 2 minute window of ground velocities as a non central chi-squared (NCX2) distribution

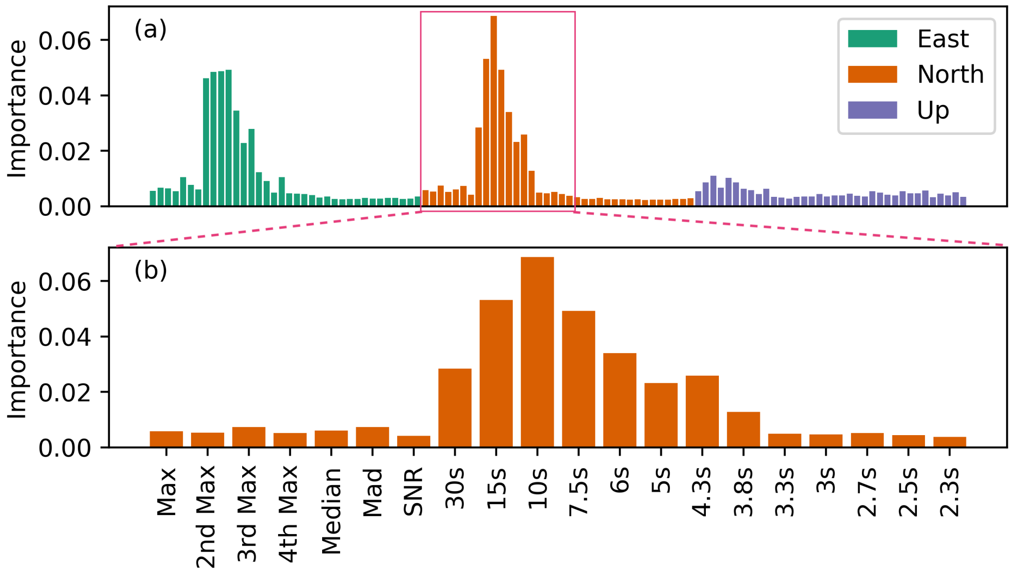


Figure 5. Feature importances from feature set #1 testing. 5a is the distribution of the importances across the horizontally concatenated, three spatial components. 5b is a close up of the east component, with the features labeled across the x axis. From the left, the first 6 of each component are time domain features (max, min, mad) within the 30 second windows; the next 15 are the power spectra from a periodogram of the 30 second 5Hz data, increasing in frequency from left to right. For reference, the periods are indicated.

with 3 degrees of freedom, and then set the 0.995 confidence level value of this distribution as a noise floor approximation. Any three dimensional GV magnitude above this noise threshold after this window is considered an event, and evaluated on whether it falls within a window labeled motion or not. RT-Shake (Psimoulis et al., 2018) evolved the previous geodetic STA/LTA algorithms (Allen & Ziv, 2011; Ohta et al., 2012) by differencing instantaneous measurements from 80 epoch moving averages and then related these values to a moving window noise threshold estimate set to three times the standard error of the previous 80 epochs. This method was run on each component independently, with a single boolean for the presence of motion on any component, and each sample window assigned a boolean based on the presence of any motion. The Dittmann et al. (2022) implementation of the threshold window in time was based upon S-Wave speeds (B. W. Crowell et al., 2013), and Psimoulis et al. (2018) modified STA/LTA correlated with surrounding stations to minimize false alerts; we did not add this logic so that we could simulate running as a stand-alone instrument.

The mean precision, F1 and accuracy from our 10 fold test of our random forest classifier outperforms the existing threshold approaches (Figure 6). In the threshold approach, recall is higher than the random forest classifier; given the large number of false positives that this method triggers, we believe this value is boosted by chance noise triggers occurring in windows of true motion triggering the motion boolean. This further demonstrates the value of optimizing on F1 as a balance of precision and recall to reduce biasing one decision criteria. Precision is low for both the threshold method and the STA/LTA, but for different reasons; while the precision values (Equation 1) are nearly identical, the threshold method suffers from a relatively high amount of false positives, whereas the STA/LTA method low score is due to a lower amount of true positives. This

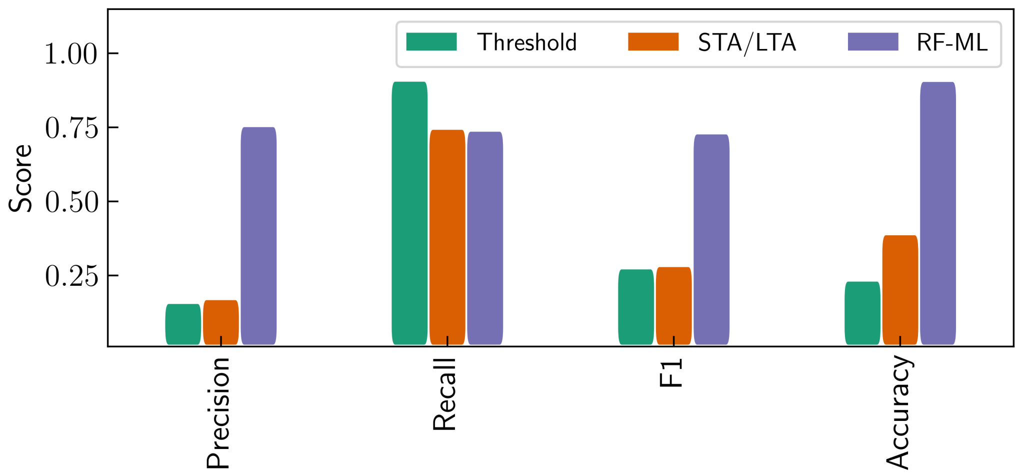


Figure 6. Performance metrics for 3 methods in stand-alone mode without external triggers or correlation. Threshold is the NCX2-995 approach used by Dittmann et al. (2022) that thresholds the noise based upon the 0.995 significance of a non-central chi-squared distribution of the ambient noise. STA/LTA is based on Psimoulis et al. (2018) GNSS motion detection modified STA/LTA algorithm. RF-ML is the method presented in the work here. Optimizing on F1 in this study allows us to balance missed detections (recall) with false alerts (precision); given the amount of false alerts of the Threshold and STA/LTA, the higher recall score could be a result of regular noise triggering events.

discrepancy is evident in the accuracy scores, where the STA/LTA outperforms the threshold approach. False positives would be decreased if using additional external information as their authors' suggest, such as stricter time window approaches and correlating in space within networks. Such an approach would also likely improve the random forest classifiers performance but limit the utility of a stand-alone detection node. Spatiotemporal information could be incorporated into future network decision criteria.

3.3 Edge Sensitivity Detection

Detecting the largest amplitude velocity waveforms relative to ambient noise does not present a significant challenge outside of mitigating false alerting from sporadic outliers (Figure 7), with a 98% true positive rate of events greater than M_W6.0 and less than 100km radius. The random forest classifier's balance of improved false alerting relative to thresholds and improved sensitivity relative to the STA/LTA is evident for these highest seismic risks. To further investigate the random forest model performance we evaluate detecting signals closer to the noise floor. For simplicity, we bin seismic motion edge case detection into two distinct classes in what is a continuous distribution: large magnitude event seismic motion detection in the far field, and smaller magnitude events detected in the nearfield.

In the relative nearfield, much of the seismic energy passes through a station in shorter duration, varied frequency signals. Earthquake focal depth and fault slip distribution in time and space can significantly vary these waveforms as observed. Critically, the waveform signatures can appear similar to those of non geophysical processing outliers which we wish to ignore for this classification. Most existing STA/LTA methods filter these noise signals but also these valuable higher frequency dynamics. In the previous threshold meth-

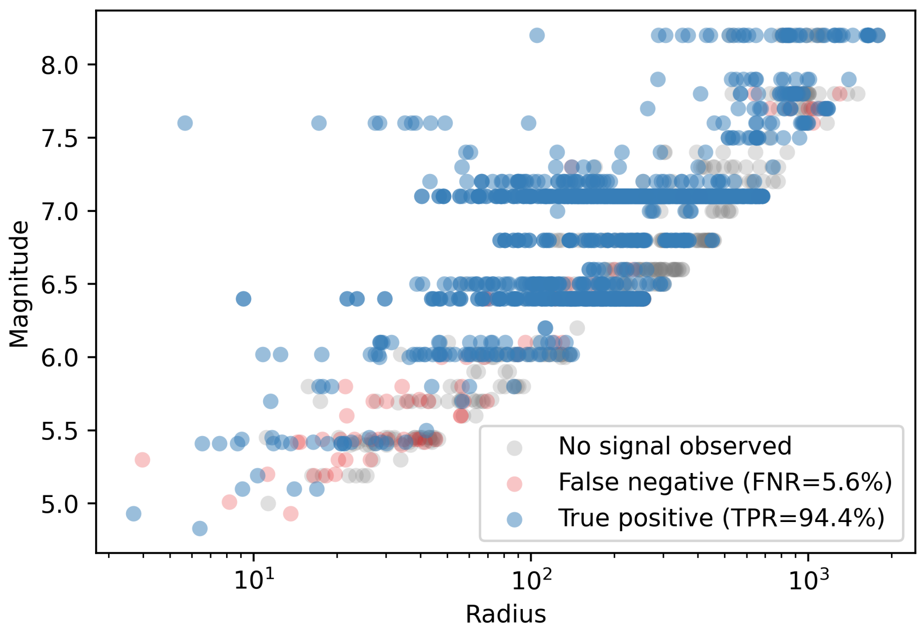


Figure 7. Performance of Random forest model developed in the work here across the entire event catalog. We reduce detection of events to a single binary for the figure. In this, each event is evaluated in a "test" split during the nested validation pipeline. This approach ensures each result depicted was evaluated as "unseen" relative to the best fit model from the training subset, and therefore representative of our model's future performance.

ods, detection of these edge cases was a function of the ambient noise level, with low precision resulting (Figure 6) as a result of a high false positive rate. Our classifier has far less false alerts than the threshold approach in these signals, but nevertheless still presents the hardest detection domain for our classifier, evident in the missed detections of Figure 7 of events less than $M_W 6.0$. Figure 8 is an example of a smaller magnitude event ($M_W 5.4$) in the relative nearfield (21km). In the top 4 panels it is evident that accurately detecting such an event using the threshold or modified STA/LTA approach is difficult; not only does the true signal barely exceed the noise floor, but there are numerous false alerts using both methods. The random forest classifier captures each labeled motion window in addition to "ignoring" the spurious signal around 100s OT that triggers all of the other methods evaluated.

The sensitivity of GNSS to long period surface waves are apparent at relatively great radii in the 5 hz TDCP velocity time series (Figure 7). The model detects teleseismic surface waves in unfiltered GNSS velocities at 1780km epicentral radius in real-time with no external corrections; Figure 9 provides an example of this detection. In Figure 9, the amplitude of the ground velocity magnitude of these long period signals is insufficient to cross the traditional noise threshold, and for that same noise threshold there are many false alerts. The modified STA/LTA RT-Shake approach does not identify the majority of the long period waves either, while the random forest classifier in the bottom panel only misses the first window.

3.4 Decision Latency

Delay in alerting is critical to EEW. While our model is trained, tested, and validated on overlapping windows every 10 seconds, we evaluate running the model at once per second, the current US EEW (ShakeAlert (Murray et al., 2018)) geodetic input rate (Figure 10). On testing data not used in model training, we find a delay relative to the P-wave ($\sim$3 seconds average at 10km) exists in the current approach. GNSS velocities using this current approach cannot reliably be used for initial phase (P-wave) picking, but can rapidly contribute to ground motion models or peak motion scaling laws (Fang et al., 2020). Given the feature importances of the classifier (Fig 5), delays are a result of coseismic energy organizing into surface waves which are confidently detected by the model. These are the signals we were visibly able to distinguish in labeling. Variance in delays in the near field are likely due to inherent limitations of modeling rupture as a point source at proximal locations (Goldberg et al., 2021). It is worth repeating that this assessment uses no external input or seismic triggering.

3.5 Ambient Noise Dataset

In addition to evaluating the performance within the bespoke event data sets, we also evaluate the performance of the method during a period of quiescence. We randomly selected 30 spatially distributed stations that supported 5 Hz downloads one hour in advance of the 2019 $M_W 6.4$ Ridgecrest Earthquake. We processed a 60 minute window to be representative of ambient noise: there were no events >M4.0 in the USGS comcat catalog, space weather indices were calm (Ap 4 nT) and all other sources of noise (signal multipath, oscillators, etc) were included. We assigned labels of *no motion* to all target vectors associated with feature extraction, and thus can evaluate ambient noise performance, or false alarm rate (Figure 11). We applied the previously trained classifier model once per second, to simulate running such a model in real-time mode.

Overall, the random forest classifier is less susceptible to spurious signals or outliers over the window tested than the threshold and STA/LTA approaches. As expected, the two threshold models are the most susceptible to false alerting; evident from the precision metric reported in Figure 6. The variations present in the random forest approach suggest that the current model has some station/time dependence not aligned with the

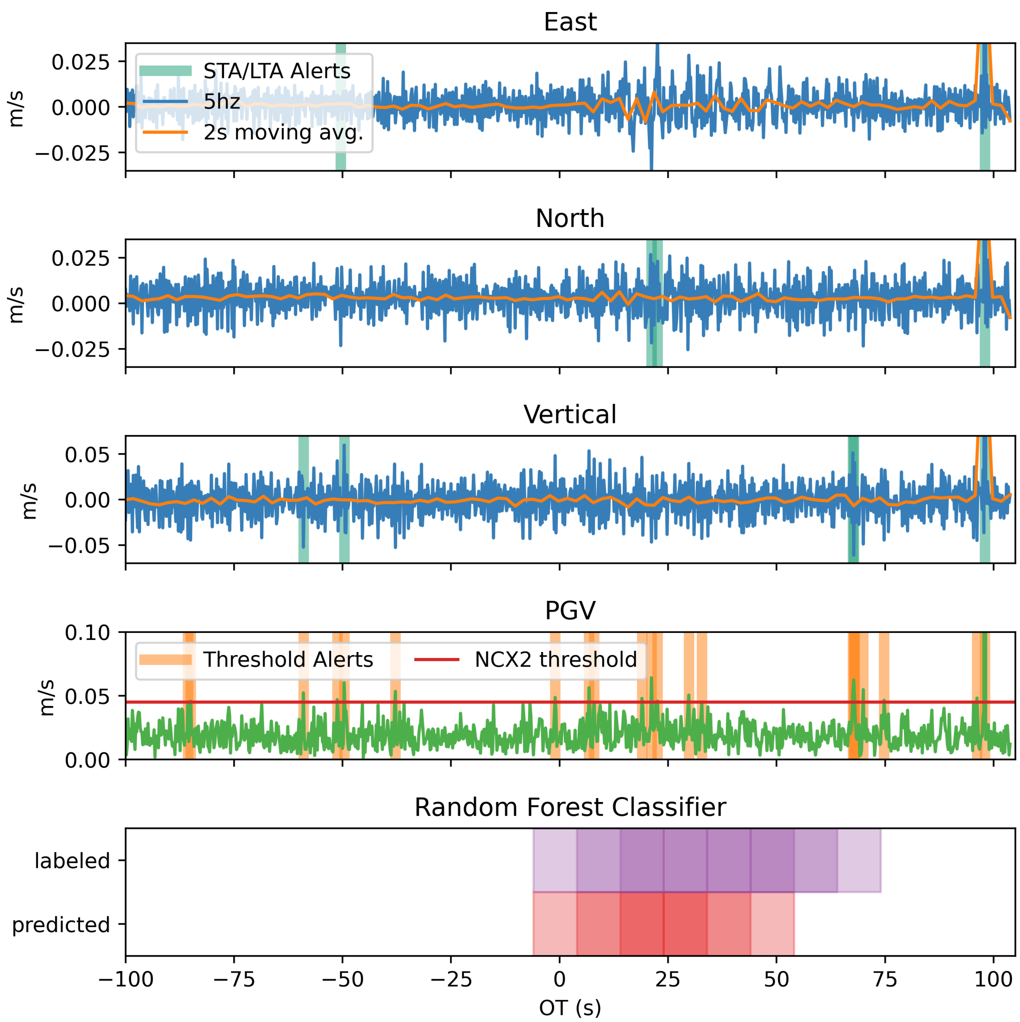


Figure 8. Velocity and detection time series from P507 observing a M5.41 at 21km epicentral radius. In the top 3 panels, we include a downsampled running mean so that the reader may readily visualize the lower frequency surface waves passing through. The teal vertical lines are alerts from the STA/LTA classifier (Psimoulis et al., 2018) on each component. The fourth panel green timeseries is the 3 component GV; the red horizontal line is the sensitivity threshold of a 0.995 non central chi squared (ncx2) noise model (Dittmann et al., 2022)), with orange vertical lines indicating a potential alert where GV greater than the threshold. The fifth panel is a comparison of the labeled feature set 1 for this event-station pair in purple, and the results of the model prediction in red. Shading is used to distinguish overlaps. This event-station pair prediction is extracted from the test or unseen event collection.

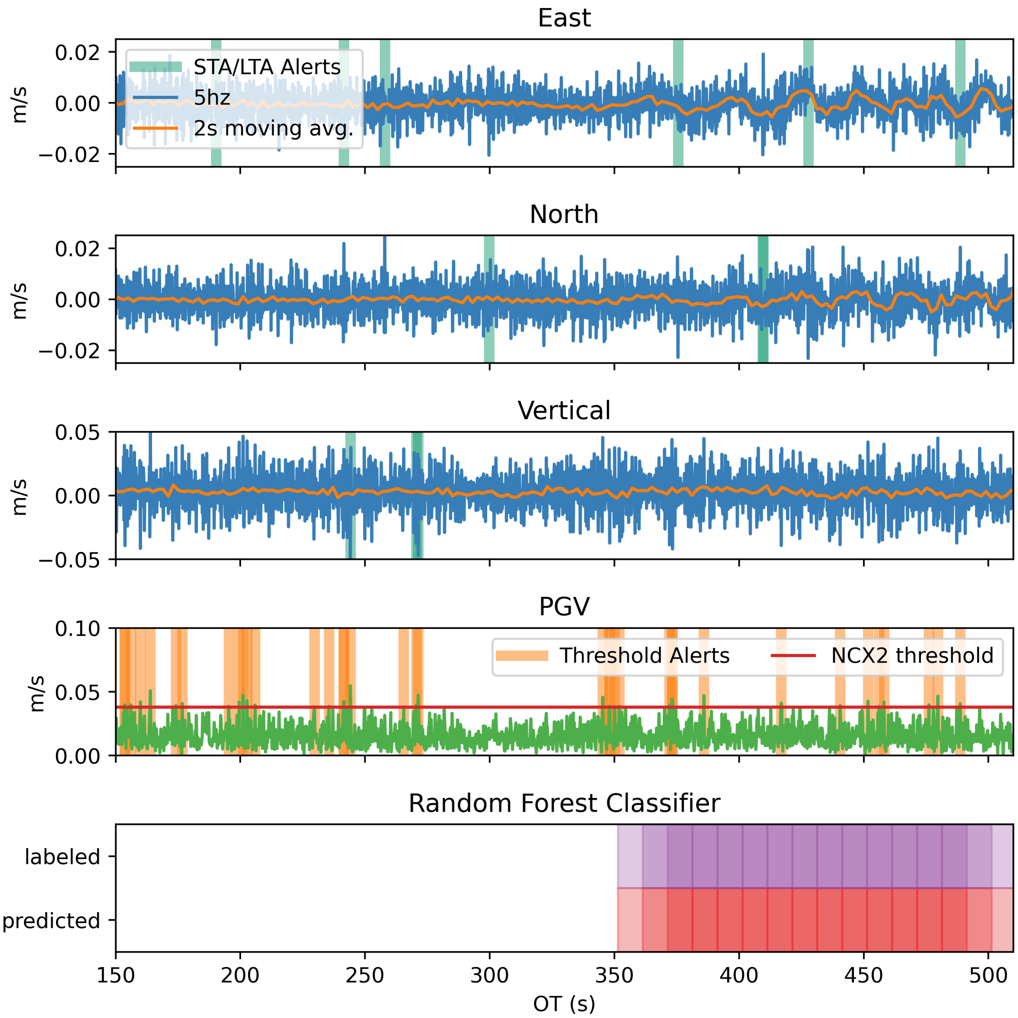


Figure 9. Velocity and detection time series from Station AB18 observing a M_W7.9 from ∼1400km epicentral radius. In the top 3 panels, we include a downsampled running mean so that the reader may readily visualize the lower frequency surface waves passing through, but these are not used in the models. The teal vertical lines are alerts from the STA/LTA classifier (Psimoulis et al., 2018) on each component. The fourth panel green timeseries is the 3 component GV; the red horizontal line is the sensitivity threshold of a 0.995 non central chi squared (ncx2) noise model (Dittmann et al., 2022), with orange vertical lines indicating a potential alert where GV greater than the threshold. The fifth panel is a comparison of the labeled feature set 1 for this event-station pair in purple, and the results of the model prediction in red. Shading is used to distinguish overlaps. This event-station pair prediction is extracted from the test or unseen event collection.

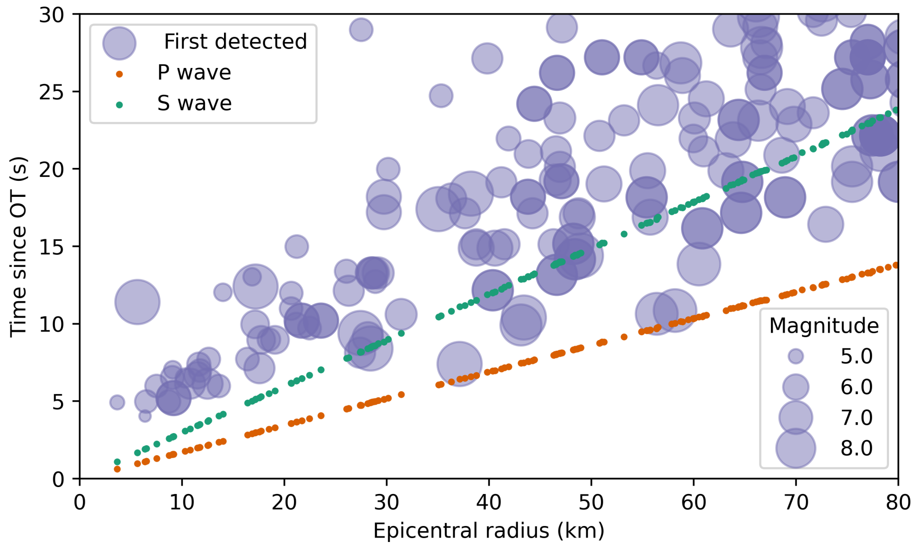


Figure 10. Time of first detection of all individual event-station pairs within 80km epicentral radius relative to earthquake origin time (OT) as a function of radius. Green dots are the estimated P- and S-wave arrivals at the event-station pairs used in this study shown for reference. Purple circles are centered on the time of first detection after the OT, where the diameter is scaled to the event magnitude. These results are from the classifier run at 1Hz on unseen testing sets to simulate a real-time operational mode.

variations of other methods. Inclusion of larger noise training datasets into our detection classifier and possibly data augmentation techniques would likely be beneficial towards training on the widest variety of noise scenarios.

4 Conclusion

We applied an existing machine learning algorithm and sample splitting pipeline techniques to training, validating and testing a seismic motion detection classifier from 5Hz TDCP GNSS velocities. We leveraged nearly 20 years of 5Hz GNSS data archives for training a classification model that outperforms existing threshold approaches for detecting motion in stand-alone mode. The classifier combines time domain and frequency domain features to match the sensitivity of the threshold method without the false alerts, and matches the minimal false alerting of the STA/LTA with improved sensitivity. Given the agreement that GNSS velocities have with existing ground motion models (B. Crowell et al., 2022) and the increased confidence in separating signal from noise demonstrated here, these GNSS velocities can operationally contribute to ground motion measurements. The alert latency of this current model does not match the sensitivity of existing inertial infrastructure. A complementary approach using the information available at the time, including lowest latency p-wave characterization from inertial sensors and unsaturated velocity estimation from GNSS provides an optimal solution for existing dense multi-sensor networks. For less dense networks of either sensor type, it is more critical to establish a decision criteria for balancing timing, noise and accuracy of these independent observation systems. Further investigation of integrating the processing and classifying approach of this manuscript with the sensitivity of co-located MEMS sensors (Goldberg & Bock, 2017) would advantageously overlap seismic and geodetic traditional boundaries.

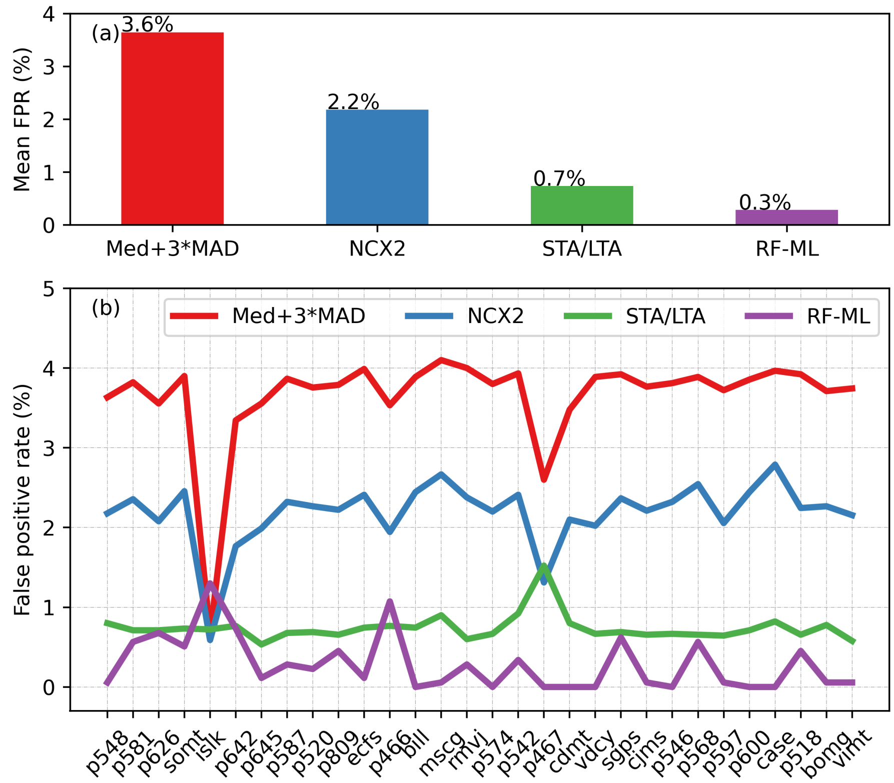


Figure 11. Panel (a) is mean false positive rates (FPR) from 30 randomly selected, spatially distributed, TDCP 5Hz velocities during the same 60 minute time window (1600-1700 4 July 2019). Methods include: median plus 3 times the median absolute deviation threshold of Hodgkinson et al. (2020), non-central chi square of Dittmann et al. (2022) NCX2 using alpha value of 0.995, the modified STA/LTA implemented by Psimoulis et al. (2018) and the random forest machine learning classifier developed in this work (RF-ML). Panel (b) is the distribution by station of each method.

Current 5 Hz GNSS observation data streams are too verbose for many bandwidth limited remote hardware; this presents an exciting opportunity for edge processing at potentially much higher rates (Shu et al., 2018), or experimental lean 5 Hz carrier phase data streams. Our method presented here does not use a sophisticated machine learning model, yet has improved detection relative to existing approaches; much improvement remains, especially with expanded datasets across global networks and/or synthetics or data augmentation for training, validation and testing of neural networks and deep learning models.

With an expanding availability and access to real-time GNSS streaming networks, the seismological community stands to benefit from this signal of opportunity for rapid ground motion detection for earthquake and tsunami source characterization. Furthermore, the vast industry of GNSS position, navigation and timing users catalyzing the expansion of these GNSS real-time networks will benefit from improved automated alerting of reference station motion onset. Future work will include integrating this classifier amongst existing and future automated GNSS carrier phase disturbance characterization methods, including space weather disturbances (Jiao et al., 2017), oscillator anomalies (Liu & Morton, 2022), radio frequency interference and signal multipath.

5 Open Research

The 5Hz GNSS data used for TDCP processing in the study are available from the Geodetic Facility for the Advancement of Geoscience (GAGE) Global Navigation Satellite Systems (GNSS) archives as maintained by UNAVCO, Inc. The data are available in RINEX (v.2.11) format at `https://data.unavco.org/archive/gnss/highrate/5-Hz/rinex/`. Earthquake depths, locations, and magnitudes came from the Advanced National Seismic System (ANSS) Comprehensive Catalog of Earthquake Events and Products (`https://earthquake.usgs.gov/data/comcat/`). Arrival times are calculated using the iasp91 velocity model as implemented by Incorporated Research Institutions for Seismology (IRIS) Web Services (`http://service.iris.edu/irisws/traveltime/`). SNIVEL code used for TDCP velocity processing is developed openly at `https://github.com/crowellbw/SNIVEL` (Accessed December 2021)(B. W. Crowell, 2021). SNIVEL 5Hz velocity timeseries used in this study are preserved at `https://doi.org/10.5281/zenodo.6588601`. Version 1.0.1 of the scikit-learn software used for random forest classification is preserved at `https://doi.org/10.5281/zenodo.5596244` and developed openly at `https://github.com/scikit-learn/scikit-learn`. (Pedregosa et al., 2011) Version v0.5.0 of PyGMT used for generating the map is preserved at `https://doi.org/10.5281/zenodo.5607255` and developed openly at `https://github.com/GenericMappingTools/pygmt`(Wessel et al., 2019)

Acknowledgments
This material is based on services provided by the GAGE Facility, operated by UNAVCO, Inc., with support from the National Science Foundation, the National Aeronautics and Space Administration, and the U.S. Geological Survey under NSF Cooperative Agreement EAR-1724794. High-rate processing and and machine learning for geoscience and hazards research is supported by NSF OAC-1835791

References

Allen, R. M., & Ziv, A. (2011). Application of real-time GPS to earthquake early warning. *Geophysical Research Letters*, *38*(16), n/a–n/a. doi: 10.1029/2011gl047947

Bishop, C. M., & Nasrabadi, N. M. (2007). Pattern Recognition and Machine Learning. *J. Electronic Imaging*, *16*, 049901.

Bock, Y., Prawirodirdjo, L., & Melbourne, T. I. (2004). Detection of arbitrarily large dynamic ground motions with a dense high-rate GPS network. *Geophysical Research Letters*, *31*(6), n/a–n/a. doi: 10.1029/2003gl019150

Branzanti, M., Colosimo, G., Crespi, M., & Mazzoni, A. (2013). Gps near-real-time coseismic displacements for the great tohoku-oki earthquake. *IEEE Geoscience and Remote Sensing Letters*, *10*(2), 372-376. doi: 10.1109/LGRS.2012.2207704

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. doi: 10.1023/a:1010933404324

Colombelli, S., Allen, R. M., & Zollo, A. (2013). Application of real-time GPS to earthquake early warning in subduction and strike-slip environments. *Journal of Geophysical Research: Solid Earth*, *118*(7), 3448–3461. doi: 10.1002/jgrb.50242

Colosimo, G., Crespi, M., & Mazzoni, A. (2011). Real-time GPS seismology with a stand-alone receiver: A preliminary feasibility demonstration. *Journal of Geophysical Research: Solid Earth (1978–2012)*, *116*(B11), n/a–n/a. doi: 10.1029/2010jb007941

Crocetti, L., Schartner, M., & Soja, B. (2021). Discontinuity Detection in GNSS Station Coordinate Time Series Using Machine Learning. *Remote Sensing*, *13*(19), 3906. doi: 10.3390/rs13193906

Crowell, B., DeGrande, J., Dittmann, T., & Ghent, J. (2022). *Validation of peak ground velocities recorded on very-highrate gnss against nga-west2 ground motion models.* (Seismological Society of America Annual Meeting 2022)

Crowell, B. W. (2021). Near-Field Strong Ground Motions from GPS-Derived Velocities for 2020 Intermountain Western United States Earthquakes. *Seismological Research Letters*, *92*(2A), 840–848. doi: 10.1785/0220200325

Crowell, B. W., Bock, Y., & Squibb, M. B. (2009). Demonstration of Earthquake Early Warning Using Total Displacement Waveforms from Real-time GPS Networks. *Seismological Research Letters*, *80*(5), 772–782. doi: 10.1785/gssrl.80.5.772

Crowell, B. W., Melgar, D., Bock, Y., Haase, J. S., & Geng, J. (2013). Earthquake magnitude scaling using seismogeodetic data. *Geophysical Research Letters*, *40*(23), 6089–6094. doi: 10.1002/2013gl058391

Dittmann, T., Hodgkinson, K., Morton, J., Mencin, D., & Mattioli, G. S. (2022). Comparing Sensitivities of Geodetic Processing Methods for Rapid Earthquake Magnitude Estimation. *Seismological Research Letters*, *93*(3), 1497–1509. doi: 10.1785/0220210265

Dybing, S., Melgar, D., Thomas, A., Hodgkinson, K., & Mencin, D. (2021). *Detecting earthquakes in noisy real-time gnss data with machine learning.* Retrieved from `\url{https://agu.confex.com/agu/fm21/meetingapp.cgi/Paper/820583}` (American Geophysical Union Fall Meeting 2021)

Fang, R., Zheng, J., Geng, J., Shu, Y., Shi, C., & Liu, J. (2020). Earthquake Magnitude Scaling Using Peak Ground Velocity Derived from High-Rate GNSS Observations. *Seismological Research Letters*, *92*(1), 227–237. doi: 10.1785/0220190347

Fratarcangeli, F., Savastano, G., D'Achille, M. C., Mazzoni, A., Crespi, M., Riguzzi, F., ... Pietrantonio, G. (2018). Vadase reliability and accuracy of real-time displacement estimation: Application to the central italy 2016 earthquakes. *Remote Sensing*, *10*(8). Retrieved from https://www.mdpi.com/2072-4292/10/8/1201 doi: 10.3390/rs10081201

Goldberg, D. E., & Bock, Y. (2017). Self-contained local broadband seismogeodetic early warning system: Detection and location. *Journal of Geophysical Research: Solid Earth*, *122*(4), 3197–3220. doi: 10.1002/2016jb013766

Goldberg, D. E., Melgar, D., Hayes, G. P., Crowell, B. W., & Sahakian, V. J. (2021, 08). A Ground-Motion Model for GNSS Peak Ground Displacement. *Bulletin* *of the Seismological Society of America*, *111*(5), 2393-2407. Retrieved from https://doi.org/10.1785/0120210042 doi: 10.1785/0120210042

Grapenthin, R., West, M., & Freymueller, J. (2017). The Utility of GNSS for Earthquake Early Warning in Regions with Sparse Seismic NetworksThe Utility of GNSS for Earthquake Early Warning in Regions with Sparse Seismic Networks. *Bulletin of the Seismological Society of America*, *107*(4), 1883–1890. doi: 10.1785/0120160317

Grapenthin, R., West, M., Tape, C., Gardine, M., & Freymueller, J. (2018). Single-Frequency Instantaneous GNSS Velocities Resolve Dynamic Ground Motion of the 2016 Mw 7.1 Iniskin, Alaska, Earthquake. *Seismological Research Letters*, *89*(3), 1040–1048. doi: 10.1785/0220170235

Habboub, M., Psimoulis, P. A., Bingley, R., & Rothacher, M. (2020). A Multiple Algorithm Approach to the Analysis of GNSS Coordinate Time Series for Detecting Geohazards and Anomalies. *Journal of Geophysical Research: Solid Earth*, *125*(2). doi: 10.1029/2019jb018104

Hodgkinson, K. M., Mencin, D. J., Feaux, K., Sievers, C., & Mattioli, G. S. (2020). Evaluation of Earthquake Magnitude Estimation and Event Detection Thresholds for Real-Time GNSS Networks: Examples from Recent Events Captured by the Network of the Americas. *Seismological Research Letters*, *91*(3), 1628–1645. doi: 10.1785/0220190269

Hohensinn, R., & Geiger, A. (2018). Stand-Alone GNSS Sensors as Velocity Seismometers: Real-Time Monitoring and Earthquake Detection. *Sensors*, *18*(11), 3712. doi: 10.3390/s18113712

Jiao, Y., Hall, J. J., & Morton, Y. T. (2017). Performance Evaluation of an Automatic GPS Ionospheric Phase Scintillation Detector Using a Machine-Learning Algorithm. *Navigation*, *64*(3), 391–402. doi: 10.1002/navi.188

Joyner, W. (August 1984). A SCALING LAW FOR THE SPECTRA OF LARGE EARTHQUAKES. *Bulletin of the Seismological Society of America*, *74*(4), 1167–1188.

Kawamoto, S., Hiyama, Y., Ohta, Y., & Nishimura, T. (2016). First result from the GEONET real-time analysis system (REGARD): the case of the 2016 Kumamoto earthquakes. *Earth, Planets and Space*, *68*(1), 190. doi: 10.1186/s40623-016-0564-4

Kong, Q., Trugman, D. T., Ross, Z. E., Bianco, M. J., Meade, B. J., & Gerstoft, P. (2019). Machine Learning in Seismology: Turning Data into Insights. *Seismological Research Letters*, *90*(1), 3–14. doi: 10.1785/0220180259

Larson, K. M. (2009). GPS seismology. *Journal of Geodesy*, *83*(3-4), 227–233. doi: 10.1007/s00190-008-0233-x

Larson, K. M., Bodin, P., & Gomberg, J. (2003). Using 1-Hz GPS Data to Measure Deformations Caused by the Denali Fault Earthquake. *Science*, *300*(5624), 1421–1424. doi: 10.1126/science.1084531

Lin, J., Melgar, D., Thomas, A. M., & Searcy, J. (2021). Early Warning for Great Earthquakes From Characterization of Crustal Deformation Patterns With Deep Learning. *Journal of Geophysical Research: Solid Earth*, *126*(10). doi: 10.1029/2021jb022703

Liu, Y., & Morton, Y. J. (2022). Improved Automatic Detection of GPS Satellite Oscillator Anomaly using a Machine Learning Algorithm. *NAVIGATION: Journal of the Institute of Navigation*, *69*(1), navi.500. doi: 10.33012/navi.500

Meier, M., Ross, Z. E., Ramachandran, A., Balakrishna, A., Nair, S., Kundzicz, P., … Yue, Y. (2019). Reliable Real-Time Seismic Signal/Noise Discrimination With Machine Learning. *Journal of Geophysical Research: Solid Earth*, *124*(1), 788–800. doi: 10.1029/2018jb016661

Melgar, D., Crowell, B. W., Geng, J., Allen, R. M., Bock, Y., Riquelme, S., … Ganas, A. (2015). Earthquake magnitude calculation without saturation from the scaling of peak ground displacement. *Geophysical Research Letters*, *42*(13), 5197–5205. doi: 10.1002/2015gl064278

Melgar, D., Crowell, B. W., Melbourne, T. I., Szeliga, W., Santillan, M., & Scrivner, C. (2020). Noise Characteristics of Operational Real-Time High-Rate GNSS Positions in a Large Aperture Network. *Journal of Geophysical Research: Solid Earth*, *125*(7). doi: 10.1029/2019jb019197

Michelini, A., Cianetti, S., Gaviano, S., Giunchi, C., Jozinović, D., & Lauciani, V. (2021). INSTANCE – the Italian seismic dataset for machine learning. *Earth System Science Data*, *13*(12), 5509–5544. doi: 10.5194/essd-13-5509-2021

Minson, S. E., Baltay, A. S., Cochran, E. S., Hanks, T. C., Page, M. T., McBride, S. K., ... Meier, M.-A. (2019). The Limits of Earthquake Early Warning Accuracy and Best Alerting Strategy. *Scientific Reports*, *9*(1), 2478. doi: 10.1038/s41598-019-39384-y

Minson, S. E., Murray, J. R., Langbein, J. O., & Gomberg, J. S. (2014). Real-time inversions for finite fault slip models and rupture geometry based on high-rate GPS data. *Journal of Geophysical Research: Solid Earth*, *119*(4), 3201–3231. doi: 10.1002/2013jb010622

Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., & Beroza, G. C. (2020). Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nature Communications*, *11*(1), 3952. doi: 10.1038/s41467-020-17591-w

Mousavi, S. M., Sheng, Y., Zhu, W., & Beroza, G. C. (2019). STanford EArthquake Dataset (STEAD): A Global Data Set of Seismic Signals for AI. *IEEE Access*, *7*, 179464–179476. doi: 10.1109/access.2019.2947848

Murray, J. R., Crowell, B. W., Grapenthin, R., Hodgkinson, K., Langbein, J. O., Melbourne, T., ... Schmidt, D. A. (2018). Development of a Geodetic Component for the U.S. West Coast Earthquake Early Warning System. *Seismological Research Letters*, *89*(6), 2322–2336. doi: 10.1785/0220180162

Nikolaidis, R. M., Bock, Y., Jonge, P. J. d., Shearer, P., Agnew, D. C., & Domselaar, M. V. (2001). Seismic wave observations with the Global Positioning System. *Journal of Geophysical Research: Solid Earth*, *106*(B10), 21897–21916. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001JB000329` doi: 10.1029/2001jb000329

Ohta, Y., Kobayashi, T., Tsushima, H., Miura, S., Hino, R., Takasu, T., ... Umino, N. (2012). Quasi real-time fault model estimation for near-field tsunami forecasting based on RTK-GPS analysis: Application to the 2011 Tohoku-Oki earthquake (Mw 9.0). *Journal of Geophysical Research: Solid Earth (1978–2012)*, *117*(B2), n/a–n/a. doi: 10.1029/2011jb008750

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, E. (2011). Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.*, *12*(null), 2825–2830.

Psimoulis, P. A., Houlié, N., Habboub, M., Michel, C., & Rothacher, M. (2018). Detection of ground motions using high-rate GPS time-series. *Geophysical Journal International*, *214*(2), 1237–1251. doi: 10.1093/gji/ggy198

Ruhl, C. J., Melgar, D., Geng, J., Goldberg, D. E., Crowell, B. W., Allen, R. M., ... D'Anastasio, E. (2019). A Global Database of Strong-Motion Displacement GNSS Recordings and an Example Application to PGD Scaling. *Seismological Research Letters*, *90*(1), 271–279. doi: 10.1785/0220180177

Shu, Y., Fang, R., Li, M., Shi, C., Li, M., & Liu, J. (2018). Very high-rate GPS for measuring dynamic seismic displacements without aliasing: performance evaluation of the variometric approach. *GPS Solutions*, *22*(4), 121. doi: 10.1007/s10291-018-0785-z

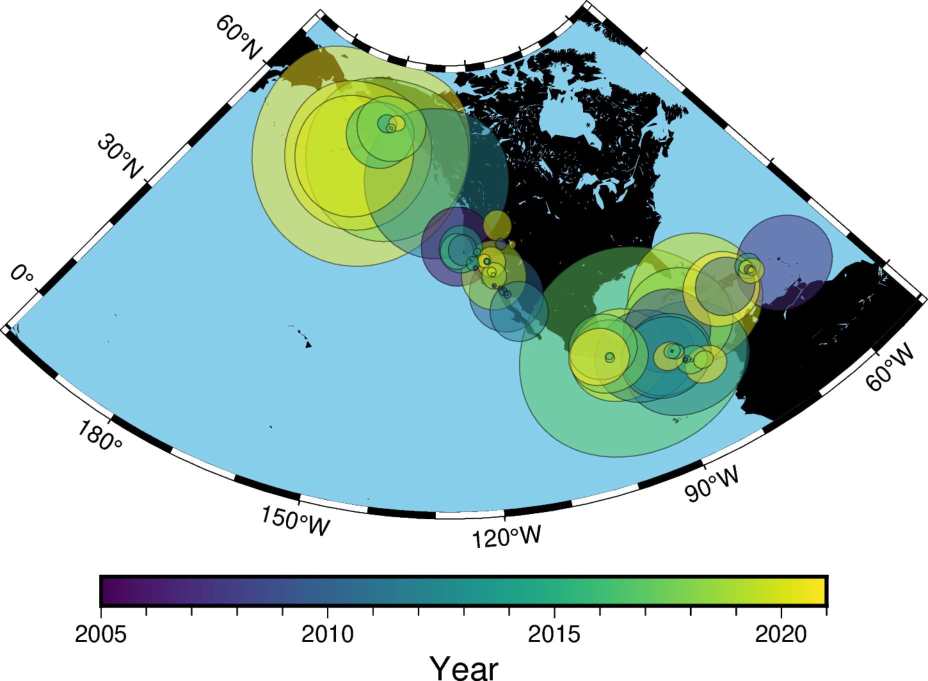
Smalley, R. (2009). High-rate GPS: How High Do We Need to Go? *Seismological Research Letters*, *80*(6), 1054–1061. doi: 10.1785/gssrl.80.6.1054

van Graas, F., & Soloviev, A. (2004). Precise Velocity Estimation Using a Stand-

Alone GPS Receiver. *Navigation*, *51*(4), 283–292. doi: 10.1002/j.2161-4296.2004.tb00359.x

Wessel, P., Luis, J. F., Uieda, L., Scharroo, R., Wobbe, F., Smith, W. H. F., & Tian, D. (2019). The generic mapping tools version 6. *Geochemistry, Geophysics, Geosystems*, *20*(11), 5556-5564. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GC008515` doi: https://doi.org/10.1029/2019GC008515

figure 1.



60°N
30°N
0°
180°
150°W
120°W
90°W
60°W
2005
2010
2015
2020
Year

figure 2.

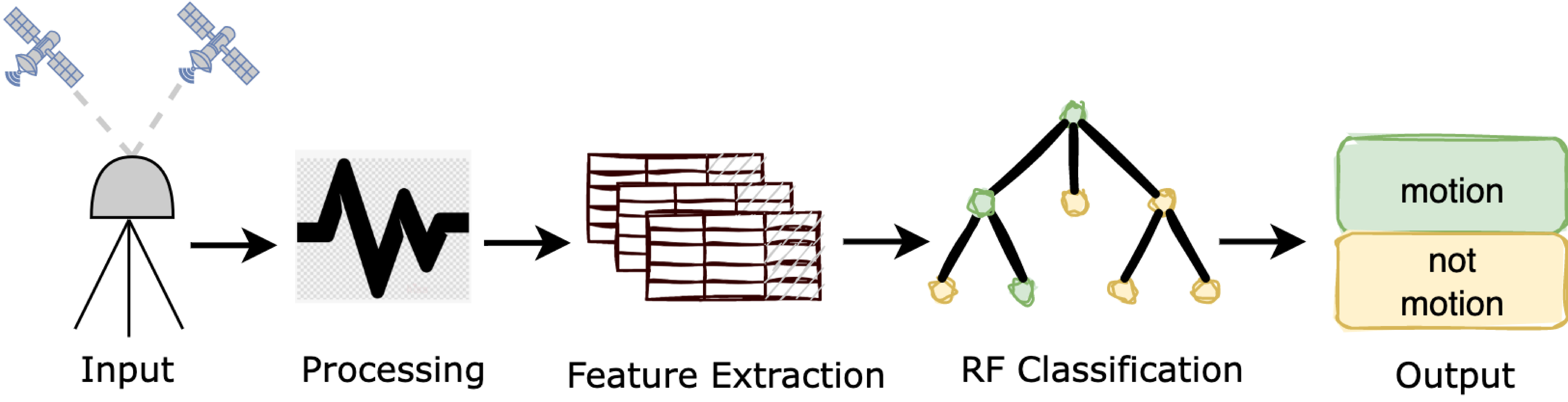

motion
not motion
Input
Processing
Feature Extraction
RF Classification
Output

figure 3.


GNSS archive
EQ catalog
TDCP SNIVEL velocities
feature selection and class labeling
training
feature vectors
target vectors
testing
feature
target
k-fold cross validation
fit random forest model
test model
10 Fold Nested Cross Validation

figure 4.

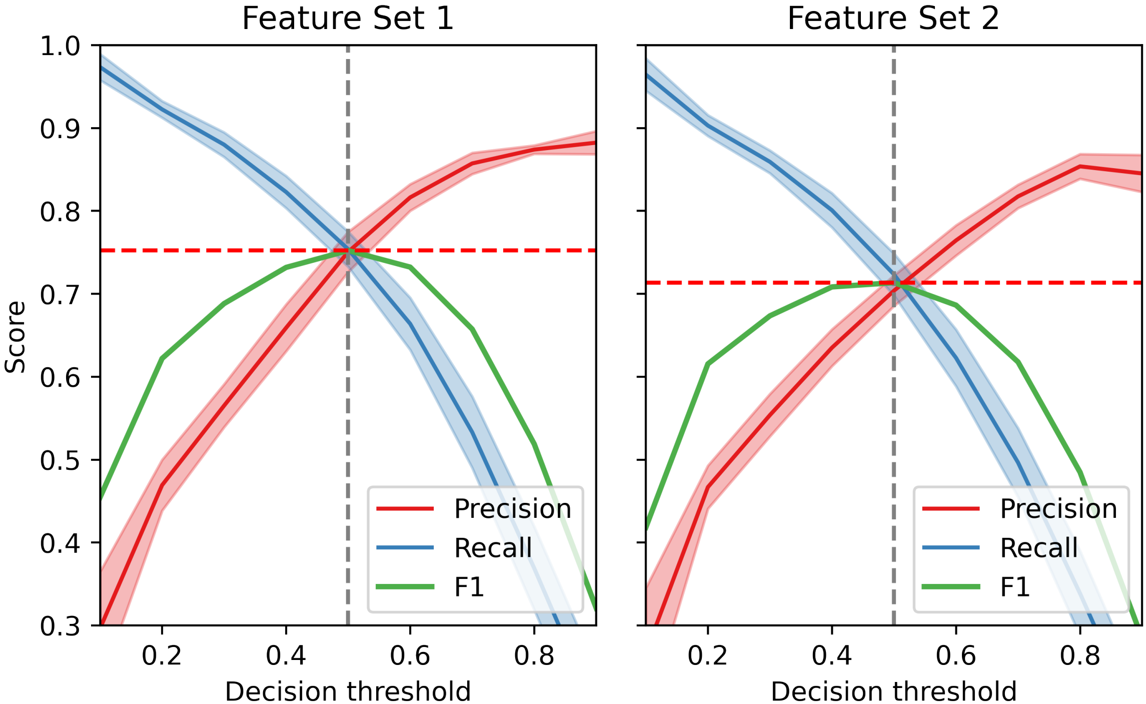
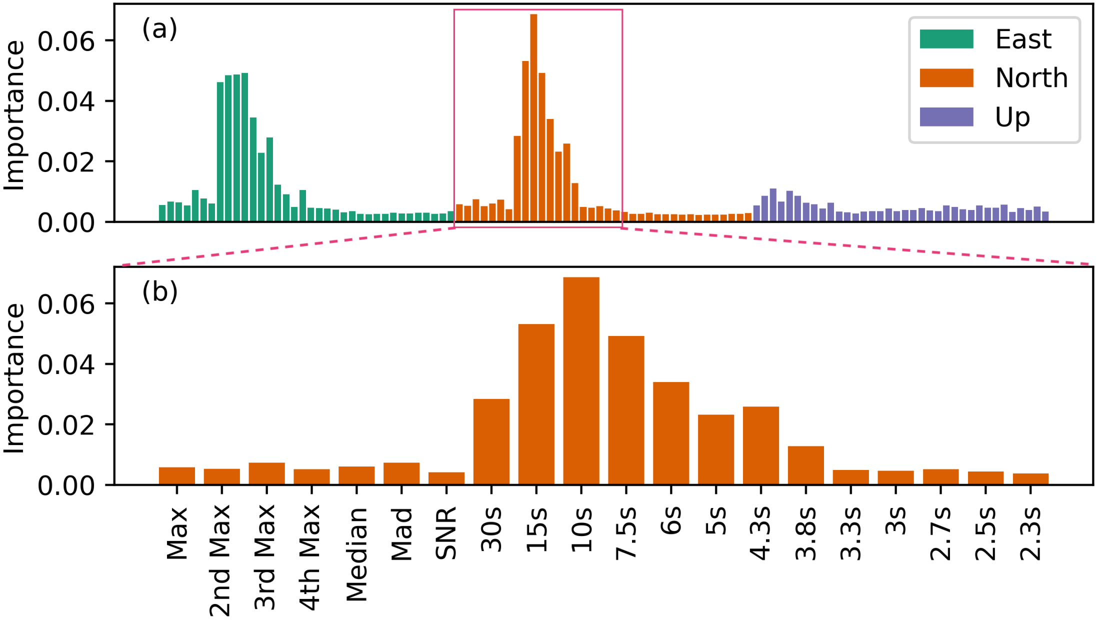
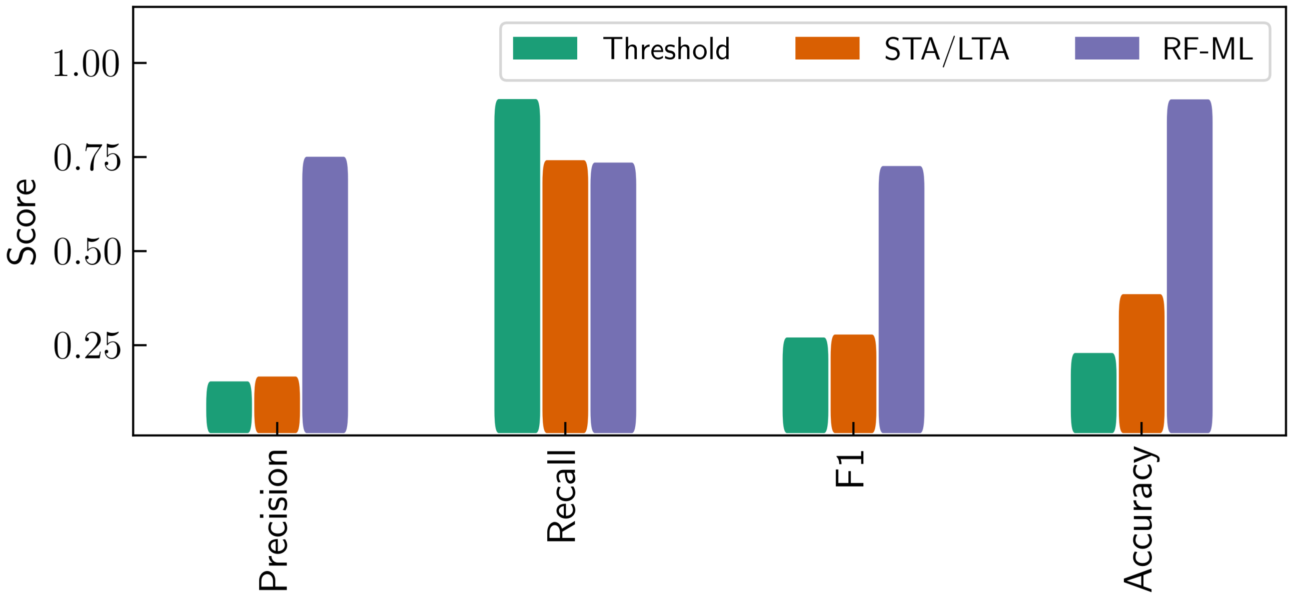

Feature Set 1
Feature Set 2
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
Score
0.2
0.4
0.6
0.8
Decision threshold
Precision
Recall
F1

figure 5.



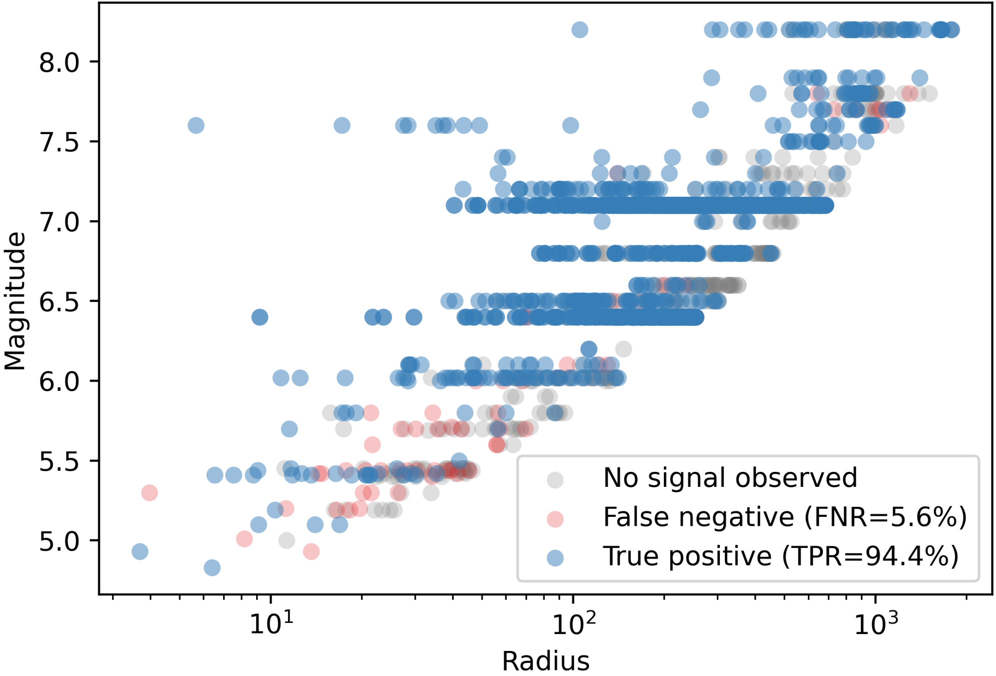
(a)
(b)
East
North
Up
Importance
0.00
0.02
0.04
0.06
Max
2nd Max
3rd Max
4th Max
Median
Mad
SNR
30s
15s
10s
7.5s
6s
5s
4.3s
3.8s
3.3s
3s
2.7s
2.5s
2.3s

figure 6.



Threshold
STA/LTA
RF-ML
Score
1.00
0.75
0.50
0.25
Precision
Recall
F1
Accuracy

figure 7.



8.0
7.5
7.0
6.5
6.0
5.5
5.0
Magnitude
10^1
10^2
10^3
Radius
No signal observed
False negative (FNR=5.6%)
True positive (TPR=94.4%)

figure 8.

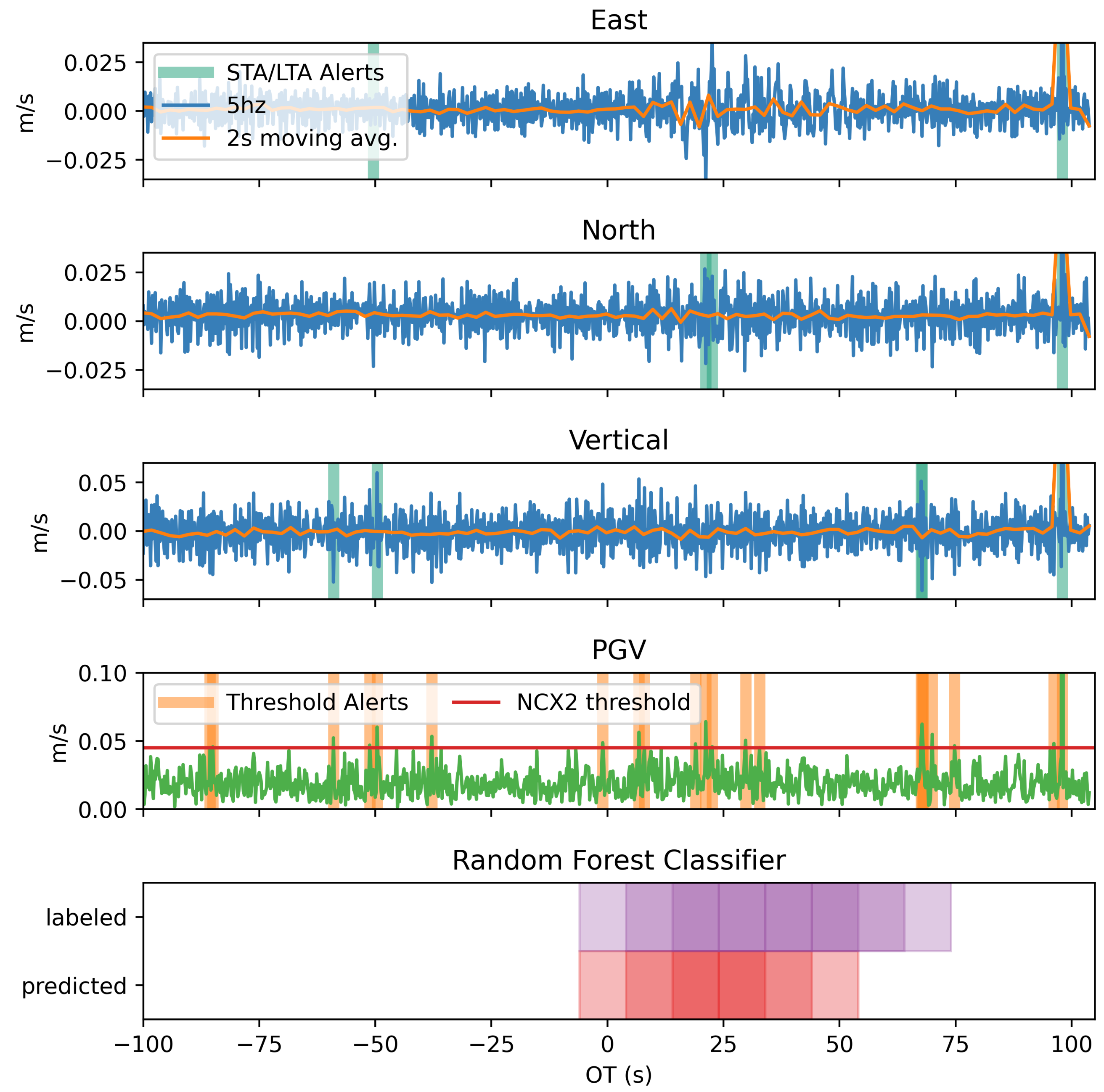

East
STA/LTA Alerts
5hz
2s moving avg.
m/s
0.025
0.000
−0.025
North
Vertical
0.05
0.00
−0.05
PGV
Threshold Alerts
NCX2 threshold
0.10
Random Forest Classifier
labeled
predicted
−100
−75
−50
−25
0
25
50
75
100
OT (s)

figure 9.

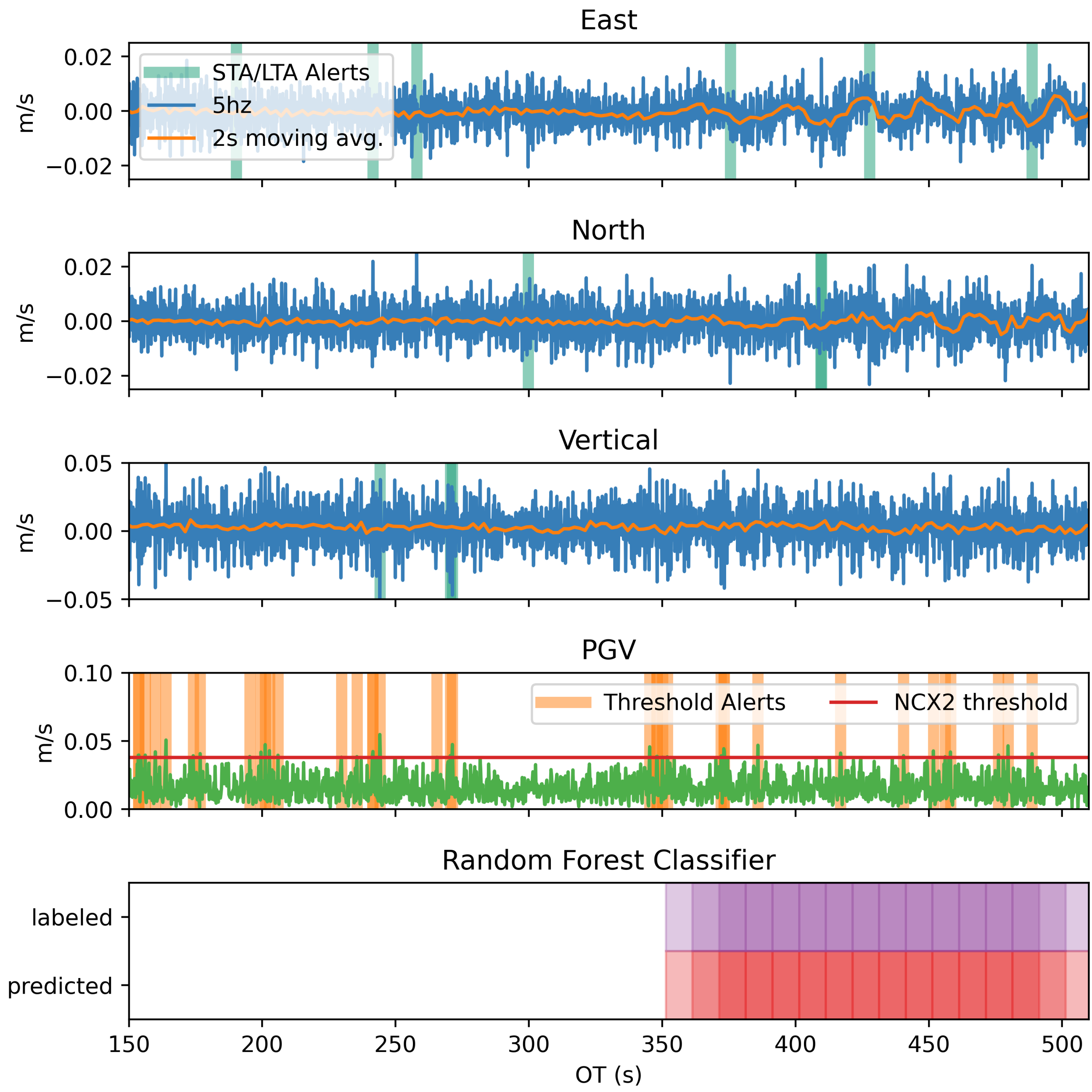
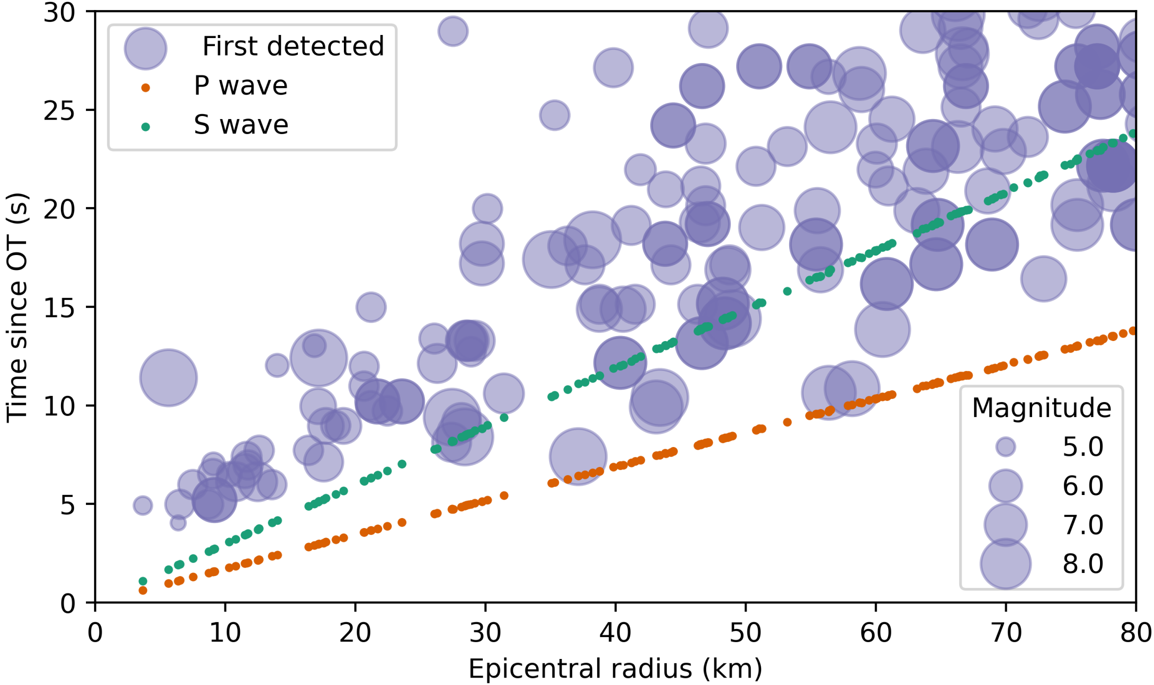
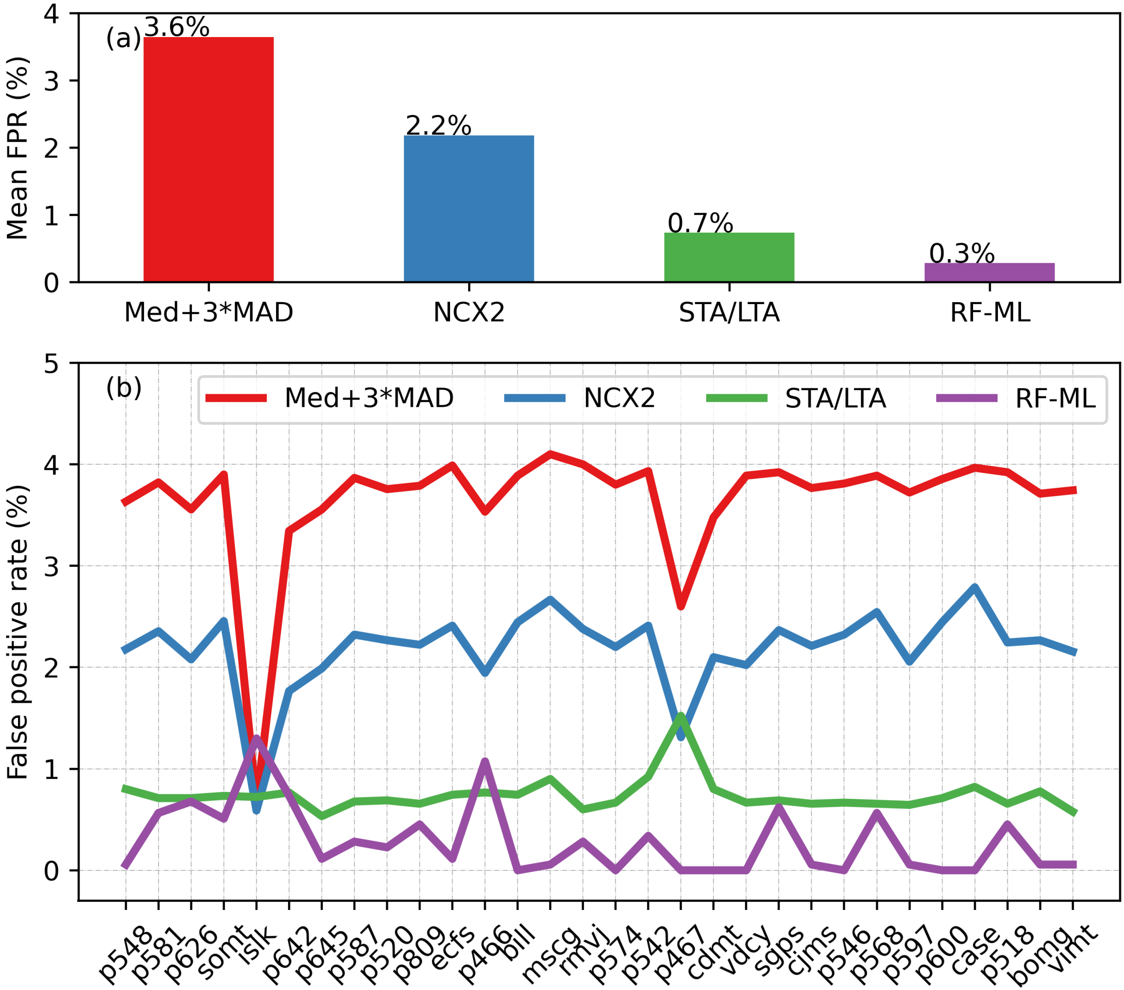

East
m/s
0.02
0.00
−0.02
STA/LTA Alerts
5hz
2s moving avg.
North
Vertical
0.05
−0.05
PGV
0.10
0.00
Threshold Alerts
NCX2 threshold
Random Forest Classifier
labeled
predicted
150
200
250
300
350
400
450
500
OT (s)

figure 10.



First detected
P wave
S wave
Magnitude
5.0
6.0
7.0
8.0
Time since OT (s)
Epicentral radius (km)
0
5
10
15
20
25
30
0
10
20
30
40
50
60
70
80

figure 11.



(a)
3.6%
2.2%
0.7%
0.3%
Mean FPR (%)
0
1
2
3
4
Med+3*MAD
NCX2
STA/LTA
RF-ML
(b)
Med+3*MAD
NCX2
STA/LTA
RF-ML
False positive rate (%)
0
1
2
3
4
5
p548
p581
p626
somt
islk
p642
p645
p587
p520
p809
ecfs
p466
bill
mscg
rmvj
p574
p542
p467
cdmt
vdcy
sgps
cjms
p546
p568
p597
p600
case
p518
bomg
vlmt