IOP Publishing Classical and Quantum Gravity

Class. Quantum Grav. **36** (2019) 085005 (14pp) https://doi.org/10.1088/1361-6382/ab0d2c

Ground motion prediction at gravitational wave observatories using archival seismic data

Nikhil Mukund[1,14], Michael Coughlin[2], Jan Harms[3,4], Sebastien Biscans[5,6], Jim Warner[7], Arnaud Pele[8], Keith Thorne[8], David Barker[7], Nicolas Arnaud[9,10], Fred Donovan[11], Irene Fiori[10], Hunter Gabbard[12], Brian Lantz[13], Richard Mittleman[11], Hugh Radkins[7] and Bas Swinkels[10]

[1] Inter-University Centre for Astronomy and Astrophysics (IUCAA), Post Bag 4, Ganeshkhind, Pune 411 007, India
[2] Division of Physics, Math, and Astronomy, California Institute of Technology, Pasadena, CA 91125, United States of America
[3] Gran Sasso Science Institute (GSSI), I-67100 L'Aquila, Italy
[4] Laboratori Nazionali del Gran Sasso, INFN, I-67100 Assergi, Italy
[5] LIGO, California Institute of Technology, Pasadena, CA 91125, United States of America
[6] LIGO, Massachusetts Institute of Technology, Cambridge, MA 02139, United States of America
[7] LIGO Hanford Observatory, Richland, WA 99352, United States of America
[8] LIGO Livingston Observatory, Livingston, LA 70754, United States of America
[9] LAL, Univ. Paris-Sud, CNRS/IN2P3, Université Paris-Saclay, F-91898 Orsay, France
[10] European Gravitational Observatory (EGO), I-56021 Cascina, Pisa, Italy
[11] LIGO Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02138, United States of America
[12] SUPA, School of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, United Kingdom
[13] Stanford University, Stanford, CA, United States of America

E-mail: nikhil@iucaa.in

Received 14 December 2018, revised 7 February 2019
Accepted for publication 6 March 2019
Published 1 April 2019

Abstract

Gravitational wave observatories have always been affected by tele-seismic earthquakes leading to a decrease in duty cycle and coincident observation time. In this analysis, we leverage the power of machine learning algorithms and archival seismic data to predict the ground motion and the state of the gravitational wave interferometer during the event of an earthquake. We

[14] Author to whom any correspondence should be addressed.

demonstrate 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. The level of accuracy achieved with this scheme makes it possible to switch control configuration during periods of excessive ground motion thus preventing the interferometer from losing lock. To further assess the accuracy and utility of our approach, we use IRIS seismic network data and obtain similar levels of agreement between the estimates and the measured amplitudes. The performance indicates that such an archival or prediction scheme can be extended beyond the realm of gravitational wave detector sites for hazard-based early warning alerts.

Keywords: GW detectors, seismic Rayleigh waves, earthquake early warning, machine learning

(Some figures may appear in colour only in the online journal)

1. Introduction

With the advent of gravitational wave (GW) astronomy, it is essential to maximize the duty cycle of second-generation gravitational-wave detectors such as the Laser Interferometer Gravitational-wave Observatory (LIGO) [1], Virgo [2], and GEO600 [3] detectors. Any increase in duty cycle increases the sensitivity of GW searches, including the observations of binary black hole mergers and binary neutron stars [4–10]. GWs from these 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 [11]. 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 environmental conditions. A subset of these is physical environmental monitor sensors, including seismometers, magnetometers, microphones, and many others. Advanced LIGO [1] and Advanced Virgo [2] have in particular driven the development of both seismic [12] and rotation [13] sensors. Seismic sensors in particular are useful for measuring any source of ground motion that can couple into the interferometers. LIGO seismic isolation systems by means of passive and active isolation provide noise suppression above 0.1 Hz [14–16] but are not effective against earthquake-related ground motion [17, 18]. The surface waves so produced hinder the process of keeping the instrument at a linear operating point and often induces higher frequency noise by up-converting low-frequency optical motion.

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 [19–28]. Worldwide, many seismic and geodetic (GPS) sensor arrays exist that produce 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 [28]. 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 ($\approx$8 km s^{-1}) and S-waves ($\approx$4 km s^{-1}). Reliability of these estimates are one of the

most important aspects of EEW systems. 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 [29]. Both of these strategies come at the cost of decreasing the warning time. 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 with high accuracy the ground motion that the earthquakes will cause as a function of location and distance.

This paper is organized as follows: section 2 talks about the sensitivity of the GW interferometers to earthquakes and the previous attempts to model them. In section 3 we describe how the seismic data was obtained at the site as well as the IRIS seismic array [30, 31]. Section 4 describes the deployed regression and clustering techniques and their requirements. Finally, the relative performance of various prediction algorithms along with the ability to guess the state of the interferometer are covered in section 5.

2. Impact assessment

Figure 1 depicts the distribution of global seismic events and their respective effect on the state of the GW interferometers during LIGO's first and second observation run. The orange circles (scaled as per the magnitude) represent scenarios where the ground motion was high enough to cause instabilities leading to a loss of resonance in the cavity (lockloss). Sensitivities to parameters such as magnitude and surface wave amplitude are shown in figure 2 where the steepness of the curve indicates higher sensitivity to the respective parameter. As expected, LLO is seen to be more vulnerable to ground shaking which can be attributed to its local geology and soil properties [32]. The primary goal of LIGO/Virgo EEW methods would be to generate reliable relations 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, displacements and predominant periods [33]. Early estimates of magnitudes tend to underestimate the energy released from the tectonic motion and hence the early estimates of ground velocity amplitudes often tend to be not as accurate as the 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 under-predicted. The loss of performance that results from use of the rapid estimates is acceptable for the purpose of issuing rapid warnings. Direct displacement measurement based on non-inertial GPS instruments has been shown to generate rapid magnitude estimates within the first minute of rupture and in many cases even before rupture is complete [34].

In previous work, we used advances in early earthquake warning to develop a low-latency earthquake early warning client named *Seismon* [18]. This system uses a real-time event messaging system of the U.S. Geological Survey (USGS) to mitigate the effects of tele-seismic events on ground-based gravitational-wave detectors. Using information about the earthquake source characteristics such as magnitude, depth, and distance from the site, ground motion velocity at the site was 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 five of the predicted value. Although the fit was derived with physical effects in mind, it was predominantly an empirical construction. To make accurate assessments of whether the gravitational-wave detectors will be affected, we prefer the relative error in the ground velocity predictions, $\mathrm{Rf}_{\mathrm{amp}}$, to be close to a factor of two, which is much smaller than the factor

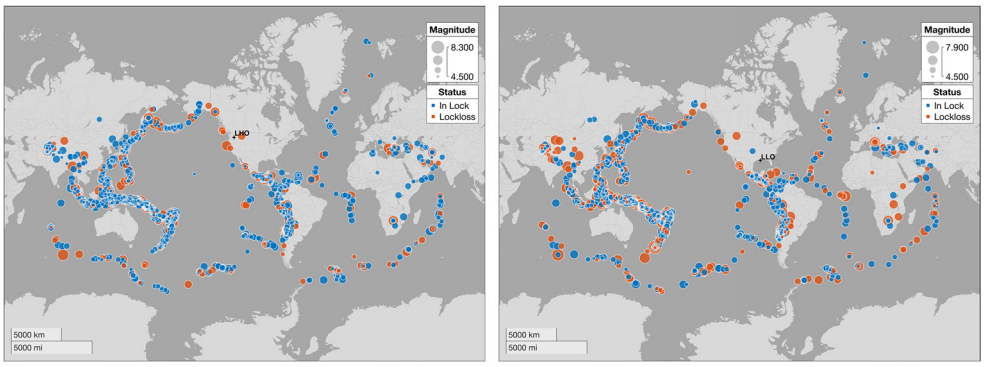

Figure 1. Impact of earthquakes happening worldwide on LIGO Interferometers at Hanford and Livingston. Points marked in orange indicate the instances when the resulting ground the motion lead to a loss of control over the ability to keep the interferometer at its operating point resulting in a lockloss.

of five scatter seen. These predictions have two purposes. Firstly, 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 [35]. The predictions also serve as inputs to the algorithms that make predictions on upcoming downtime, which could be utilized to perform opportunistic maintenance to rectify problems typically scheduled for weekly maintenance periods.

3. Methods: seismic data

The first part of our analysis uses data obtained from the GW sites. For each earthquake event from the archival database, we take the vertical component of broadband data which is converted to ground velocity (m/s) using a frequency dependent calibration factor appropriate for each seismometer. Time-series are chosen to encompass the P-wave arrival and surface wave calculated assuming a (very conservative) seismic velocity of 2 km s^{-1} [36]. As the seismometers located at the end and center stations observe similar values for the relevant frequency band, we use the center station sensor for rest of our analysis.

We also perform an analysis of seismic time-series that were made available through IRIS seismometer array, covering the last ten years. These stations covering all the conterminous US states (see figure 4) have time-series with response between 10 mHz to 10 Hz. A variety of sources including anthropogenic and atmospheric disturbances contribute to the observed seismic signal. One source present across the world is the oceanic microseism around 0.3 Hz that dominate seismic ground spectra everywhere on Earth [37–40]. Systematic processing of IRIS data from January 2006 to November 2017 lead us to create a database of around 733 208 earthquakes. The analysis was restricted to earthquakes with magnitude ranging from 6.0 to 9.1, which covers the range that is likely to significantly effect the gravitational-wave detectors. In section 5 we report the performance of the prediction scheme in the presence of variation in local seismic spectra arising from affects such as local geology, topography and proximity to urban settlements.

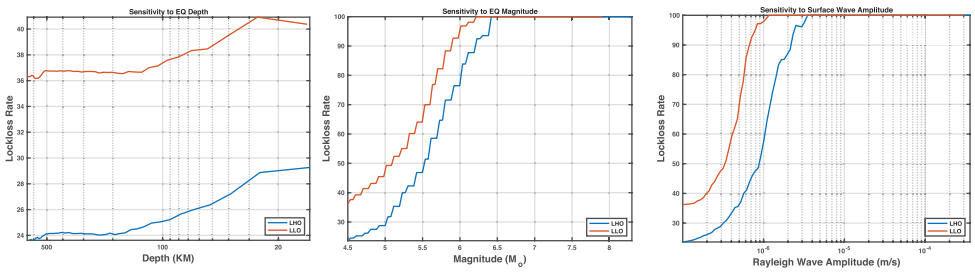

Figure 2. Plot shows the sensitivity of both the LIGO detectors to earthquakes in terms of distance to the hypocenter from the surface, earthquake magnitude and the Rayleigh wave amplitude.

4. Methods: MLA description

Machine learning (ML) has recently become an important aspect of EEW and seismology in general. For example, the *MyShake* EEW system uses artificial neural networks to differentiate earthquake and human motions, with 98% of earthquake records within 10 km of the epicenter correctly identified, and only 7% of people-induced transients appearing to be earthquakes to the algorithm [41]. They have been used to differentiate earthquakes from other seismic transients [42–44], discriminate between deep and shallow micro-earthquakes [45] and to add to undersampled or missing traces [46]. In addition, they have been used to make full-wave tomography images [47]. The idea of our analysis is to compare historical ground velocity measurements to predictions made using different machine learning algorithm techniques. The parameters that enter the predictions are M, the magnitude of the earthquake, h, the depth, $\log(r)$, the distance to the detectors in log scale, θ and ϕ, the latitude and longitude, and α, the earthquake azimuth relative to the detector. 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 as additional parameters which could be used in the future to improve the prediction accuracy, albeit not at low latency. The target variables correspond to peak ground velocities measured using seismometers. This scheme is advantageous over the analytical equation in a few ways. First of all, by switching to ML, we eliminate the dependence 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 of two different machine learning approaches: regression and clustering. Within regression, we evaluate the performance of the Tensorflow implementation of deep neural networks (DNN) [48], stacked ensemble regressors [49] and Gaussian process regression (GPR) [50], while in clustering we use a Mahalanobis distance [51] based similarity search to make the predictions. The performance of each algorithm is accessed using real and simulated datasets. For each dataset, 80% is used for training with the remaining 20% used for predictions. As for simulated data, new samples are generated from each of the original datasets by creating a Gaussian jitter distribution centered around the parameter value followed by a random draw of samples from these distributions. Artificially adding noise (or jitter) to the predictor and response variables in a controlled fashion helps to improve the learning and prevent early stopping. The presence of noise enhances the ability of the ML algorithms to better learn and generalize to the underlying smooth, non-linear function. Variance of the jitter distribution should be chosen such that the synthetic samples created are neither completely nonsensical nor very much identical to the original data set [52]. In most of the cases, USGS do provide us with the level of uncertainty associated with

the estimated earthquake parameters. We set the variance of the jitter distribution such that the new samples generated around the original values are well within the measurement uncertainties of each of the variables. For a given input taken from the jittered population, the predicted output value would be the weighted average of the outputs of the original training samples with the weights proportional to the densities of the jittered distribution. For a very large number of such samples, the average predicted value turns out to be the Nadaraya–Watson kernel regression estimator [53] where the variance in jitter corresponds to the bandwidth of the estimator [54].

The deep neural network (DNN) that we employ to carry out the nonlinear regression has a topology inspired from generalized regression neural networks [55], but we back-propagate the errors and update the weights by training it through several epochs. DNNs, in general, require larger data sets to learn the underlying function without overfitting the data and tend to be sensitive to the network architecture and the activation functions. We use a sequential network with nine dense layers with exponential linear unit activation and the first-order gradient-based Adam optimizer [56]. In recent times, stacked ensemble regressors have also gained much prominence and are seen to consistently outperform other competing algorithms in several datasets hosted at the Kaggle challenge [57]. The first level consists of a set of base learners that are individually trained and cross-validated. Their predictions form input to a second level meta-learner regressor which is further trained to generate the final ensemble prediction. Such systems are theoretically guaranteed to be the optimal learners in an asymptotic sense [58]. Success with DNN and ensemble techniques crucially depends on the amount of training data and could be sensitive to the hyper-parameters. As for the GPR, we optimize the respective hyper-parameters using Bayesian optimization and use a squared exponential kernel for covariance estimation. But the method scales as $O(n^3)$ thus resulting in high memory requirements and training time for large data sets.

Mahalanobis distance is the multi-dimensional generalization of z-score which tells you how many sigmas the data is away from the mean distribution. It is observed to be a 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 1 compares the performance of the different MLAs. We find the clustering-based predictor to be the most accurate and select is as the default algorithm for EEW pipeline shown in figure 3. In addition to having the best prediction accuracy, it has the following advantages. Firstly, as there is no training involved, the need for hyper-parameter tuning is eliminated. As we are constantly monitoring the seismic data and appending the earthquake database, with time we expect a decrease in prediction error along with a reduction in the number of outliers.

5. Results

Figures 5 and 6 show the prediction results from applying Mahalanobis based clustering on the simulated and real earthquake data. The events have been ordered by their measured peak ground velocity (in orange), and the error bar (in grey) corresponds to a factor of 2.5 within the measured value. As compared to the previously used analytic expression, this machine learning based scheme leads to a performance improvement from a factor of 5 to 2.5 in scatter of the error in the predicted ground velocity. For both real and simulated cases, accuracy in predictions (in blue) is observed to be above 90% for LIGO and 85% for Virgo datasets. We can attribute the improvement to the increased availability of data, the inclusion of more

Table 1. Rf amplitude prediction performance for different ML algorithms. The top and bottom panel give the percentage of events correctly predicted within a factor of 2.5 by each of the ML algorithm for simulated and real data.

	Deep neural nets	Stacked ensemble	GPR	Clustering
LIGO Livingston	89 % / 85 %	93 % / 89 %	94% / 87 %	98% / 94%
LIGO Hanford	86% / 84%	91% / 88%	92% / 89%	97% / 92%

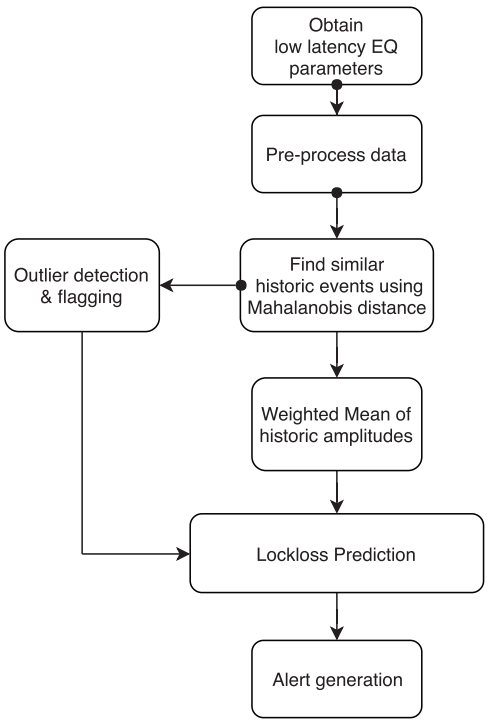

Figure 3. Process flowchart depicting the low latency earthquake warning pipeline.

earthquake parameters and the usage of robust algorithms. This level of accuracy is sufficient to switch control configuration during periods of excessive ground motion thus preventing the interferometer from losing lock. As for the outliers, they do not seem to show any specific dependence concerning the input parameters. One reason could be that their parameter combination is rather uncommon, so the predicted amplitude is averaged across not-so-similar events. The general trend among outliers seems to be that the higher amplitude events are underestimated, and the lower amplitude events are over-estimated. Such outliers should decrease as we gather more training data. Bringing down the mismatch, especially at the lower amplitude side, would require improvements in signal to noise ratio of the ground velocity measurements carried out at the site along with the need for more detailed initial earthquake information. Availability of low latency information regarding the centroid moment tensors associated with the earthquake could be one such solution.

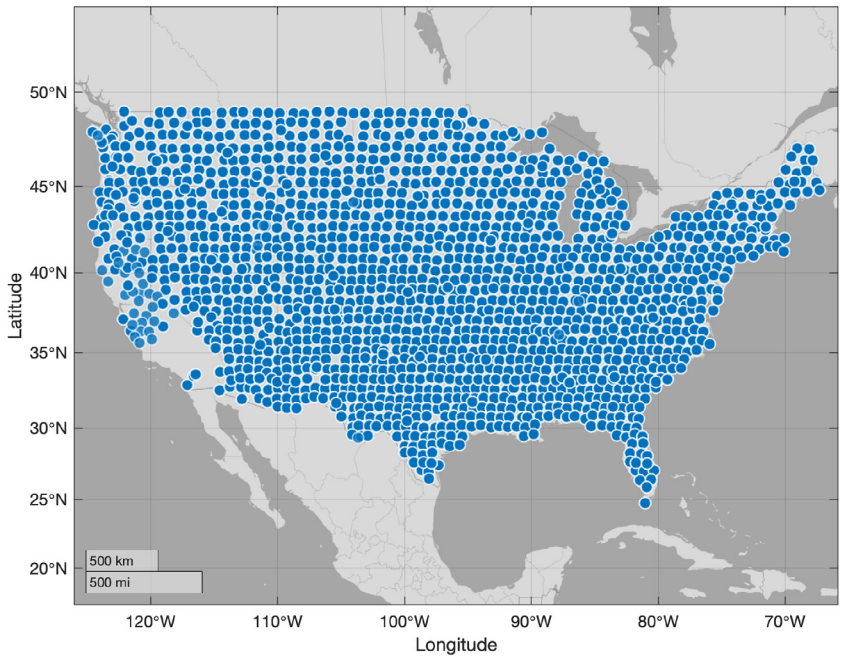

Figure 4. Locations of IRIS seismometer array used to collect the earthquake data from January 2006 to November 2017.

We also demonstrate the resourcefulness of the above scheme by making predictions across the United States using the data recorded by the IRIS network. For ground motion recorded in each state, we perform the similarity search using archival data and compare the prediction accuracy as shown in figure 7. For each row representing a state, we show how well we can make ground velocity predictions based on archival events from itself as well as other states. For at least 24 states, the accuracy is seen to be above 90% when we use its training data for predicting the new events. 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. High values seen along several of the off-diagonal terms are due to the nearly identical response to tele-seismic events could mean the corresponding similarity within their geological properties such as shear velocity profiles, elasticity, and local soil density, etc. In future, state-wise seismic modeling could use this information to augment the state-wise data with the ones from similar states especially if the original dataset is sparse. The performance shows that the archival data-based prediction scheme can be extended beyond the realm of gravitational-wave detector sites for hazard-based early warning alerts. For the state of Louisiana, the data from previous states do not seem to make predictions with a high level of accuracy. As expected, comparatively better performance is obtained from the neighboring states of Texas, Arkansas, and Mississippi. Since similar anomalies are not observed in Florida and Texas data, proximity to the Gulf of Mexico does not seem to a strong reason for this behavior. Usage of multiple sensors within each state and the adoption of standardized data analysis procedure which includes the individual sensor calibration and data preprocessing minimizes the likelihood of systematic and instrumental effects to bias the results. Another

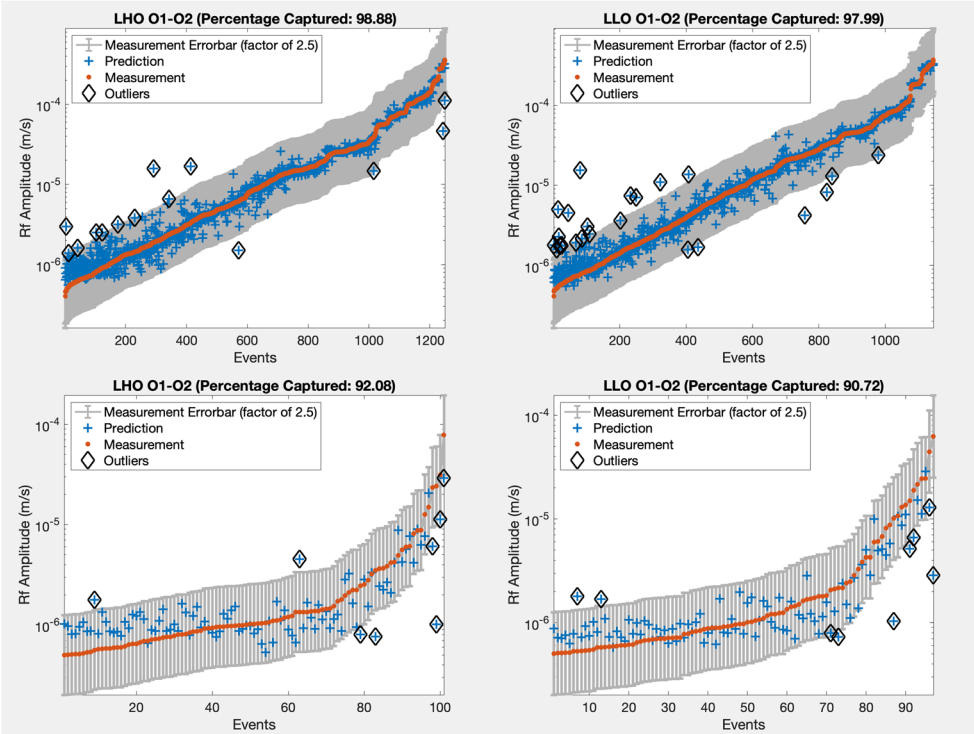

Figure 5. Fit of peak velocities seen during O1–O2 at the interferometers (LHO, LLO) using Mahalanobis distance-based clustering. Results on simulated and real data are respectively shown in the top and bottom rows. The events have been ordered by their measured peak ground velocity (in orange), and grey error bar corresponds to a factor of 2.5 within the measured value. More than 90% of predictions (in blue) are within a factor of 2.5 of the measured value.

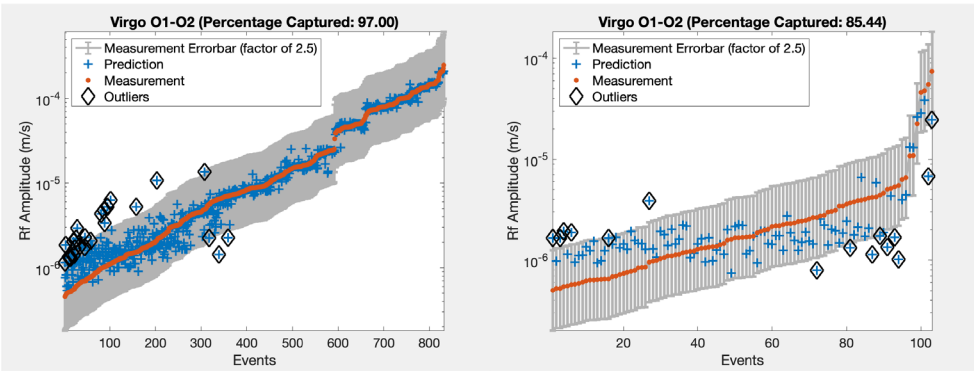

Figure 6. Fit of peak velocities seen during O1–O2 at the Virgo interferometer.

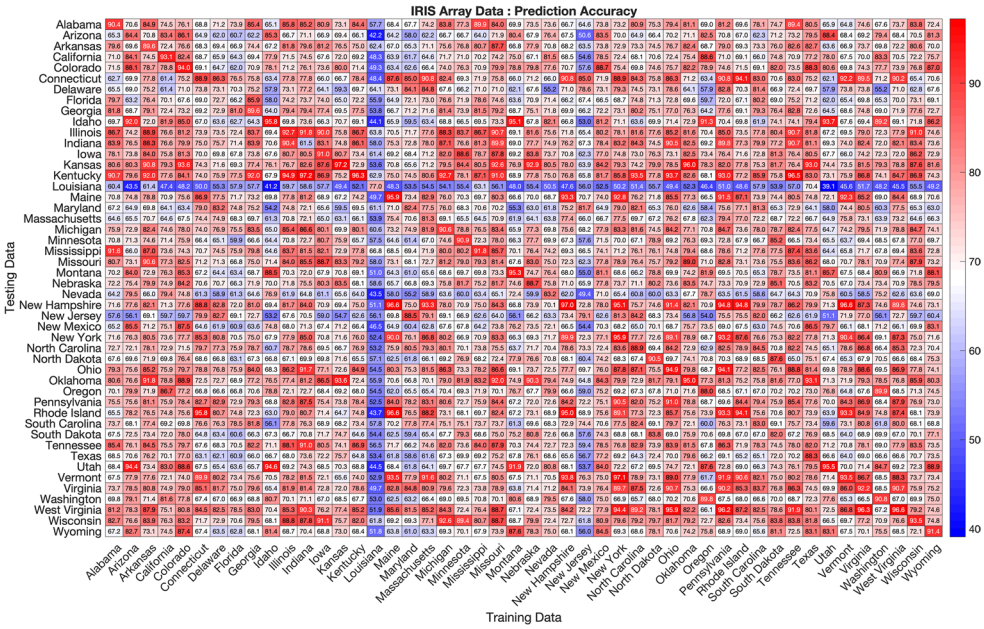

Figure 7. The heat-map on the right shows the ground motion prediction accuracy within each state making use of archival data from all the conterminous US states.

Table 2. Performance analysis of lockloss prediction models for LHO and LLO. Each of them respectively has an accuracy of 92% and 93%.

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

possibility is from peculiarities in local geology, but further studies would be required to understand the exact cause of this anomaly.

The main benefit of ground velocity predictions for gravitational-wave detectors is to inform predictions of whether an earthquake will cause the loss of data for the detector. We have previously developed techniques for preventing earthquakes from causing the loss of data taking if advanced notice is given [35]. In this work, we use the previously described

clustering technique 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. Acknowledging that there is a trade-off between false-alarm probability and efficiency standard, we can make predictions for the inliers with an accuracy above 92%, keeping the associated false-alarm probability to be less than 10% (See table 2). Given this scenario of an impending lockloss, it would be more desirable to switch to a configuration that provides enough freedom for common mode motion but at the same time enforces the best possible suppression for the local differential motion.

6. Conclusions

In conclusion, we have used MLAs to predict peak ground velocities from tele-seismic earthquakes. The estimated ground velocity is used to forecast the potential effect of earthquakes on gravitational-wave detectors and issue near real-time alerts at the site. The alert system based on this scheme has been implemented at the Advanced GW observatories and will be used shortly to switch seismic filters at very low frequencies. 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 community. While we focus on the prediction of peak ground velocity here, in future we will explore the possibility of using historical seismometer data along with CMT parameters to predict radiation patterns associated with the fault rupture. 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.

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

Acknowledgments

NM acknowledges Council for Scientific and Industrial Research (CSIR), India, for providing financial support as Senior Research Fellow. MC was supported by the David and Ellen Lee Postdoctoral Fellowship at the California Institute of Technology. Authors express thanks to Duncan Agnew, Rich Ormiston and Brian O'Reilly for their valuable comments and suggestions. 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-P1800312. 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 Science Foundation under Cooperative Agreement EAR-126168.

ORCID iDs

Nikhil Mukund https://orcid.org/0000-0002-8666-9156
Jan Harms https://orcid.org/0000-0002-7332-9806
Sebastien Biscans https://orcid.org/0000-0002-9635-7527
Nicolas Arnaud https://orcid.org/0000-0001-6589-8673
Bas Swinkels https://orcid.org/0000-0002-3066-3601

References

[1] Aasi J *et al* 2015 Advanced LIGO *Class. Quantum Grav.* **32** 074001
[2] Acernese F *et al* 2015 Advanced Virgo: a second-generation interferometric gravitational wave detector *Class. Quantum Grav.* **32** 024001
[3] Grote H and For the LIGO Scientific Collaboration 2010 The GEO 600 status *Class. Quantum Grav.* **27** 084003
[4] Abbott B P *et al* 2016 Observation of gravitational waves from a binary black hole merger *Phys. Rev. Lett.* **116** 061102
[5] Abbott B P *et al* 2016 GW151226: observation of gravitational waves from a 22-solar-mass binary black hole coalescence *Phys. Rev. Lett.* **116** 241103
[6] Abbott B P *et al* 2017 GW170104: observation of a 50-solar-mass binary black hole coalescence at redshift 0.2 *Phys. Rev. Lett.* **118** 221101
[7] Abbott B P *et al* 2017 GW170814: a three-detector observation of gravitational waves from a binary black hole coalescence *Phys. Rev. Lett.* **119** 141101
[8] Abbott B P *et al* 2017 GW170608: observation of a 19 solar-mass binary black hole coalescence *Astrophys. J. Lett.* **851** L35
[9] Abbott B P *et al* 2017 GW170817: observation of gravitational waves from a binary neutron star inspiral *Phys. Rev. Lett.* **119** 161101
[10] The LIGO Scientific Collaboration and Virgo Collaboration 2018 GWTC-1: a gravitational-wave transient catalog of compact binary mergers observed by ligo and virgo during the first and second observing runs (arXiv:1811.12907)
[11] The LIGO Scientific Collaboration and Virgo Collaboration 2016 Characterization of transient noise in advanced LIGO relevant to gravitational wave signal GW150914 *Class. Quantum Grav.* **33** 134001
[12] Beker M, Campman X, Oven J, Walk W, Levell J, Tang Z, Danilouchkine M, Smit D, den Brand J, Koley S and Bader M 2016 Innovations in seismic sensors driven by the search for gravitational waves *Leading Edge* **35** 590–3
[13] Venkateswara K, Hagedorn C A, Turner M D, Arp T and Gundlach J H 2014 A high-precision mechanical absolute-rotation sensor *Rev. Sci. Instrum.* **85** 015005
[14] Abbott R *et al* 2002 Seismic isolation for Advanced LIGO *Class. Quantum Grav.* **19** 1591
[15] Stochino A *et al* 2009 The seismic attenuation system (sas) for the Advanced LIGO gravitational wave interferometric detectors *Nucl. Instrum. Methods Phys. Res.* A **598** 737–53
[16] Matichard F *et al* 2015 Seismic isolation of Advanced LIGO: Review of strategy, instrumentation and performance *Class. Quantum Grav.* **32** 185003
[17] Coughlin M, Stubbs C, Barrientos S, Claver C, Harms J, Chris Smith R and Warner M 2015 Real-time earthquake warning for astronomical observatories *Exp. Astron.* **39** 387–404
[18] Coughlin M *et al* 2017 Limiting the effects of earthquakes on gravitational-wave interferometers *Class. Quantum Grav.* **34** 044004
[19] Allen R M 2012 Transforming earthquake detection? *Science* **335** 297–8
[20] Kuyuk H S and Allen R M 2013 A global approach to provide magnitude estimates for earthquake early warning *Geophys. Res. Lett.* **40** 6329–33
[21] Kuyuk H S and Allen R M 2013 Optimal seismic network density for earthquake early warning: a case study from California *Seismol. Res. Lett.* **84** 946–54
[22] Kuyuk H S *et al* 2014 Designing a network-based earthquake early warning system for California: ElarmS-2 *Bull. Seismol. Soc. Am.* **104** 162–73

[23] Cochran E, Lawrence J, Christensen C and Chung A 2009 A novel strong-motion seismic network for community participation in earthquake monitoring *IEEE Instrum. Meas.* **12** 8–15
[24] Cochran E, Lawrence J, Christensen C and Jakka R 2009 The quake-catcher network: citizen science expanding seismic horizons *Seismol. Res. Lett.* **80** 26–30
[25] Bose M *et al* 2014 CISN shakealert: an earthquake early warning demonstration system for California *Early Warning for Geological Disasters* (*Advanced Technologies in Earth Sciences*) (Berlin: Springer) pp 49–69
[26] Hoshiba M, Kamigaichi O, Saito M, Tsukada S and Hamada N 2008 Earthquake early warning starts nationwide in Japan *Eos, Trans. Am. Geophys. Union* **89** 73–4
[27] Hoshiba M, Iwakiri K, Hayashimoto N, Shimoyama T, Hirano K, Yamada Y, Ishigaki Y and Kikuta H 2011 Outline of the 2011 off the Pacific coast of Tohoku earthquake (Mw 9.0)—earthquake early warning and observed seismic intensity *Earth Planets Space* **63** 7
[28] Strauss J A and Allen R M 2016 Benefits and costs of earthquake early warning *Seismol. Res. Lett.* **87** 765–72
[29] Kuyuk H S, Colombelli S, Zollo A, Allen R M and Erdik M O 2015 Automatic earthquake confirmation for early warning system *Geophys. Res. Lett.* **42** 5266–73
[30] IRIS data services products: usarraygmv-super (http://ds.iris.edu/ds/products/usarraygmv-super/) (Accessed: 15 June 2018)
[31] Trabant C, Hutko A R, Bahavar M, Karstens R, Ahern T and Aster R 2012 Data products at the iris dmc: stepping stones for research and other applications *Seismol. Res. Lett.* **83** 846–54
[32] Daw E J, Giaime J A, Lormand D, Lubinski M and Zweizig J 2004 Long-term study of the seismic environment at LIGO *Class. Quantum Grav.* **21** 2255
[33] Douglas J 2003 Earthquake ground motion estimation using strong-motion records: a review of equations for the estimation of peak ground acceleration and response spectral ordinates *Earth Sci. Rev.* **61** 43–104
[34] Melgar D, Crowell B W, Geng J, Allen R M, Bock Y, Riquelme S, Hill E M, Protti M and Ganas A 2015 Earthquake magnitude calculation without saturation from the scaling of peak ground displacement *Geophys. Res. Lett.* **42** 5197–205
[35] Biscans S *et al* 2018 Control strategy to limit duty cycle impact of earthquakes on the LIGO gravitational-wave detectors *Class. Quantum Grav.* **35** 055004
[36] Shearer P M 2009 *Introduction to Seismology* (Cambridge: Cambridge University Press) (https://doi.org/10.1017/CBO9780511841552)
[37] Haubrich R A, Munk W H and Snodgrass F E 1963 Comparative spectra of microseisms and swell *BSSA* **53** 27
[38] Toksöz M N and Lacoss R T 1968 Microseisms: mode structure and sources *Science* **159** 872
[39] Cessaro R K 1994 Sources of primary and secondary microseisms *BSSA* **84** 142
[40] Friedrich A, Krüger F and Klinge K 1998 Oceangenerated microseismic noise located with the Gräfenberg array *J. Seism.* **2** 47
[41] Kong Q, Allen R M, Schreier L and Kwon Y-W 2016 Myshake: a smartphone seismic network for earthquake early warning and beyond *Sci. Adv.* **2** e1501055
[42] Kuyuk H S, Yildirim E, Dogan E and Horasan G 2011 An unsupervised learning algorithm: application to the discrimination of seismic events and quarry blasts in the vicinity of Istanbul *Nat. Hazards Earth Syst. Sci.* **11** 93–100
[43] Kortström J, Uski M and Tiira T 2016 Automatic classification of seismic events within a regional seismograph network *Comput. Geosci.* **87** 22–30
[44] Perol T, Gharbi M and Denolle M 2017 Convolutional neural network for earthquake detection and location *Science Advances* **4** e1700578
[45] Mousavi S M, Horton S P, Langston C A and Samei B 2016 Seismic features and automatic discrimination of deep and shallow induced-microearthquakes using neural network and logistic regression *Geophys. J. Int.* **207** 29–46
[46] Jia Y and Ma J 2017 What can machine learning do for seismic data processing? An interpolation application *Geophysics* **82** V163–77
[47] Diersen S, Lee E-J, Spears D, Chen P and Wang L 2011 Classification of seismic windows using artificial neural networks *Proc. Comput. Sci.* **4** 1572–81
[48] Abadi M *et al* 2016 Tensorflow: a system for large-scale machine learning *Proc. of the 12th USENIX Conf. on Operating Systems Design and Implementation (Berkeley, CA, USA)* (USENIX Association) pp 265–83
[49] Wolpert D H 1992 Stacked generalization *Neural Netw.* **5** 241–59

[50] Rasmussen C E and Williams C K 2006 *Gaussian Processes for Machine Learning. 2006* vol 38 (Cambridge, MA: MIT Press) pp 715–9
[51] Mahalanobis P C 1936 On the generalised distance in statistics *Proc. National Institute of Sciences of India* pp 49–55
[52] Holmstrom L and Koistinen P 1992 Using additive noise in back-propagation training *IEEE Trans. Neural Netw.* **3** 24–38
[53] Nadaraya E A 1964 On estimating regression *Theory Probab. Appl.* **9** 141–2
[54] Scott D W 2015 *Multivariate Density Estimation: Theory, Practice, and Visualization* (New York: Wiley)
[55] Specht D F 1991 A general regression neural network *IEEE Trans. Neural Netw.* **2** 568–76
[56] Kingma D P and Ba J 2014 Adam: a method for stochastic optimization (arXiv:1412.6980)
[57] Kaggle website www.kaggle.com (Accessed: 15 June 2018)
[58] Van der Laan M J, Polley E C and Hubbard A E 2007 Super learner *Stat. Appl. Genet. Mol. Biol.* **6**