Geophysical Research Letters

RESEARCH LETTER
10.1029/2018GL077870

Machine Learning Seismic Wave Discrimination: Application to Earthquake Early Warning

Zefeng Li[1], **Men-Andrin Meier**[1], **Egill Hauksson**[1], **Zhongwen Zhan**[1], **and Jennifer Andrews**[1]

[1]Seismological Laboratory, Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA

Abstract Performance of earthquake early warning systems suffers from false alerts caused by local impulsive noise from natural or anthropogenic sources. To mitigate this problem, we train a generative adversarial network (GAN) to learn the characteristics of first-arrival earthquake *P* waves, using 300,000 waveforms recorded in southern California and Japan. We apply the GAN critic as an automatic feature extractor and train a Random Forest classifier with about 700,000 earthquake and noise waveforms. We show that the discriminator can recognize 99.2% of the earthquake *P* waves and 98.4% of the noise signals. This state-of-the-art performance is expected to reduce significantly the number of false triggers from local impulsive noise. Our study demonstrates that GANs can discover a compact and effective representation of seismic waves, which has the potential for wide applications in seismology.

Plain Language Summary Earthquake early warning systems are sometimes accidentally triggered by impulsive noise signals, rather than by real earthquake signals, which leads to false alerts. This may cause unnecessary economic loss and public concern. Here we use machine learning tools to determine if the waveforms are generated by earthquakes or local noise sources. We train the algorithms with about 700,000 waveforms recorded by southern California and Japan. We demonstrate that the trained machine learning discriminator can recognize 99.2% of the earthquakes and 98.4% of the noise. This discriminator can reduce a large number of false alerts and significantly improve the robustness of early warning systems.

Key Points:
- We train machine learning algorithms with a large data set to discriminate earthquake *P* waves from local impulsive noise
- The trained discriminator achieves accuracy of 99.2% for *P* waves and 98.4% for impulsive noise
- The discriminator can significantly reduce false alerts in earthquake early warning systems

Supporting Information:
- Supporting Information S1

Correspondence to:
Z. Li,
zefengli@gps.caltech.edu

Citation:
Li, Z., Meier, M.-A., Hauksson, E., Zhan, Z., & Andrews, J. (2018). Machine learning seismic wave discrimination: Application to earthquake early warning. *Geophysical Research Letters*, *45*, 4773–4779. https://doi.org/10.1029/2018GL077870

Received 9 MAR 2018
Accepted 6 MAY 2018
Accepted article online 11 MAY 2018
Published online 29 MAY 2018

1. Introduction

Rapid growth in the quantity of seismic data has posed challenges for processing and analysis techniques in modern seismology. Many popular techniques used in major data centers and seismological research originated from the age when seismic data were small and computational power was limited. Today machine learning provides a large collection of tools to extract valuable information from voluminous data. Trained with sufficient data, machine learning algorithms have human-like capability to recognize natural objects and make expert-level decisions in various disciplines. In addition, although the computational cost for training is usually high, their online operation is low cost (e.g., Perol et al., 2018). These advantages make machine learning particularly suitable for applications in real-time seismology and earthquake early warning (EEW).

One of the major challenges in EEW is rapid and reliable detection of earthquakes in seismically noisy environments. Impulsive natural or anthropogenic noise (e.g., traffic industry activities, lightning, and device malfunction) near the seismometers may accidentally trigger the system and eventually lead to false alerts (Böse et al., 2014; Wurman et al., 2007). False alerts may halt industrial activities and transportation systems and cause unnecessary public concern. Therefore, a reliable method to discriminate earthquake waves from other impulsive noise is of particular importance for robust EEW.

A simple discrimination approach is to define a set of waveform features (peak amplitude, dominant frequency, etc.) and set up the criteria that are exclusively satisfied by an authentic *P* wave. However, due to the natural complexity of *P* waves and unpredictability of noise waveforms, finding such criteria is very difficult. Alternatively, we can process the features using a machine learning algorithm (e.g., decision tree and multilayer perceptron) and train the algorithm to find an implicit determination of signal types (e.g., Kong et al., 2016). In either method, these features are manually extracted and subjectively chosen, in the hope that they fully characterize the waveforms. In comparison, deep neural networks (e.g., convolutional neural

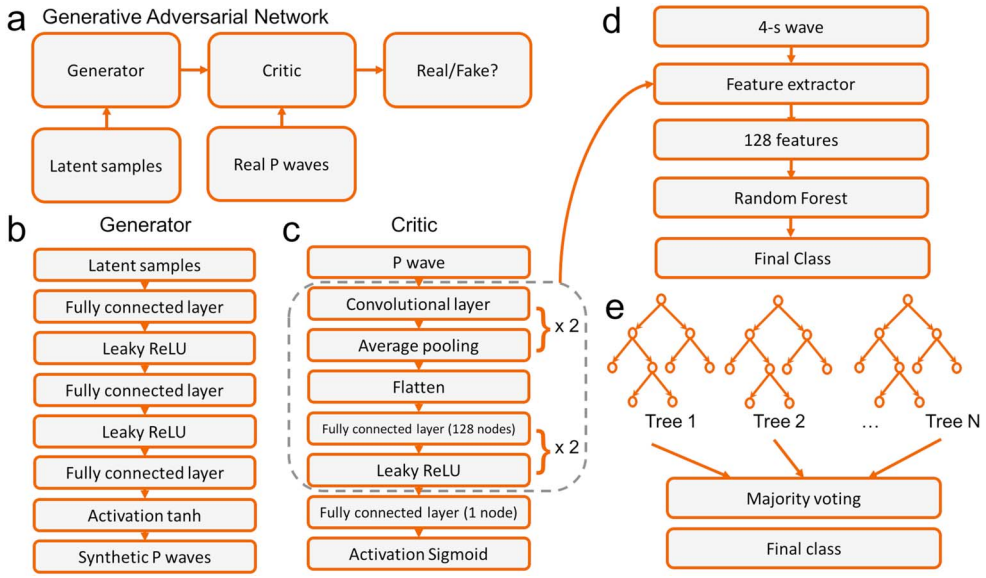

Figure 1. The architectures of generative adversarial networks. (a) A basic model of a GAN. (b) Architectures of the generator. (c) Architectures of the critic. The dashed box marks the components that are later used as a waveform feature extractor. (d) The basic workflow for seismic wave discrimination combining the critic feature extractor and Random Forest. (e) Structure of Random Forest.

networks and recurrent neural networks) can directly take waveforms as input and infer implicit but relevant features that accurately represent the waveforms (e.g., Perol et al., 2018).

In this study, we address the discrimination problem in EEW using a combination of generative adversarial networks (GANs) and Random Forests. Proposed by Goodfellow et al. (2014), GANs are unsupervised learning algorithms that consist of two neural networks, a generator and a critic, competing with each other (Figure 1a). The generator is designed to produce synthetic samples that are as realistic as possible in order to fool the critic, whereas the critic is designed to distinguish the difference between the generator output and the real data. We first train the GAN so that it can recognize and produce realistic synthetic *P* waves. Then we use the trained critic as an automatic waveform feature extractor and combine it with a Random Forest classifier to be an earthquake *P* wave discriminator. After training the networks with a large number of waveforms, we demonstrate that such a discriminator can achieve the state-of-the-art performance in distinguishing earthquake waveforms in EEW.

2. Data

We use waveforms of 342,228 local *P* waves from southern California and Japan and 373,731 noise waveforms from southern California. The earthquake data set consists of broadband and strong motion waveforms from the Southern California Seismic Network (SCSN) and strong motion waveforms from Japan (K-NET and KiK-net, surface stations only). For SCSN, we include earthquakes with M $>$ =3.0 from January 1990 to November 2016 (91,057 records). Where available, we use the onset arrival times provided by the SCSN. Otherwise we apply the Suspension Bridging Picking (SBPx) algorithm, which maximizes a ratio of integrated weighted amplitudes before and after a series of candidate picks to get a maximally precise pick (Meier et al., 2015). We use all available Japanese waveforms with $M_{JMA} > 4$ from June 1996 to October 2017 (251,171 records) and obtain picks with the SBPx algorithm. All earthquake waveforms are verified by a series of automated quality checks, such as a comparison of the measured *P* wave onset with theoretical arrival times.

We do not impose an explicit limit on epicentral distance but only include records that have signal-to-noise ratio (SNR) $>$ =5, which is defined as the signal power ratio of 2-s window before and after the pick. This effectively limits the included distance ranges as a function of earthquake size. Furthermore, we require the standard deviation of the 1.0-s waveform after the arrival to be greater than that of the 1.0-s waveform before the

arrival for both velocity and acceleration records. These selection criteria reduce the number of *P* waves to 317,500.

The noise data are waveforms that triggered the short-term average/long-term average filter of the real-time ShakeAlert EEW system (Kohler et al., 2017) from January 2015 to April 2017 across the SCSN. We remove all triggers within 2 min of any local or regional earthquakes, as well as quarry blasts in the SCEDC catalog, to avoid routinely determined earthquakes and blasts in the noise data set. For each station, we use a maximum number of 2,000 waveforms, but we make sure to include all records with nonzero quality metric *Q*. The *Q*-value is an estimate of how likely a signal corresponds to local earthquake signals, based on the peak displacement and predominant period in the initial waveform (modified from Böse et al., 2009). This ensures that all the difficult cases (i.e., noise signals that closely resemble real earthquake signals) are included in the data set. This study includes 373,731 such triggered noise signals. Owing to the automated procedure, the noise data set might mistakenly include some uncatalogued small earthquakes. Conversely, some impulsive noise picks that happen to occur around the earthquakes may be included as earthquake *P* waves. However, we visually inspect a subset of the data and find that such mislabeled cases are very rare. Besides, some of the smallest earthquakes categorized as noise would not affect EEW, because EEW is tuned for earthquakes with moderate to large magnitudes.

Before training our machine learning models, we perform simple preprocessing steps for both data sets as follows: (1) remove the prearrival mean, (2) apply a second-order causal Butterworth high-pass filter with a corner frequency of 0.075 Hz, (3) convert the waveforms into acceleration traces at a sample rate of 100 Hz, (4) cut the waveforms 1 s before and 3 s after the trigger time, and (5) normalize each waveform by its absolute maximum amplitude. We include the 1 s prearrival noise to accommodate some pick error. Because the task is to recognize the waveform shape, absolute amplitude information is not necessary. All the algorithm training and subsequent analysis are applied to the vertical acceleration waveforms.

3. Methods

3.1. Training the GAN

A GAN consists of the generative network and the critic network that compete with each other (Figure 1a). During the training, the generator makes more and more realistic synthetic waveforms, while the critic improves its ability to discriminate between the real and the generated waveforms. The critic takes input of a 4-s waveform (3-s *P* wave and 1-s prearrival noise) and outputs a continuous scalar that is indicative of the possibility for a real *P* wave (between 0 and 1, 1 for real and 0 for fake). We construct the critic as a shallow convolutional network, which consists of two convolutional layers and two fully connected layers (Figure 1c). Each convolutional layer has 16 filters, kernel size = 3, and stride = 2, followed by an average pooling layer. Each fully connected layer has 128 neurons and is followed by a leaky rectifier unit. The critic has a total of 66,769 trainable parameters. In comparison, the generator takes input of a latent sample (a 50-element vector drawn from a normal distribution) and outputs a 4-s waveform. It is composed of three fully connected layers (128, 128, and 400 neurons, respectively), each followed by a leaky rectifier unit (Figure 1b). The generator has a total of 74,640 trainable parameters, comparable to those of the critic.

The GAN is trained with only *P* waves. We split 80% of the data as training set and 20% as testing set. In the training, we need to train the critic slightly faster than the generator and maintain balance between them in the learning progress. After several experiments, we find that when the critic learns at a rate 2 times that of the generator, and the critic is trained for 5 iterations for each generator iteration, the learning process achieves balance. These parameters are inherited from other GAN examples as well as explored in this study to produce stable results (e.g., Arjovsky et al., 2017). We summarize the training settings in Table S1.

3.2. Training the Random Forest

With the trained GAN, we remove the last two layers of the critic and use it as a feature extractor (Figure 1). This is based on the assumption that the trained critic has learned to recognize the key features of earthquake *P* waves by analyzing a large number of them. Using the extracted features, we set up a Random Forest to classify *P* waves and noise waveforms (Figure 1d). A Random Forest is an ensemble of decision trees that

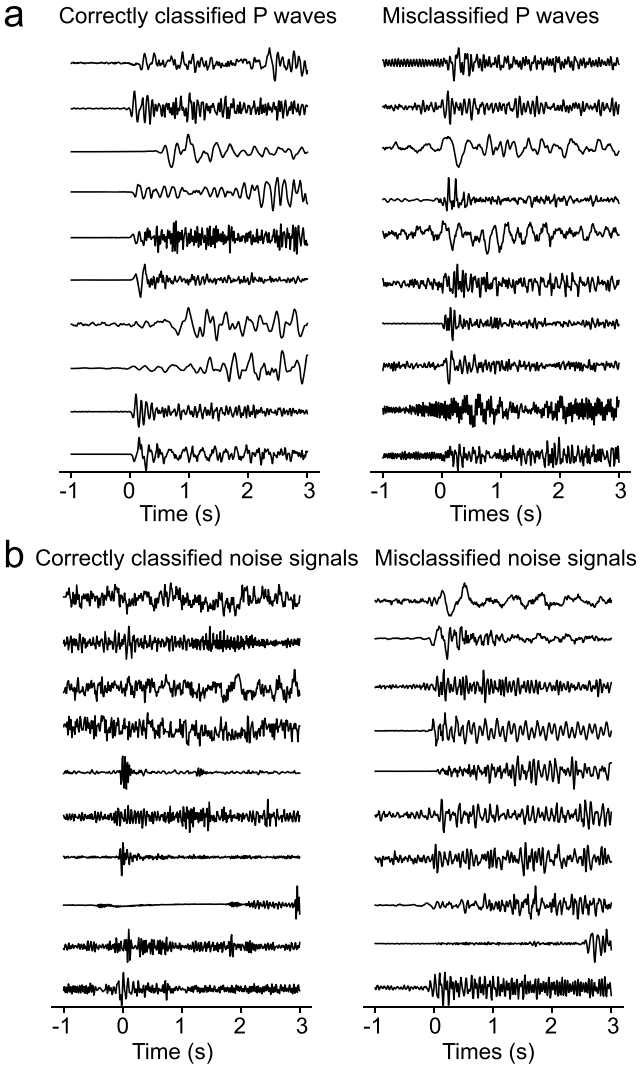

Figure 2. Examples of correctly classified and misclassified earthquake *P* waves and noise signals.

are trained via data bootstrapping and output aggregation (also known as bagging; Ho, 1995). Random Forests increase the tree diversity by introducing extra randomness, such as subsampling feature subspaces when growing the decision trees (Ho, 1998). These characteristics make Random Forests among the most powerful machine learning algorithms available today. Using the Random Forest as a classifier, rather than the GAN critic itself, leads to a much higher classification performance (see section 5).

In the Random Forest training, we include the data of 317,500 *P* waves and 373,731 triggered noise signals. We take a random subset of 250,000 *P* waves and 250,000 triggered noise signals for training, and use the remaining data for testing. The waveforms are first fed into the feature extractor and go through two convolutional layers and two fully connected layers, and are finally transformed into vectors of 128 features. The Random Forest takes inputs of 128 dimensional features and makes a decision on the signal class. To obtain high test accuracy while maintaining a reasonable training time, we search the parameters of the tree depth and the total number of decision trees (Figure S1). Other settings follow the default values in *sklearn.ensemble.RandomForestClassifier* (Pedregosa et al., 2011). Finally, we settle at 100 decision trees and the tree depth of 45 (Figure S1), and the following analysis is based on these settings. Table S2 summarizes the major parameters of the Random Forest training.

4. Results

After training, the discriminator, as a combination of the GAN critic and the Random Forest classifier, achieves 99.2% accuracy for *P* waves and 98.4% accuracy for noise signals in the test data set. In other words, we have 0.8% chance to mistake a *P* wave as noise and 1.6% chance to mistake a noise signal as a *P* wave. Figure 2 shows examples of the classified and misclassified waveforms (whose predicted labels are consistent or inconsistent with their existing labels). Note that a variety of *P* waves can be properly classified, which suggests that the discriminator can recognize a wide range of *P* wave characteristics (e.g., low/high dominant frequency, temporal amplitude decrease/increase, and various SNRs). The misclassified *P* waves generally have either relatively peculiar waveform shapes or insignificant onsets. The tested noise waveforms also vary drastically in waveform characteristics but most of them are properly classified (Figure 2). In comparison, the noise signals misclassified as *P* waves have relatively good SNR and an enduring wave train, which is similar to authentic *P* waves to some extent. Some of them are possibly from small earthquakes that are not listed in the standard catalog (Figure S3), which suggest potential applications of our discriminator in earthquake detection without using templates.

For EEW, we are particularly interested in robust performance for short-distance recordings of large-magnitude earthquakes. In Figure 3a, we analyze the dependence of prediction accuracy on magnitudes, using the stations with hypocentral distance $<$100 km. For M 3–4, we miss a fraction of 1.2% *P* waves, which is likely due to the low SNR in data recorded by a few noisy stations. For M 4–7, the accuracy reaches up to 99.7%. For M $>$ 7, we predict all the 37 test records correctly. This is a desired behavior for an EEW discriminator in which larger events need to be reliably identified to avoid missing critical alerts. The data plotted in Figure 3b confirm that the classifier performs better in the higher SNR range.

Besides the class label, the Random Forest classifier computes a continuous scalar suggestive of the probability of a *P* wave (between 0 and 1, with 1 for true and 0 for false). This enables us to examine the prediction reliability and evaluate the threshold-dependent performance of the discriminator. Figure 4a shows that

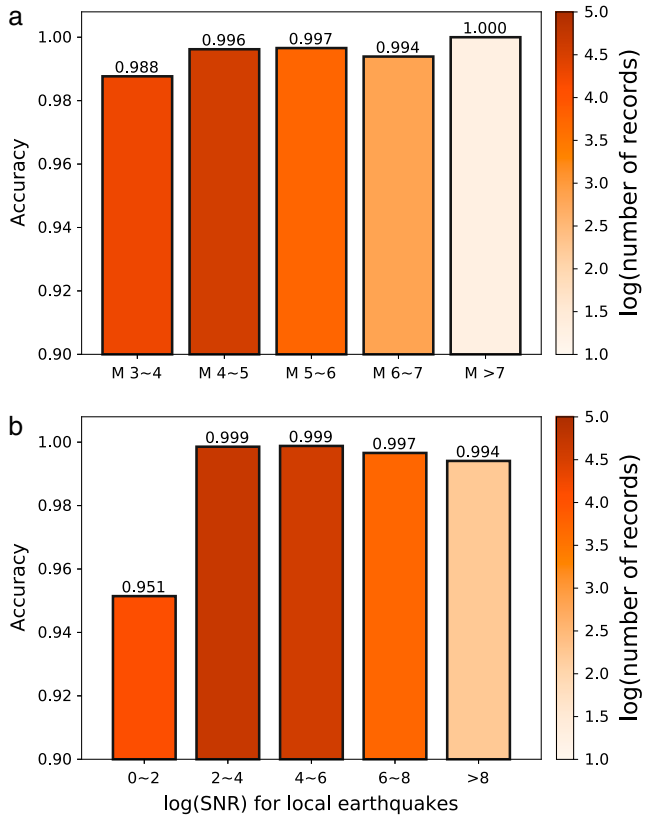

Figure 3. Dependence of the discriminator performance on earthquake magnitude and SNR. (a) Discrimination accuracy (marked on the column top) for earthquakes with different magnitudes. The numbers marked in the column are the number of test waveforms in the bins. (b) Discrimination accuracy for earthquakes with different SNRs.

92.4% of the test noise and 91.9% of earthquake data concentrate on two ends of the range of this probability proxy (0–0.1 and 0. 9–1.0). This demonstrates that the majority of the earthquake and noise data are classified with very high confidence. Conventionally, the probability of 0.5 is chosen as the decision boundary. However, in an EEW system, this threshold can be tuned according to the occurrence frequency of noise triggers at individual stations. In Figure 4b, we examine the influence of relative trigger frequencies for noise triggers and earthquake triggers on precision (ratio of true positive cases to all positive cases) and recall (ratio of true positive cases to the total of true positive and false negative cases) relationship. The results show that the precision decreases with the number of noise triggers. Therefore, the frequency of noise triggers is an important factor for threshold setting.

As a by-product, the trained GAN generator has gained the ability to produce realistic synthetic waveforms by mapping random latent samples (Creswell et al., 2018). Figure S2 shows examples of GAN synthetic waveforms and the comparison with the real earthquake waveforms. We observe that the synthetic waveforms show several essential features similar to those of real *P* waveforms, including clear *P* wave onsets, coda wave decay after the first arrival, and relatively low frequency at the beginning (Figure S2). This demonstrates that the GAN generator has actually learned the statistical characteristics from the input earthquake *P* waves. It is worth noting that the GAN synthetics are neither exact copies nor noisy versions of the real data. Owing to the GAN structure (Figure 1), the generator does not have direct access to the real data during the training process. Such synthetic waveforms could be useful in waveform data augmentation in deep learning applications in seismology.

5. Discussions

We trained a GAN with more than 300,000 earthquake *P* waves recorded in southern California and Japan. Using the GAN critic as a feature extractor, we trained a Random Forest as a *P* wave discriminator, with the goal to reduce the number of false triggers in EEW. The test results show that the discriminator can identify a *P* wave with 99.2% accuracy and eliminate a false trigger with 98.4% accuracy. Our machine learning method is designed to learn human-like capabilities of recognizing what an earthquake waveform looks like, rather than to act as a phase picker. The discriminator can be triggered by, for example, a short-term average/long-term average phase detector, so that any potential earthquake signal can be classified as real, or discarded as noise. By installing this discriminator after a phase picker in EEW systems, we expect to reduce large numbers of potential false alerts and thus increase the EEW robustness to local noise.

A key motivation behind GANs is to find a representation, either implicit or explicit, of the real-world data (Creswell et al., 2018). Specifically, in our study, we use it to capture the statistical distribution of *P* waves. The critic defines some form of similarity metric, and the training narrows down the gap between the generator outputs and the real *P* waves. Therefore, a well-trained critic can grasp the key features of the real data and thus can be used as a feature extractor. As compared to manually defining and subjectively choosing a set of waveform features (e.g., Kong et al., 2016; Rouet-Leduc et al., 2017), the critic automatically extracts the features that are highly representative of the study target. However, automatic features have a drawback of difficulty in interpretation, whereas a model using sets of standard statistical features can be interpreted relatively easily. In EEW, we aim to discriminate *P* waves out of all the triggered signals. Hence, the GAN extracted features are a natural fit for such a one-against-others classification task.

Radford et al. (2016) have shown that the image features extracted from a trained critic are useful for classification even with a simple linear model. Inspired by Radford et al. (2016) and Salimans et al. (2017), we use

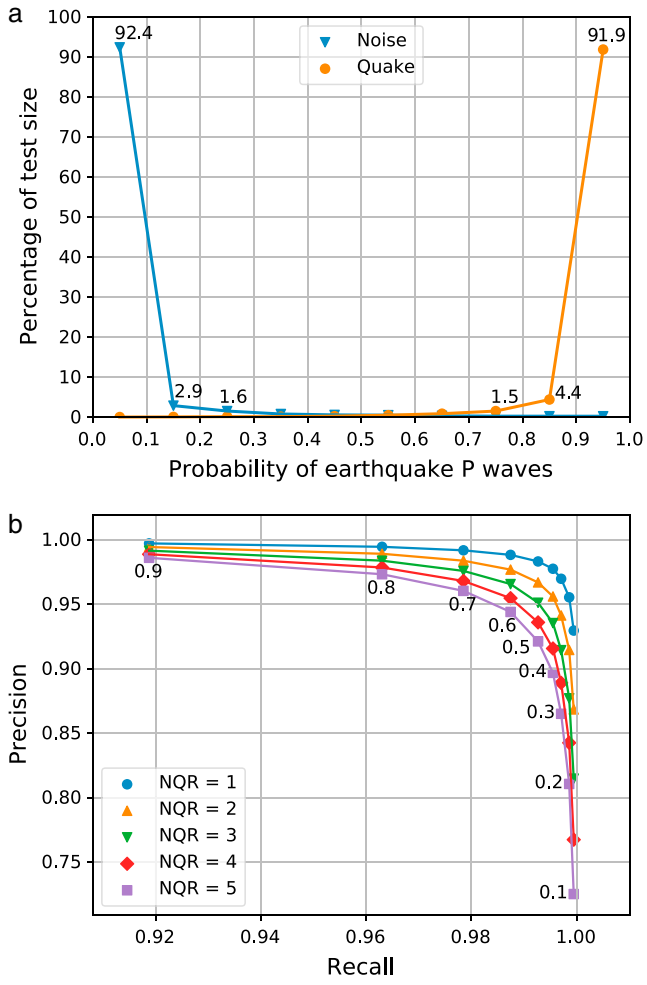

Figure 4. Statistical analysis of the discriminator performance. (a) Normalized histogram of discriminator output probability of being a real *P* waves for quake and noise test data. Note that the quake data concentrate on the 0.9–1.0 and the noise data concentrate on the 0.0–0.1. (b) Precision-recall changes with relative numbers of noise triggers and quake triggers, assuming the constant noise and quake accuracy as in the test. NQR represents the ratio of the number of noise triggers to the number of quake triggers. The numbers marked near the purple dots are the probability threshold for *P* waves.

the Random Forest for classification with the features extracted by the critic, instead of directly using the trained GAN critic as a discriminator. In fact, if the trained critic is directly used as a quake-to-noise discriminator, the accuracy for *P* waves and for noise signals is 81% and 36%, respectively, which is far from the performance of our combined discriminator. This is because the critic itself is trained to distinguish the real data and the generator synthetics, but has no access to the noise data. In comparison, trained with the earthquake and noise data, the Random Forest learns the feature ranges of the noise data and the earthquake data so that it can classify them with much higher accuracy.

As in any classification problem, there generally exists a trade-off between false positives (false alerts) and false negatives (missed alerts), which can be modified by the trigger criterion (Figure 4). Apart from the relative numbers of noise triggers, two other important factors can affect the discrimination threshold setting: robustness of the initial phase picker and the end-users' tolerance of false alerts. A robust phase picker can remove most of the noise triggers at the first stage, thus reducing the workload of the discriminator. Tolerance on false alerts varies among different users. For example, industry users with high costs of false alerts may be less tolerant of false alerts and may prefer a high alerting threshold. Personal users who receive little loss due to false alerts might be more tolerant, and therefore, a relatively low threshold could be used, thus minimizing the chance of missed alerts.

Because most current EEW systems use multiple stations to trigger an alert, it is of interest to evaluate the performance of our discriminator at a network level. Here we consider a simple scenario: an EEW system that requires two stations to detect an earthquake to issue a warning. For a true earthquake, the possibility of a missed alert is 1–0.99^2 = 0.02. For a noise trigger, the possibility of a false alert is 0.02 * 0.02 = 0.0004. Compared to the performance of individual stations, the possibility of false alerts decreases from 0.02 to 0.0004, whereas the possibility of missed alerts increases from 0.01 to 0.02. The actual performance using more sophisticated warning criteria would be different from this simple calculation. However, generally, with a stricter criterion, false alerts are reduced significantly at the expense of slight increased chance of missed alerts.

It is clear that the setup of the neural network architectures (both the critic and the generator) is not unique. Besides the network presented here (Figure 1), we have explored a range of popular structures in image processing, such as DCGAN (Radford et al., 2016) and Wasserstein GAN (Arjovsky et al., 2017). They have not performed well in our experiments. It could be either because the architecture for image processing may not be suitable for waveform processing or because parameters have not been optimally tuned in our experiments. Further exploration of better architectures and training parameters may improve their performance.

Our preferred approach is applicable to many other discrimination and/or detection tasks related to seismic waveforms. The GAN can easily identify the waveform features that are most relevant to the target, which does not require subjective choices of feature sets. The Random Forest incorporates additional power by its proven robustness in classification of different objects. Because these tools are very new to seismology, the applications of their powerful capabilities are still under exploration. For example, it is possible to train successfully a GAN to generate synthetic waveforms that approximate wave-equation simulated ones, which significantly reduces the simulation time (Krishcher & Fichtner, 2017). Based on other applications in computer vision such as image in-painting and image superresolution, we expect that the GAN can have much broader seismological applications that are beyond the scope of this study.

6. Conclusion

We have constructed an earthquake *P* wave discriminator to address the noise trigger problem in EEW, by making use of two powerful machine learning algorithms, GANs and Random Forests. The GAN is trained to learn the primary features of early *P* waves, which saves us from having manually to define and select the waveform features. Random Forests take advantage of these features and classify the waveform types with high accuracy. Both algorithms were trained with a large amount of seismic data, that is, 300,000 *P* waves and more than 350,000 noise waveforms recorded in southern California and Japan. Combining the GAN critic and the Random Forest, we achieved the start-of-the-art performance in discriminating earthquakes against other impulsive noise triggers, which can significantly reduce false triggers in EEW. Our study makes a compelling case that GANs have capability to discover compact representation of seismic waves, which has potential for wide applications in seismology.

Acknowledgments
We thank Qingkai Kong and another anonymous reviewer for constructive comments. This research was supported by a Gordon and Betty Moore Foundation grant to Caltech and by the Swiss National Science Foundation. The Japanese waveform data can be downloaded from http://www.kik.bosai.go.jp/ (last accessed October 2017). For southern California we have used waveforms and parametric data from the Caltech/USGS Southern California Seismic Network (doi:10.7914/SN/CI) stored at the Southern California Earthquake Data Center (doi:10.7909/C3WD3xH1). The algorithms were written with Python packages Keras (https://keras.io/) and Scikit-learn (http://scikit-learn.org/). The GAN was trained for 2.0 hr, and the Random Forest was trained for 20 min on a PC (NVIDIA GeForce GTX 1050 Ti 4 GB, Intel Core i5-7300HQ 2.50GHz).

References

Arjovsky, M., Chintala, S., & Bottou, L. (2017). Wasserstein generative adversarial networks. *Proceedings of the 34th International Conference on Machine Learning, PMLR*, *70*, 214–223.

Böse, M., Allen, R., Brown, H., Gua, G., Fischer, M., Hauksson, E., et al. (2014). CISN ShakeAlert: An earthquake early warning demonstration system for California. In *Early Warning for Geological Disasters* (pp. 49–69). Berlin, Heidelberg: Springer.

Böse, M., Hauksson, E., Solanki, K., Kanamori, H., Wu, Y. M., & Heaton, T. H. (2009). A new trigger criterion for improved real-time performance of onsite earthquake early warning in southern California. *Bulletin of the Seismological Society of America*, *99*(2A), 897–905. https://doi.org/10.1785/0120080034

Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. A. (2018). Generative adversarial networks: An overview. *IEEE Signal Processing Magazine*, *35*(1), 53–65. https://doi.org/10.1109/MSP.2017.2765202

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2014). Generative adversarial nets. In *Proceedings of Neural Information Processing Systems* (pp. 2672–2680).

Ho, T. K. (1995). Random Decision Forests. *Proceedings of the 3rd International Conference on Document Analysis and Recognition* (pp. 278–282). Montreal, QC, 14–16 August 1995.

Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(8), 832–844. https://doi.org/10.1109/34.709601

Kohler, M. D., Cochran, E. S., Given, D., Guiwits, S., Neuhauser, D., Henson, I., et al. (2017). Earthquake early warning ShakeAlert system: West coast wide Production Prototype. *Seismological Research Letters*, *89*(1), 99–107.

Kong, Q., Allen, R. M., Schreier, L., & Kwon, Y.-W. (2016). Myshake: A smartphone seismic network for earthquake early warning and beyond. *Science Advances*, *2*(2), e1501055. https://doi.org/10.1126/sciadv.1501055

Krishcher, L., & Fichtner, A. (2017). *Generating seismograms with deep neural networks, abstract S41D-03*. Presented at 2017 fall meeting, AGU, New Orleans, LA, 11–15 December.

Meier, M.-A., Heaton, T., & Clinton, J. (2015). The Gutenberg algorithm: Evolutionary bayesian magnitude estimates for earthquake early warning with a filter bank. *Bulletin of the Seismological Society of America*, *105*(5), 2774–2786. https://doi.org/10.1785/0120150098

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Perol, T., Gharbi, M., & Denolle, M. (2018). Convolutional neural network for earthquake detection and location. *Science Advances*, *4*(2), e1700578. https://doi.org/10.1126/sciadv.1700578

Radford, A., Metz, L., & Chintala, S. (2016). Unsupervised representation learning with deep convolutional generative adversarial networks. In *Proceedings of the 5th International Conference on Learning Representations (ICLR)* - Workshop Track, 2016.

Rouet-Leduc, B., Hulbert, C., Lubbers, N., Barros, K., Humphreys, C. J., & Johnson, P. A. (2017). Machine learning predicts laboratory earthquakes. *Geophysical Research Letters*, *44*, 9276–9282. https://doi.org/10.1002/2017GL074677

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., & Chen, X. (2017). Improved techniques for training GANs. In *Proceedings of Neural Information Processing Systems* (pp. 2234–2242).

Wurman, G., Allen, R. M., & Lombard, P. (2007). Toward earthquake early warning in northern California. *Journal of Geophysical Research*, *112*, B08311. https://doi.org/10.1029/2006JB004830