

Prediction of surface wave velocities with historical seismic data

Nikhil Mukund

*Inter-University Centre for Astronomy and Astrophysics (IUCAA),
Post Bag 4, Ganeshkhind, Pune 411 007, India*

Michael Coughlin

*Division of Physics, Math, and Astronomy,
California Institute of Technology, Pasadena, CA 91125, USA*

Jan Harms

*INFN, Sezione di Firenze, Sesto Fiorentino, 50019, Italy
Università degli Studi di Urbino “Carlo Bo”, I-61029 Urbino, Italy*

Nicolas Arnaud

*LAL, Univ. Paris-Sud, CNRS/IN2P3,
Université Paris-Saclay, F-91898 Orsay, France and
European Gravitational Observatory (EGO), I-56021 Cascina, Pisa, Italy*

David Barker, Hugh Radkins, and Jim Warner

LIGO Hanford Observatory, Richland, WA 99352, USA

Sebastien Biscans, Fred Donovan, and Richard Mittleman

LIGO Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02138, USA

Eric Coughlin

*Department of Computer Science, Luther College,
700 College Dr, Decorah, IA 52101, USA*

Paul Earle

U.S. Geological Survey, Golden, CO 80401, USA

Jeremy Fee and Michelle Guy

United States Geological Survey, Golden, CO 80401, USA

Irene Fiori and Bas Swinkels

European Gravitational Observatory (EGO), I-56021 Cascina, Pisa, Italy

Hunter Gabbard

*Albert-Einstein-Institut, Max-Planck-Institut für
Gravitationsphysik, D-30167 Hannover, Germany*

Brian Lantz

Stanford University, Stanford, CA, USA

Arnaud Pele and Keith Thorne

LIGO Livingston Observatory, Livingston, LA 70754, USA

Earthquake early warning (EEW) is a burgeoning field dedicated to the rapid detection and characterization of earthquakes as well as the dissemination of that information to people and infrastructure in their path [1–10]. As these systems minimize the time required to calculate the source parameters of earthquakes (i.e. their location and magnitude), it becomes important to predict the ground motion that the earthquakes will cause as a function of location and distance with high accuracy. In this analysis, we leverage the power of machine learning algorithms to make both ground velocity predictions and gravitational-wave detector lockloss predictions. We demonstrate an improvement from a factor of 5 to a factor of 2.5 in scatter of the error in the predicted ground velocity over a previous model fitting based approach. To assess the accuracy and utility of our approach, we compare the estimates based only on a rapid magnitude and location estimates to the amplitudes observed. We find agreement within a factor of 2 by this metric. Further, we compare measurements that include the less timely earthquake slip inversion and CMT information to the original amplitudes observed, resulting in a factor of 2 agreement.

With the advent of gravitational-wave astronomy, it is essential to maximize the duty cycle of second-generation gravitational-wave detectors such as the Laser Interferometer Gravitational-wave Observatory (LIGO) [11], Virgo [12], and GEO600 [13] detectors. Any increase in duty cycle increases the sensitivity of gravitational-wave searches, including the observations of binary black hole mergers [14–17] and binary neutron stars [18]. One source of ground motion that destabilizes the detectors are earthquakes [19, 20], despite seismic isolation systems designed to minimize such effects [21–23]. The surface waves, the highest amplitude component from earthquakes with the longest duration, adversely affect the detectors. This occurs by making it impossible to keep the detectors operable or induce higher frequency noise by upconverting low-frequency optical motion.

Many seismic and geodetic (GPS) sensor arrays exist that are producing rapid earthquake information products, from magnitude and location estimates to regional centroid moment tensors (CMTs) and advanced slip inversions. With wide-ranging public warning systems in Mexico and Japan and smaller-scale systems in many other countries, warnings from seconds to minutes are now available to reduce the impact of earthquakes on society [10]. The short warning times arise out of the physical processes that drive the earthquake rupture,

where the warning is given by seismometers measuring P-waves (≈ 8 km/s) and S-waves (≈ 4 km/s). Reliability of these estimates are one of the most important aspects of EEW systems, and their improvements generally rely on increasing the number of stations involved in the warning decisions as well as increasing alarm thresholds on ground motion, both seeking to limit the number of false positives [24]. Both of these strategies come at the cost of decreasing the warning time.

The main goal of EEW methods is to generate reliable relations (sometimes called source-scaling laws) between earthquake source parameters and ground motion metrics. Examples in the time domain include peak ground acceleration, peak ground velocity, and peak ground displacement, while in the frequency domain there are spectral accelerations, velocities, and displacements as well as predominant periods [25]. These source-scaling laws are applied to early portions of seismograms to make predictions about the magnitude for EEW [26], important for hypocenter and magnitude computations in tsunamis [27], hazard computations in engineering seismology [28], and computation of the elastic response spectrum [29].

Early estimates of magnitudes tend to underestimate the energy released due to the non-instantaneous pattern of slip. For this reason, the early estimates of the ground velocity amplitudes are often not as accurate as later values. The effects of these errors are particularly pronounced for larger earthquakes, where the estimates of the fault lengths become more important. Thus, these larger earthquakes tend to have their amplitudes underpredicted. The loss of performance that results from use of the rapid estimates is acceptable to use as rapid warnings. [27] showed that real-time GPS waveforms can rapidly determine the magnitude within the first minute of rupture and in many cases before rupture is complete.

Machine learning has recently become an important aspect of EEW and seismology in general. The *MyShake* EEW system uses artificial neural networks to differentiate earthquake and human motions, with 98% of earthquake records within 10 km correctly identified, and only 7% of people-induced transients appearing to be earthquakes to the algorithm [30]. Machine learning algorithms are also used to differentiate earthquakes from other seismic transients [31–33]. In addition, they have been used to discriminate between deep and shallow microearthquakes [34]. It can also be used to add to undersampled or missing traces [35]. In addition, they have been used to make full-wave tomography images [36].

One of the key aspects of the system is the ground velocity predictions, Rf_{amp} , for each site.

These predictions have two purposes. First of all, they provide a meaningful metric which on-site-staff at the detectors can use to plan the response to the incoming earthquake. The response could be in the form of switching seismic isolation loops to steer the interferometer to a more robust configuration keeping it locked although with a lesser sensitivity [37]. The predictions also serve as inputs to the algorithms which make lockloss predictions, which we will describe in the following. Any such information about upcoming downtime can be utilized to perform opportunistic maintenance to rectify problems typically scheduled for weekly maintenance periods.

In previous work, we used advances in early earthquake warning to develop a low-latency earthquake early warning client named *Seismon* [20]. This system uses a real-time event messaging system of the U.S. Geological Survey (USGS) to mitigate the effects of teleseismic events on ground-based gravitational-wave detectors. Using information about the earthquake source characteristics such as time, location, depth, and magnitude, predictions as to the arrival time and ground velocity induced by the earthquakes were predicted. In the initial version of the algorithm, we used an empirical fit to an equation derived to account for physical effects. This equation succeeded in predicting peak ground velocity such that 90% of events had a measured ground velocity within a factor of 5 of the predicted value. There were a few downsides to this empirical fit. First of all, while it was derived with physical effects in mind, it was predominantly an empirical construction. It was also found that the parameters in the model were quite degenerate, which meant that parameters derived to be physically meaningful quantities showed significant differences from site to site which were unlikely to actually be very different. Finally, to be useful to the detectors, there is a goal of a factor of 2 in relative error in the ground velocity predictions in order to make accurate assessments of whether the gravitational-wave detectors will be affected, which is much smaller than the factor of 5 scatter seen.

The idea of this analysis is to compare historical ground velocity measurements to predictions made using different machine learning algorithm techniques. The inputs to the algorithm are the earthquake magnitude, latitude, longitude, distance, depth, and earthquake azimuth relative to the detector. The target output is the measured ground velocity. This improves on the equation in a few ways. First of all, the algorithm leverages the power of machine learning algorithms, which is not reliant on a functional form. Second, it trivially includes more parameters, such as latitude, longitude, and earthquake azimuth relative to

the detector above and beyond the initial analytical formalism.

In particular, we compare the efficiency as the two different machine learning approaches: regression and clustering. Within regression we check the performance of tensorflow implementation of deep neural networks (DNN) [38], stacked ensemble regressor [39, 40] and Gaussian Process Regression (GPR) [41] while in clustering we use Mahalanobis distance [42] based similarity search to make the predictions. We expand on the specific implementations in the Supplementary Materials. We use historical earthquakes to train our machine learning algorithms. The parameters that enter the predictions are M , the magnitude of the earthquake, h , the depth, r , the distance to the detectors, θ and ϕ , the latitude and longitude, and α , the earthquake azimuth relative to the detector. All of these variables are available in low latency from the USGS. On longer timescales, the earthquake slip inversion, strike, rake, and dip, and the moment tensor values, M_{rt} , M_{tp} , M_{rp} , M_{tt} , M_{rr} , and M_{pp} are also available, serving as additional parameter to the machine learning algorithms. The target variables correspond to peak ground velocities measured using seismometers. Time-series are chosen to encompass the P-wave arrival to surface waves calculated assuming a ground velocity of 2 km/s. We take the vertical component of broadband (velocity) data that is filtered using an acausal 0.1 Hz low-pass Butterworth filter. The data is calibrated into ground velocity using a constant V to m/s value appropriate for each seismometer. We take the peak ground velocities measured for each earthquake and use a machine learning algorithm to make predictions of that velocity.

To improve the learning and prevent early stopping, we augment the training data by artificially adding noise (or jitter) to the predictor and response variables in a controlled fashion. The presence of noise enhances the ability of the MLA to better learn and generalize to the underlying smooth, non-linear function. New samples are generated from each of the original datasets by creating a Gaussian jitter distribution centered around the parameter value followed by random draw of samples from these distributions. Selective boosting is done to minimize the imbalances in the dataset using **S**ynthetic **M**inority **O**versampling **T**Echnique (SMOTE) [43].

Figure 1 shows the prediction results from applying Mahalanobis based clustering on the simulated and real earthquake data. We obtain an accuracy above 90% in both the cases. This performance improvement in prediction accuracy from a factor of five to 2.5 can be attributed to increased availability of data, inclusion of more earthquake parameters and

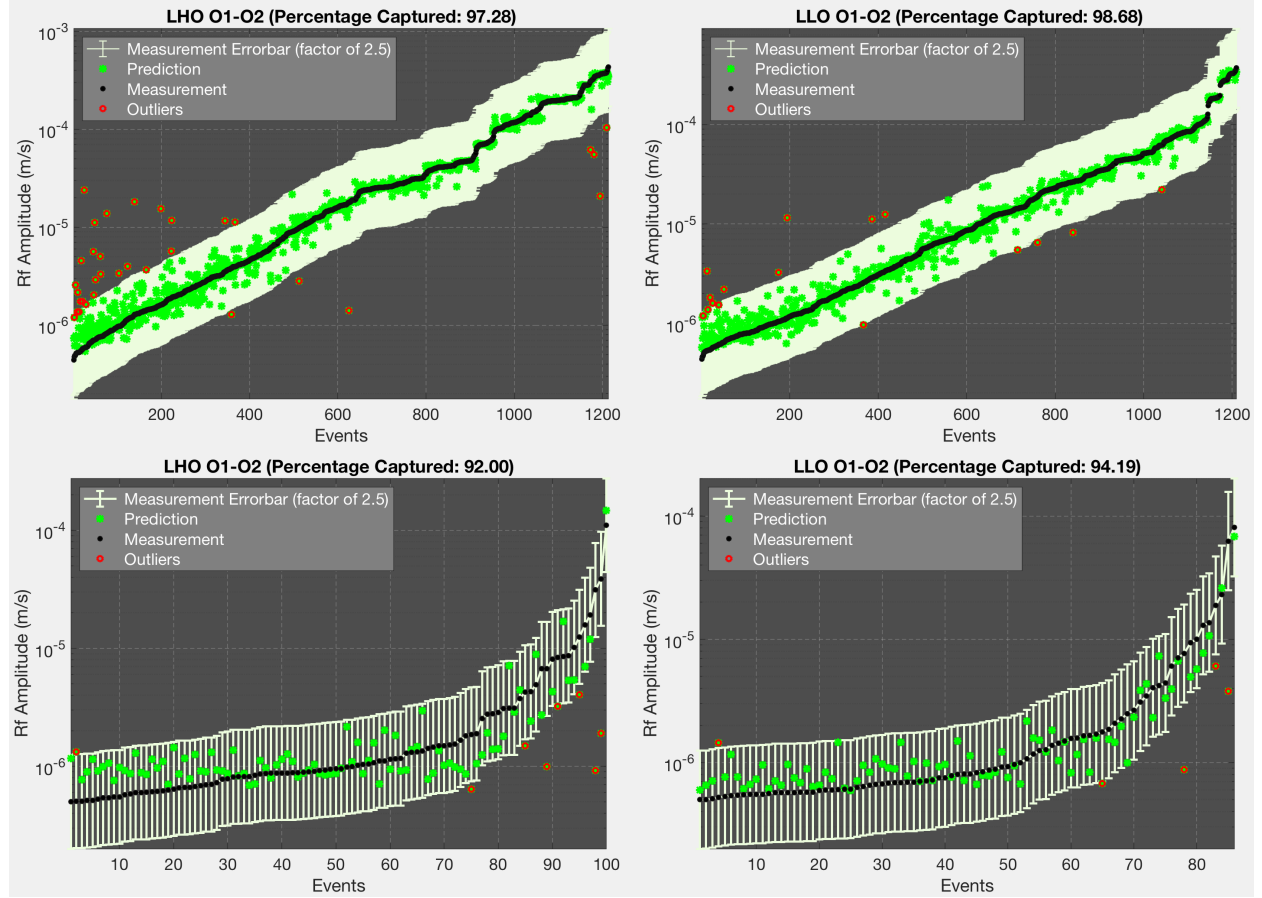


FIG. 1: Fit of peak velocities seen during O1-O2 at the interferometers (LHO, LLO) using Mahalanobis distance based clustering. Results on simulated and real data are respectively shown in the top and bottom rows. The events have been ordered by their measured peak ground velocity (in grey) and yellow error bar corresponds to a factor of 2.5 within the predicted value. More than 90% of events are within a factor of 2 of the predicted value.

the usage of robust algorithms.

We also demonstrate the resourcefulness of the above scheme by making predictions across the United States using the data recorded by the IRIS transportable array network. For ground motion recorded in each state, we perform the similarity search using archival data from the same state as well as the other states and compare the prediction accuracy as shown in Figure 2. For at-least 24 states the accuracy is seen to be above 90%. The variation seen in predictability along the diagonal might be due to the differences in local geology across the US. This observation of unpredictability could be beneficial for future site selection surveys looking for suitable locations for next generation interferometers. The fact

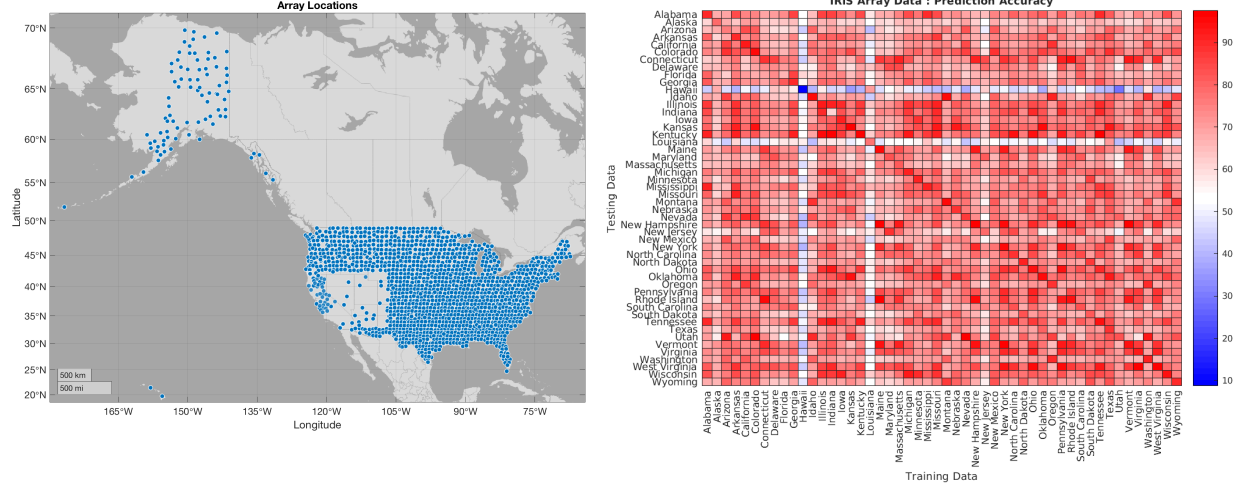


FIG. 2: Figure to the left shows the locations of seismometer array used to collect the earthquake data. Heatmap on the right shows the ground motion prediction accuracy within each state making use of archival data from the same as well as every other state.

that nearby states observe similar levels of ground motion for a given earthquake is evident from the corresponding off-diagonal terms.

The performance shows that the archival data based prediction scheme can be extended beyond the individual gravitational-wave detector sites to issue state specific early warning alerts.

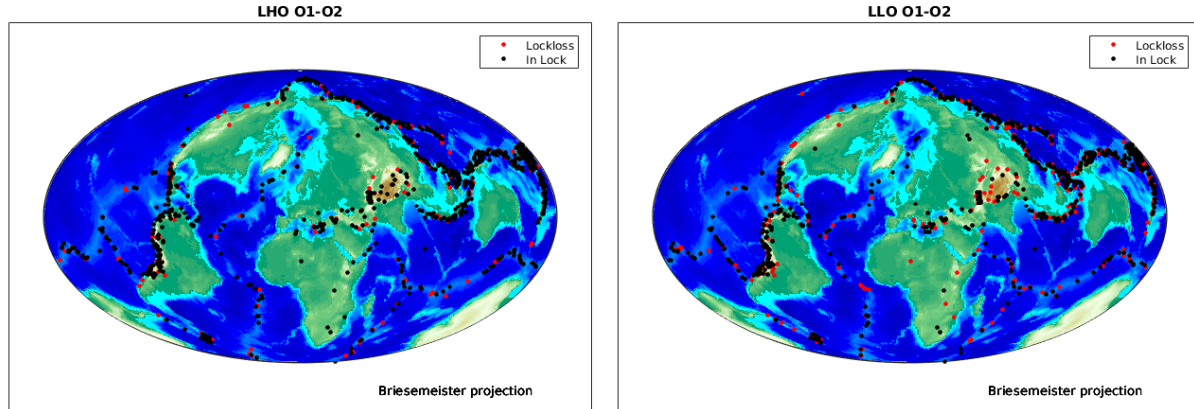


FIG. 3: Impact of earthquakes happening worldwide on GW inteferometers. Points marked in red indicate the instances when the resulting ground motion caused the interferometer to go out of lock.

The main benefit of ground velocity predictions for gravitational-wave detectors is to

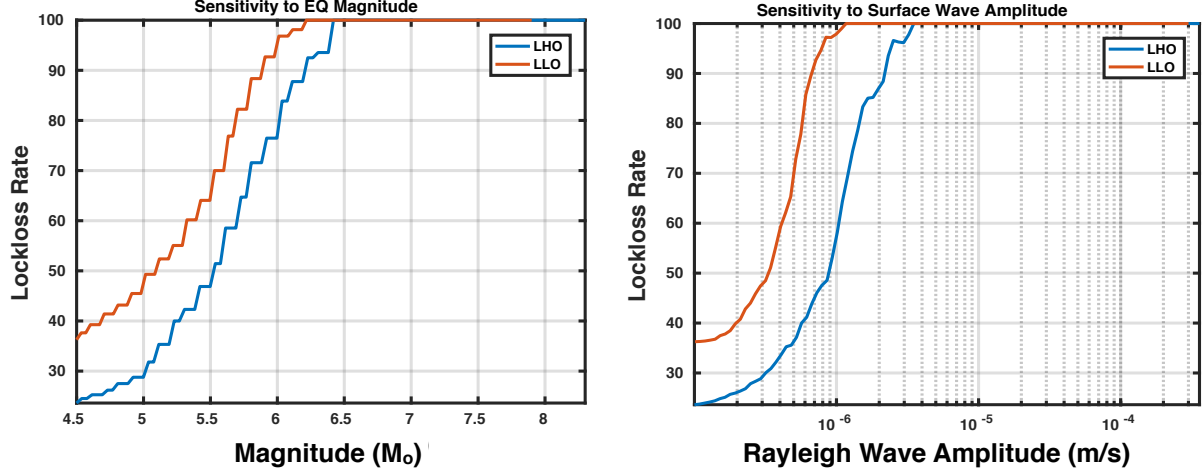


FIG. 4: Plot shows the lockloss rate associated with earthquake magnitude and Rayleigh wave amplitude at both the LIGO detectors.

inform predictions of whether an earthquake will cause the loss of data for the detector. Figure 4 shows the sensitivity of the interferometers to earthquake magnitude and Rayleigh wave amplitude. We have previously developed techniques for preventing earthquakes from causing the loss of data taking if advanced notice is given [37]. In the following, we will use a MLA to develop a lockloss prediction model as well. We use the same set of inputs to the algorithm as in the ground velocity prediction case, but also include the ground velocity predictions themselves as inputs. To generate the target variable, we take times when the gravitational-wave detectors lost the ability to take data during an earthquake and assign a value of 1, and a 0 otherwise. In figure 3, we showed world maps of earthquakes that both caused and did not cause the lockloss. We use the same clustering based algorithm as applied for the ground velocity predictions and allows for outlier detection. Acknowledging that there is a trade-off between false-alarm probability and efficiency standard, we are able to make predictions for the inliers with an accuracy above 92%, keeping the associated false-alarm probability to be less than 10%.

In conclusion, we have used MLAs to predict peak ground velocities from teleseismic earthquakes. We use the peak ground velocity predictions to predict the potential effect of earthquakes on gravitational-wave detectors and issue near realtime alerts at the site 5. Given the significant interest in accurate ground velocity predictions for EEW systems in general, we believe the techniques here are beneficial beyond the gravitational-wave commu-

nity. While we focus on the prediction of peak ground velocity here, another possibility is to use other seismometers to predict the time-series themselves. This would provide a way to directly measure transfer functions between ground motion very near the earthquake source and those in areas of significant seismological hazard, such as in the Los Angeles basin.

LHO	In Lock	Lockloss	LLO	In Lock	Lockloss
True Positives	39	10	True Positives	49	16
False Positives	1	3	False Positives	3	2
True Negatives	3	1	True Negatives	2	3
False Negatives	10	39	False Negatives	16	49
Precision	0.97	0.77	Precision	0.94	0.89
Sensitivity	0.93	0.91	Sensitivity	0.96	0.84
Specificity	0.91	0.93	Specificity	0.84	0.96

TABLE I: Performance analysis of lockloss prediction models for LHO and LLO. Each of them respectively has an accuracy of 92% and 93%.

Code availability. The code to reproduce the analysis is open-source and available at <https://github.com/ligovirgo/seismon/> for public download.

Acknowledgments. MC was supported by the David and Ellen Lee Postdoctoral Fellowship at the California Institute of Technology. NM acknowledges Council for Scientific and Industrial Research (CSIR), India, for providing financial support as Senior Research Fellow. LIGO was constructed by the California Institute of Technology and Massachusetts Institute of Technology with funding from the National Science Foundation and operates under cooperative agreement PHY-0757058. This paper has been assigned LIGO document number LIGO-?.

Global Seismographic Network (GSN) is a cooperative scientific facility operated jointly by the Incorporated Research Institutions for Seismology (IRIS), the United States Geological Survey (USGS), and the National Science Foundation (NSF), under Cooperative Agreement EAR-1261681. The facilities of IRIS Data Services, and specifically the IRIS Data Management Center were used for access to waveforms, related metadata, and derived products used in this study. IRIS Data Services are funded through the Seismological Facilities for the Advancement of Geoscience and EarthScope (SAGE) Proposal of the National

- [1] RM. Allen. Transforming earthquake detection? *Science*, 335:297–298, 2012.
- [2] Kuyuk, H.S. and RM. Allen. A global approach to provide magnitude estimates for earthquake early warning. *Geophysical Research Letters*, 40, 2013.
- [3] Kuyuk, H.S. and RM. Allen. Optimal seismic network density for earthquake early warning: A case study from California. *Seismological Research Letters*, 84(6):946–954, 2013.
- [4] Kuyuk, H.S. et al. Designing a network-based earthquake early warning system for California: ElarmS-2. *Bulletin of Seismological Society of America*, 104(1), 2014.
- [5] Cochran E., Lawrence J., Christensen C., and Chung A. A novel strong-motion seismic network for community participation in earthquake monitoring. *IEEE Inst and Meas*, 12(6):8–15, 2009.
- [6] Cochran E., Lawrence J., Christensen C., and Jakka R. The quake-catcher network: Citizen science expanding seismic horizons. *Seismological Research Letters*, 80:26–30, 2009.
- [7] M. Bose, R. Allen, H. Brown, G. Gua, M. Fischer, E. Hauksson, T. Heaten, M. Hellweg, M. Liukis, D. Neuhauser, P. Maechling, K. Solanki, M. Vinci, I. Henson, O. Khainovski, S. Kuyuk, M. Carpio, M.-A. Meier, and T. Jordan. CISN ShakeAlert: An Earthquake Early Warning Demonstration System for California. In Friedemann Wenzel and Jochen Zschau, editors, *Early Warning for Geological Disasters*, Advanced Technologies in Earth Sciences, pages 49–69. Springer Berlin Heidelberg, 2014.
- [8] Mitsuyuki Hoshiba, Osamu Kamigaichi, Makoto Saito, Shin’ya Tsukada, and Nobuo Hamada. Earthquake early warning starts nationwide in japan. *Eos, Transactions American Geophysical Union*, 89(8):73–74, 2008.
- [9] Mitsuyuki Hoshiba, Kazuhiro Iwakiri, Naoki Hayashimoto, Toshihiro Shimoyama, Kazuyuki Hirano, Yasuyuki Yamada, Yuzo Ishigaki, and Haruyuki Kikuta. Outline of the 2011 off the Pacific coast of Tohoku Earthquake (Mw 9.0) —Earthquake Early Warning and observed seismic intensity—. *Earth, Planets and Space*, 63(7):7, 2011.
- [10] Jennifer A. Strauss and Richard M. Allen. Benefits and Costs of Earthquake Early Warning. *Seismological Research Letters*, 87(3):765–772, 2016.
- [11] J Aasi et al. Advanced ligo. *Classical and Quantum Gravity*, 32(7):074001, 2015.
- [12] F Acernese et al. Advanced virgo: a second-generation interferometric gravitational wave

- detector. *Classical and Quantum Gravity*, 32(2):024001, 2015.
- [13] Grote H. for the LIGO Scientific Collaboration. The GEO 600 status. *Class. Quantum Grav.*, 27:084003, 2010.
 - [14] Abbott, B. P. et al. Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett.*, 116:061102, Feb 2016.
 - [15] Abbott, B. P. et al. Gw151226: Observation of gravitational waves from a 22-solar-mass binary black hole coalescence. *Phys. Rev. Lett.*, 116:241103, Jun 2016.
 - [16] Abbott, B. P. et al. Gw170104: Observation of a 50-solar-mass binary black hole coalescence at redshift 0.2. *Phys. Rev. Lett.*, 118:221101, Jun 2017.
 - [17] Abbott, B. P. et al. Gw170814: A three-detector observation of gravitational waves from a binary black hole coalescence. *Phys. Rev. Lett.*, 119:141101, Oct 2017.
 - [18] Abbott, B. P. et al. Gw170817: Observation of gravitational waves from a binary neutron star inspiral. *Phys. Rev. Lett.*, 119:161101, Oct 2017.
 - [19] Michael Coughlin, Christopher Stubbs, Sergio Barrientos, Chuck Claver, Jan Harms, R. Chris Smith, and Michael Warner. Real-time earthquake warning for astronomical observatories. *Experimental Astronomy*, 39(2):387–404, 2015.
 - [20] Michael Coughlin, Paul Earle, Jan Harms, Sebastien Biscans, Christopher Buchanan, Eric Coughlin, Fred Donovan, Jeremy Fee, Hunter Gabbard, Michelle Guy, Nikhil Mukund, and Matthew Perry. Limiting the effects of earthquakes on gravitational-wave interferometers. *Classical and Quantum Gravity*, 34(4):044004, 2017.
 - [21] R Abbott, R Adhikari, et al. Seismic isolation for advanced ligo. *Classical and Quantum Gravity*, 19(7):1591, 2002.
 - [22] Alberto Stochino, Benjamin Abbot, Yoichi Aso, Mark Barton, Alessandro Bertolini, Valerio Boschi, Dennis Coyne, Riccardo DeSalvo, Carlo Galli, Yumei Huang, Alex Ivanov, Szabolcs Marka, David Ottaway, Virginio Sannibale, Chiara Vanni, Hiroaki Yamamoto, and Sanichiro Yoshida. The seismic attenuation system (sas) for the advanced LIGO gravitational wave interferometric detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 598(3):737 – 753, 2009.
 - [23] F Matichard, B Lantz, R Mittleman, K Mason, J Kissel, B Abbott, S Biscans, J McIver, R Abbott, S Abbott, E Allwine, S Barnum, J Birch, C Celerier, D Clark, D Coyne, D DeBra, R DeRosa, M Evans, S Foley, P Fritschel, J A Giaime, C Gray, G Grabeel, J Hanson, C Hard-

- ham, M Hillard, W Hua, C Kucharczyk, M Landry, A Le Roux, V Lhuillier, D Macleod, M Macinnis, R Mitchell, B O'Reilly, D Ottaway, H Paris, A Pele, M Puma, H Radkins, C Ramet, M Robinson, L Ruet, P Sarin, D Shoemaker, A Stein, J Thomas, M Vargas, K Venkateswara, J Warner, and S Wen. Seismic isolation of advanced ligo: Review of strategy, instrumentation and performance. *Classical and Quantum Gravity*, 32(18):185003, 2015.
- [24] H. S. Kuyuk, S. Colombelli, A. Zollo, R. M. Allen, and M. O. Erdik. Automatic earthquake confirmation for early warning system. *Geophysical Research Letters*, 42(13):5266–5273, 2015. 2015GL063881.
- [25] J. Douglas. Earthquake ground motion estimation using strong-motion records: a review of equations for the estimation of peak ground acceleration and response spectral ordinates. *Earth Science Reviews*, 61:43–104, April 2003.
- [26] Richard M. Allen, Paolo Gasparini, Osamu Kamigaichi, and Maren Böse. The status of earthquake early warning around the world: An introductory overview. *Seismological Research Letters*, 80(5):682–693, 2009.
- [27] Diego Melgar, Brendan W. Crowell, Jianghui Geng, Richard M. Allen, Yehuda Bock, Sebastian Riquelme, Emma M. Hill, Marino Protti, and Athanassios Ganas. Earthquake magnitude calculation without saturation from the scaling of peak ground displacement. *Geophysical Research Letters*, 42(13):5197–5205, 2015. 2015GL064278.
- [28] G. F. Panza, C. L. Mura, A. Peresan, F. Romanelli, and F. Vaccari. Seismic Hazard Scenarios as Preventive Tools for a Disaster Resilient Society. *Advances in Geophysics*, 53:93–165, 2012.
- [29] Anil K. Chopra. Elastic response spectrum: a historical note. *Earthquake Engineering & Structural Dynamics*, 36(1):3–12, 2007.
- [30] Qingkai Kong, Richard M. Allen, Louis Schreier, and Young-Woo Kwon. Myshake: A smart-phone seismic network for earthquake early warning and beyond. *Science Advances*, 2(2), 2016.
- [31] H. S. Kuyuk, E. Yildirim, E. Dogan, and G. Horasan. An unsupervised learning algorithm: application to the discrimination of seismic events and quarry blasts in the vicinity of Istanbul. *Natural Hazards and Earth System Sciences*, 11:93–100, January 2011.
- [32] Jari Kortström, Marja Uski, and Timo Tiira. Automatic classification of seismic events within a regional seismograph network. *Computers & Geosciences*, 87(Supplement C):22 – 30, 2016.
- [33] T. Perol, M. Gharbi, and M. Denolle. Convolutional Neural Network for Earthquake Detection

- and Location. *ArXiv e-prints*, February 2017.
- [34] S. Mostafa Mousavi, Stephen P. Horton, Charles A. Langston, and Borhan Samei. Seismic features and automatic discrimination of deep and shallow induced-microearthquakes using neural network and logistic regression. *Geophysical Journal International*, 207(1):29–46, 2016.
 - [35] Yongna Jia and Jianwei Ma. What can machine learning do for seismic data processing? an interpolation application. *GEOPHYSICS*, 82(3):V163–V177, 2017.
 - [36] Steve Diersen, En-Jui Lee, Diana Spears, Po Chen, and Liqiang Wang. Classification of seismic windows using artificial neural networks. *Procedia Computer Science*, 4(Supplement C):1572 – 1581, 2011. Proceedings of the International Conference on Computational Science, ICCS 2011.
 - [37] Sebastien Biscans, Jim Warner, Richard Mittleman, Christopher Buchanan, Michael W Coughlin, Matthew Evans, Hunter Gabbard, Jan Harms, Brian Lantz, Nikhil Mukund, Arnaud Pele, Charles Pezerat, Pascal Picart, Hugh Radkins, and Thomas Shaffer. Control strategy to limit duty cycle impact of earthquakes on the ligo gravitational-wave detectors. *Classical and Quantum Gravity*, 2018.
 - [38] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, Manjunath Kudlur, Josh Levenberg, Rajat Monga, Sherry Moore, Derek G. Murray, Benoit Steiner, Paul Tucker, Vijay Vasudevan, Pete Warden, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. Tensorflow: A system for large-scale machine learning. In *Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation*, OSDI’16, pages 265–283, Berkeley, CA, USA, 2016. USENIX Association.
 - [39] David H Wolpert. Stacked generalization. *Neural networks*, 5(2):241–259, 1992.
 - [40] Mark J Van der Laan, Eric C Polley, and Alan E Hubbard. Super learner. *Statistical applications in genetics and molecular biology*, 6(1), 2007.
 - [41] Carl Edward Rasmussen and Christopher KI Williams. Gaussian processes for machine learning. 2006. *The MIT Press, Cambridge, MA, USA*, 38:715–719, 2006.
 - [42] Prasanta Chandra Mahalanobis. On the generalised distance in statistics. In *Proceedings of the National Institute of Sciences of India.*, UAI’95, page 49–55, 1936.
 - [43] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. SMOTE: Synthetic Minority Over-sampling Technique. *ArXiv e-prints*, June 2011.

- [44] S. Bonnefoy-Claudet, F. Cotton, and P.-Y. Bard. The nature of noise wavefield and its applications for site effects studies; A literature review. *Earth-Science Rev.*, 79:205, 2006.
- [45] J. Peterson. Observation and modeling of seismic background noise. *Open-file report*, 93-322, 1993.
- [46] J. Berger, P. Davis, and G. Ekström. Ambient Earth Noise: A survey of the Global Seismographic Network. *Journal of Geophysical Research*, 109:B11307, 2004.
- [47] D. E. McNamara, C. R. Hutt, L. S. Gee, H. M. Benz, and R. P. Buland. A Method to Establish Seismic Noise Baselines for Automated Station Assessment. *Seism. Res. Lett.*, 80:628, 2009.
- [48] R. A. Haubrich, W. H. Munk, and F. E. Snodgrass. Comparative spectra of microseisms and swell. *BSSA*, 53:27, 1963.
- [49] M. N. Toksöz and R. T. Lacoss. Microseisms: Mode Structure and Sources. *Science*, 159:872, 1968.
- [50] R. K. Cessaro. Sources of Primary and Secondary Microseisms. *BSSA*, 84:142, 1994.
- [51] A. Friedrich, F. Krüger, and K. Klinge. Oceangenerated microseismic noise located with the Gräfenberg array. *J. Seism.*, 2:47, 1998.
- [52] Mark Beker, Xander Campman, Julesvan Oven, Wim Walk, Jack Levell, Zijian Tang, Mike Danilouchkine, Dirk Smit, Johannesvan den Brand, Soumen Koley, and Maria Bader. Innovations in seismic sensors driven by the search for gravitational waves. *The Leading Edge*, 35(7):590–593, 2016.
- [53] Krishna Venkateswara, Charles A. Hagedorn, Matthew D. Turner, Trevor Arp, and Jens H. Gundlach. A high-precision mechanical absolute-rotation sensor. *Review of Scientific Instruments*, 85(1):015005, 2014.
- [54] The LIGO Scientific Collaboration and Virgo Collaboration. Characterization of transient noise in advanced ligo relevant to gravitational wave signal gw150914. *Classical and Quantum Gravity*, 33(13):134001, 2016.

Methods.

Seismic data.

We use multiple sources of seismic data. We perform an analysis of seismic time series that were made available through IRIS, covering the last ten years. These stations have time-series with response between 10 mHz to 10 Hz. The noise for these instruments is determined by a variety of sources including anthropogenic and atmospheric disturbances, earthquakes

and ocean waves [44]. Seismic noise models are developed using global seismometer arrays [45–47]. We use stations across the world to explore the effects of a variety of different sites, which can have noise spectra that have significant variation, due to location aspects such as topography and proximity to urban settlements. One source present across the world is the oceanic microseism around 0.3 Hz that dominate seismic ground spectra everywhere on Earth [48–51].

We systematically downloaded and processed data from all stations with channel names BH?. Stations are supplied with Nanometrics T240, Streckeisen STS-1/STS-2, Güralp CMG-3T and Geotech KS-54000 broadband seismometers. IRIS contains data for some stations as far back as the early 1970’s, and we analyze data from 2005 – 2017. We analyze data from (21412) earthquakes from (January 2005 to May 2017). The magnitudes range from (6.0) to (9.2), chosen to cover the range of earthquake magnitudes likely to significantly effect the gravitational-wave detectors.

We also use seismic data from gravitational-wave detectors covering the last two years. The Advanced LIGO [11] and the Advanced Virgo [12] detectors are multi-kilometer Michelson-based interferometers which have driven the development of both seismic [52] and rotation [53] sensors. Gravitational waves induce small displacements in the detectors, which are designed to be free from environmental disturbances and limited only by processes of fundamental physics. These detectors are subject to non-Gaussian noise transients due to either internal behavior of the instrument or interactions between the detector and its environment. To minimize the effect of the environment, the LIGO detectors contain 200,000 auxiliary channels which are designed to monitor both the behavior of the instrument and the environment conditions. A subset of these sensors is physical environmental monitor sensors dedicated to monitoring the environment, including seismometers, magnetometers, microphones, and many others. The LIGO and Virgo detectors contain arrays of seismometers, from which we take a seismometer in each of the central buildings [54]. These are useful for measuring any source of ground motion that can couple into the interferometers.

MLA description

In previous work, [20] had developed an empirical equation to predict the Rayleigh wave amplitude based magnitude, depth and distance to the earthquake. The model was able to capture 90% of the observational data within a factor of 5. To improve the accuracy further, we required more complex models that incorporate the physical mechanism of the

fault rupture and along with the directional information of the earthquake.

The deep neural network (DNN) that we employ to carry out the nonlinear regression has a topology inspired from generalized regression neural networks but we back-propagate the errors and update the weights by training it through several epochs. DNNs require larger data sets to learn the underlying function without overfitting the data. As the observed data could have measurement uncertainties, we boosted our data set by drawing additional samples from a normal distribution centered around each observation. This step prevented the early stopping during the training period and avoided the network to memorize the data. The trained network can capture the radiation pattern associated with the fault rupture making it a good candidate for earthquake warning systems.

Recently stacked ensemble regressors have gained much prominence and are consistently outperforming others in several datasets hosted at Kaggle. The first level consists of a set of base learners who are individually trained and cross-validated. Their predictions form input to second level meta-learner regressor which is further trained to generate the final ensemble prediction. Such systems are theoretically guaranteed to present be the optimal learners in the asymptotic sense .

For the Gaussian Process Regression or kriging model we use squared exponential kernel with prediction based on block coordinate descent. We optimize the hyper-parameters using bayesian optimization.

Mahalanobis distance is the multi-dimensional generalization of z-score which tell you how many sigma is the data away from the mean distribution. It is observed to be very robust technique as it takes into account the covariances between the variables. Our clustering technique makes use of this metric to find the closest matching earthquakes that happened in the past. This scheme naturally lets one identify outlier earthquakes with no similar events in the archival data.

TABLE II: Rf amplitude prediction performance of different ML algorithms for real and simulated data

	Deep Neural Nets	Stacked Ensemble	GPR	Clustering
<i>LIGO Livingston</i>	89% 85%	93% 89%	94% 87%	98% 94%
<i>LIGO Hanford</i>	86% 84%	91% 88%	92% 89%	97% 92%

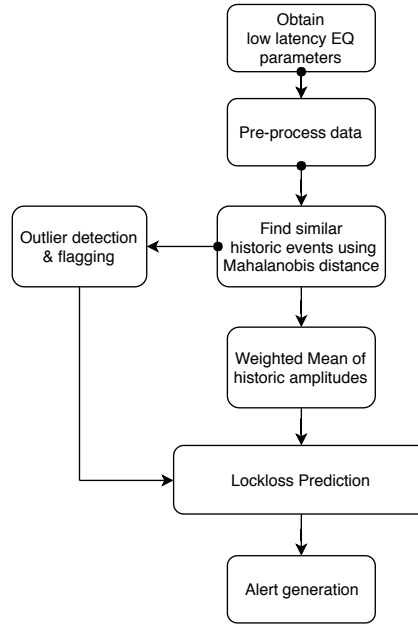


FIG. 5: Process flowchart depicting the low latency earthquake warning pipeline.