Predicting surface wave velocities at gravitational wave observatories using historical seismic data

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

[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, USA*
[3]*Gran Sasso Science Institute (GSSI), I-67100 L'Aquila, Italy*
INFN, Laboratori Nazionali del Gran Sasso, I-67100 Assergi, Italy
[4]*LAL, Univ. Paris-Sud, CNRS/IN2P3, Université Paris-Saclay, F-91898 Orsay, France*
[5]*European Gravitational Observatory (EGO), I-56021 Cascina, Pisa, Italy*
[6]*LIGO Hanford Observatory, Richland, WA 99352, USA*
[7]*LIGO Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02138, USA*
[8]*Albert-Einstein-Institut, Max-Planck-Institut für Gravitationsphysik, D-30167 Hannover, Germany*
[9]*Stanford University, Stanford, CA, USA*
[10]*LIGO Livingston Observatory, Livingston, LA 70754, USA*
[11]*Department of Physics, Cypress College, Cypress, CA 90630, USA*

In this analysis, we leverage the power of machine learning algorithms and historic seismic data to predict the ground velocity and the state of the gravitational wave interferometer during the event of an earthquake. We 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 archival data-based prediction scheme can be extended beyond the realm of gravitational wave detector sites for hazard-based early warning alerts.

I. 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 [4–7] and binary neutron stars [8]. 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 [9]. 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 sensors is physical environmental monitor sensors dedicated to monitoring the environment, including seismometers, magnetometers, microphones, and many others. The Advanced LIGO [1] and the Advanced Virgo [2] have in particular driven the development of both seismic [10] and rotation [11] sensors. Among these, seismometers are useful for measuring any source of ground motion that can couple into the interferometers. One source of ground motion that destabilizes the detectors are earthquakes [12, 13], despite seismic isolation systems designed to minimize such effects [14–16]. The surface waves, typically the highest amplitude component from earthquakes in the frequency band of interest 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 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 [17–26]. 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 [26]. 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

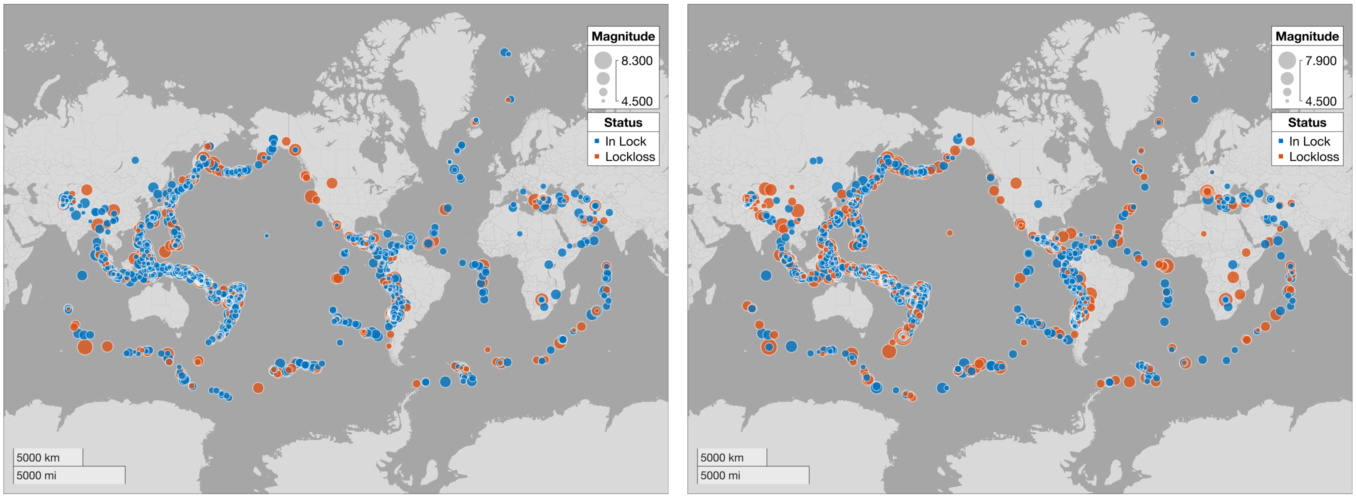

FIG. 1: Impact of earthquakes happening worldwide on LIGO Interferometers at Hanford and Livingston. Points marked in red indicate the instances when the resulting ground the motion caused the interferometer to go out of lock..

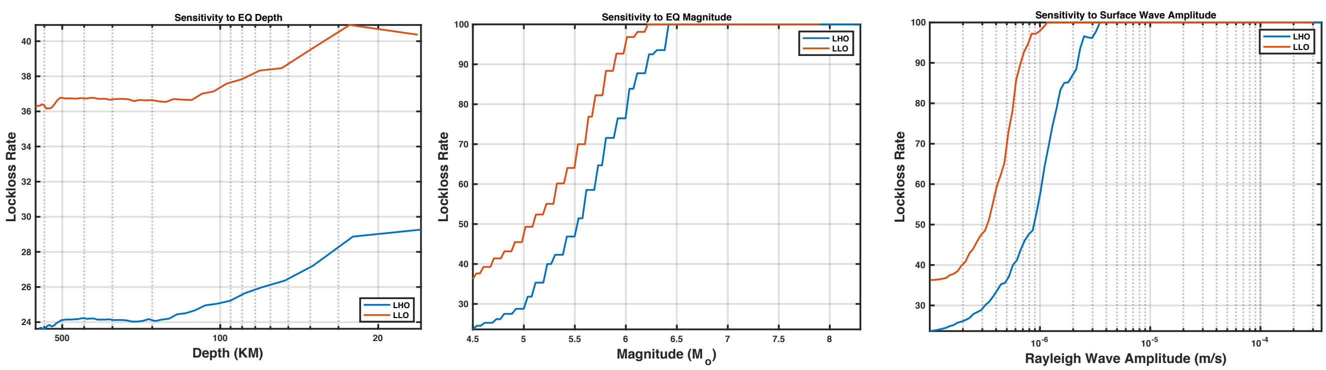

FIG. 2: Plot shows the lockloss rate associated with earthquake magnitude and Rayleigh wave amplitude at both the LIGO detectors.

positives [27]. 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 the ground motion that the earthquakes will cause as a function of location and distance with high accuracy.

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 correctly identified, and only 7% of people-induced transients appearing to be earthquakes to the algorithm [28]. They have been used to differentiate earthquakes from other seismic transients [29–31], discriminate between deep and shallow microearthquakes [32] and to add to undersampled or missing traces [33]. In addition, they have been used to make full-wave tomography images [34]. In this paper, we use advances in ML techniques to make ground velocity predictions using seismic data at both GW observatories and across the country. This paper is organized as follows. Section II discusses the sensitivity of the GW interferometers to earthquakes and the previous attempts to model them. In Section III, we describe how the seismic data was obtained at the GW detector sites as well as the IRIS seismic array from around the world. Section IV 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 V. Our conclusions are presented in Section VI.

II. IMPACT ASSESSMENT

Fig.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 blue circles (scaled as per the magnitude) represent scenarios where the ground motion was high enough to cause to instabilities leading to a state of lockloss. Sensitivities to parameters such as magnitude and surface wave amplitude are shown in Fig.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 could be attributed to its local geology and soil properties. The main goal of LIGO/Virgo EEW methods would be to generate reliable relations between and 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 [35]. 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 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 underpredicted. The loss of performance that results from use of the rapid estimates is acceptable to use as rapid warnings. [36] showed that real-time GPS waveforms could rapidly determine the magnitude within the first minute of rupture and in many cases before rupture is complete.

In previous work, we used advances in early earthquake warning to develop a low-latency earthquake early warning client named *Seismon* [13]. 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 magnitude, depth and distance 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 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 to make accurate assessments of whether the gravitational-wave detectors will be affected, which is much smaller than the factor of 5 scatter seen.

One of the key aspects of the system is the ground velocity predictions, $\mathrm{Rf}_{\mathrm{amp}}$, 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.

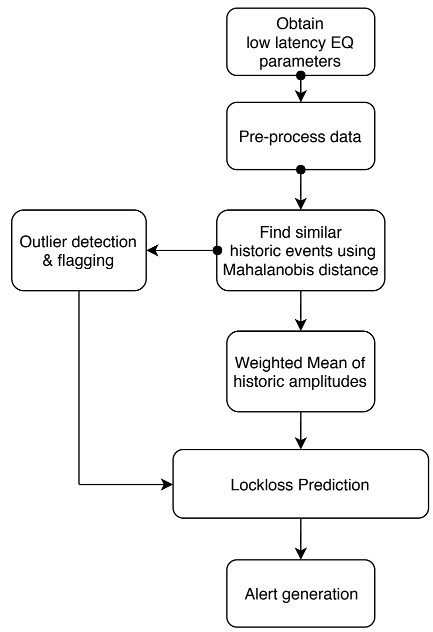

FIG. 3: Process flowchart depicting the low latency earthquake warning pipeline.

III. 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 (velocity) data that is filtered using an acausal 0.1 Hz low-pass Butter-worth filter. The data is calibrated into ground velocity using a constant V to m/s value appropriate for each seismometer. Time-series are chosen to encompass the P-wave arrival to surface waves calculated assuming a (very conservative) seismic velocity of 2 km/s. As the seismometers located at the end and center stations observed similar values for the relevant frequency band, we use the center station sensor for the GW detector analyses.

We also 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 [38]. Seismic noise models are developed using global seismometer arrays [39–41]. We use stations across 50 US states 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 [42–45]. IRIS contains data for some stations as far back as the early 1970's, of which we systematically downloaded and processed data from almost all stations from (January 2005 to May 2017) consisting of (21412) earthquakes. Stations are supplied with Nanometrics T240, Streckeisen STS-1/STS-2, Güralp CMG-3T and Geotech KS-54000 broadband seismometers. 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.

IV. MACHINE LEARNING ALGORITHM DESCRIPTION

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 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 as the two different machine learning approaches: regression and clustering. Within regression, we evaluate the performance of the Tensorflow implementation of deep neural networks (DNN) [46], stacked ensemble regressors [47, 48] and Gaussian Process Regression (GPR) [49], while in clustering we use a Mahalanobis distance [50] based similarity search to make the predictions. 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 parameters which could be help the prediction process. The target variables correspond to peak ground velocities measured using seismometers.

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 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 MLA to better learn and generalize to the underlying smooth, non-linear function. Selective boosting is done to minimize the imbalances in the dataset using Synthetic Minority Oversampling TEchnique (SMOTE) [51]. Prediction results are quoted in this paper are based on the unseen test data which account for 20% of the total data samples.

The deep neural network (DNN) that we employ to carry out the nonlinear regression has a topology inspired from generalized regression neural networks[52], 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 we observed them 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 [53].

Recently stacked ensemble regressors have gained much prominence and are consistently outperforming others in several datasets hosted at Kaggle[54]. 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. Success with DNN and ensemble techniques crucially depends on the number of training data and is sensitive to the hyper-parameters. As for the Gaussian Process Regression or kriging, the results are seen to be better than the above two when we use squared exponential kernel with prediction based on block coordinate descent. The GPR hyper-parameters are optimized using Bayesian optimization. But the method scales as $O(n^3)$ resulting in high memory requirements and training time for large data sets.

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 a very robust technique as it takes into account the covariance 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.

Among the various MLAs, we choose clustering-based prediction for the EEW pipeline 3. In addition to having the best prediction accuracy, it has the following advan-

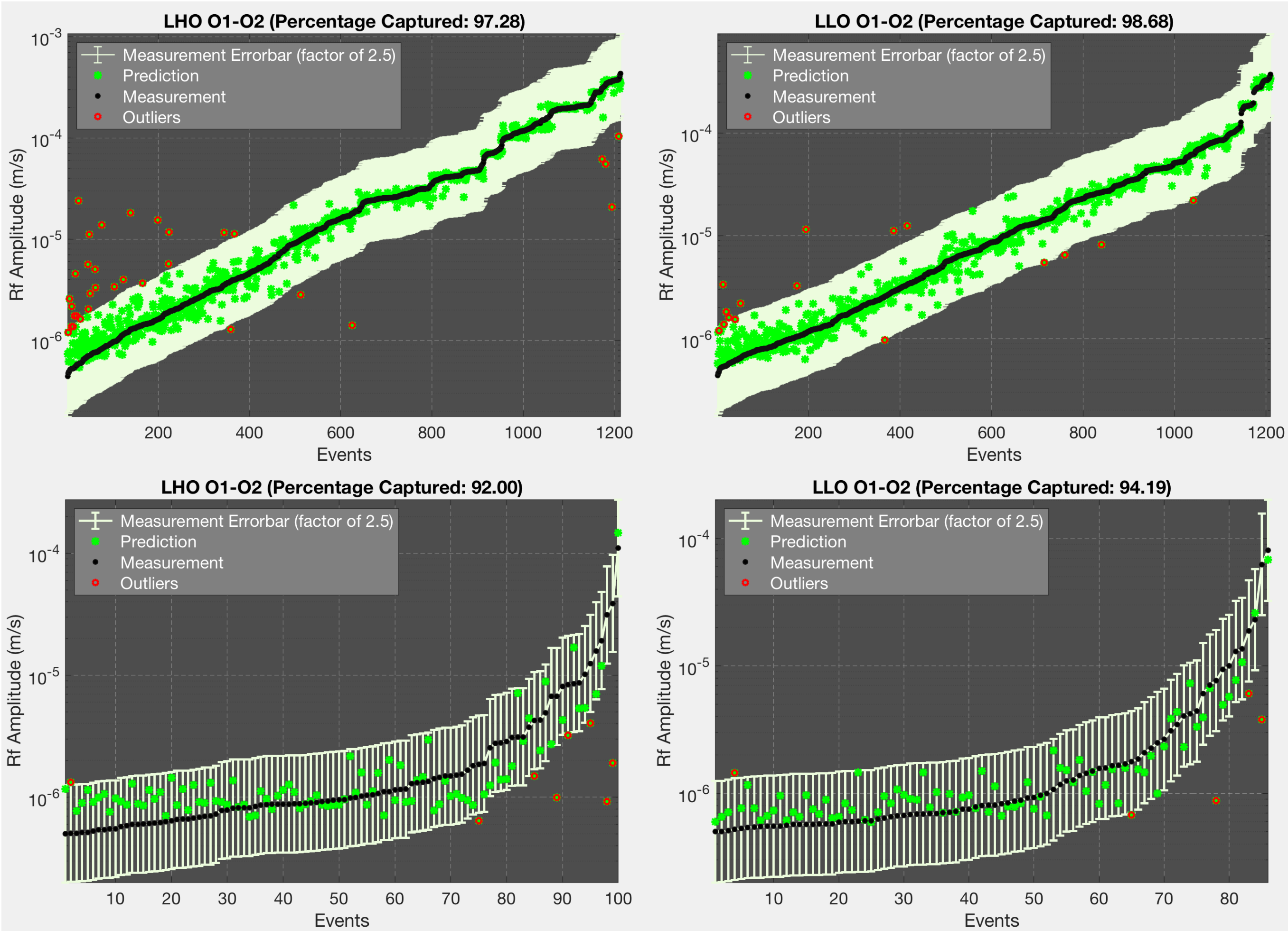

FIG. 4: Fit of peak velocities seen during O1-O2 at the interferometers (LHO, LLO) using Mahalanobis distance-based clustering. Results on simulated and real data are respectively shown in the top and bottom rows. The events have been ordered by their measured peak ground velocity (in grey), and yellow error bar corresponds to a factor of 2.5 within the predicted value. More than 90% of events are within a factor of 2.5 of the predicted value.

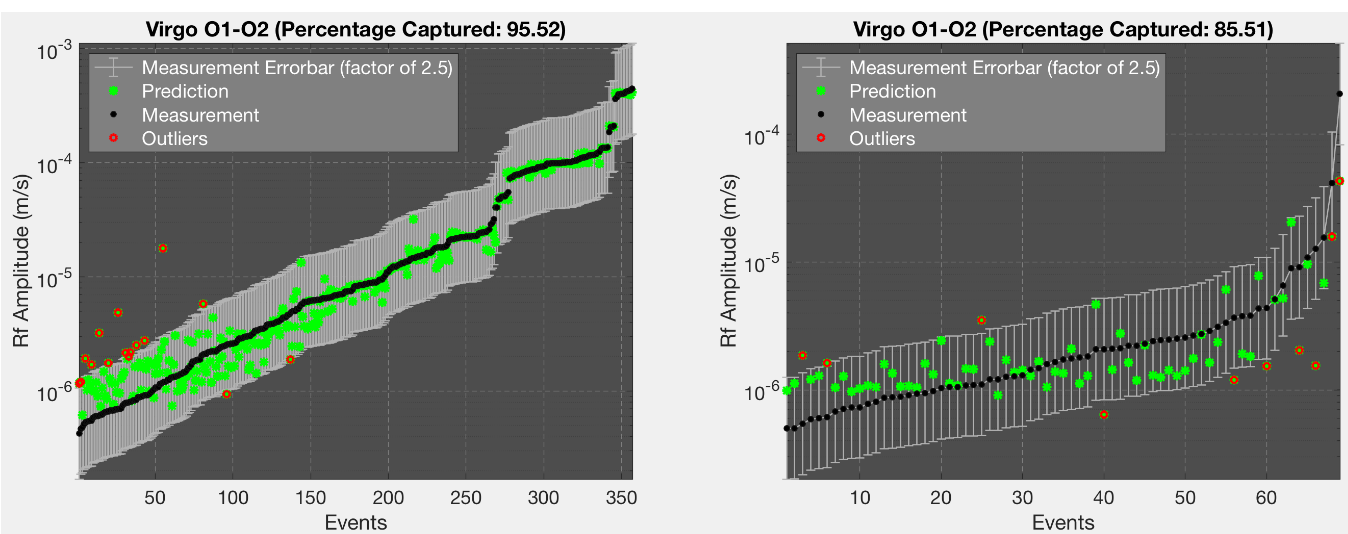

FIG. 5: Fit of peak velocities seen during O1-O2 at the Virgo Interferometer

tages. Firstly, as there is no training involved, the need for hyper-parameter tuning is eliminated. As we are con-

TABLE I: Rf amplitude prediction performance of different ML algorithms for real and simulated 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%

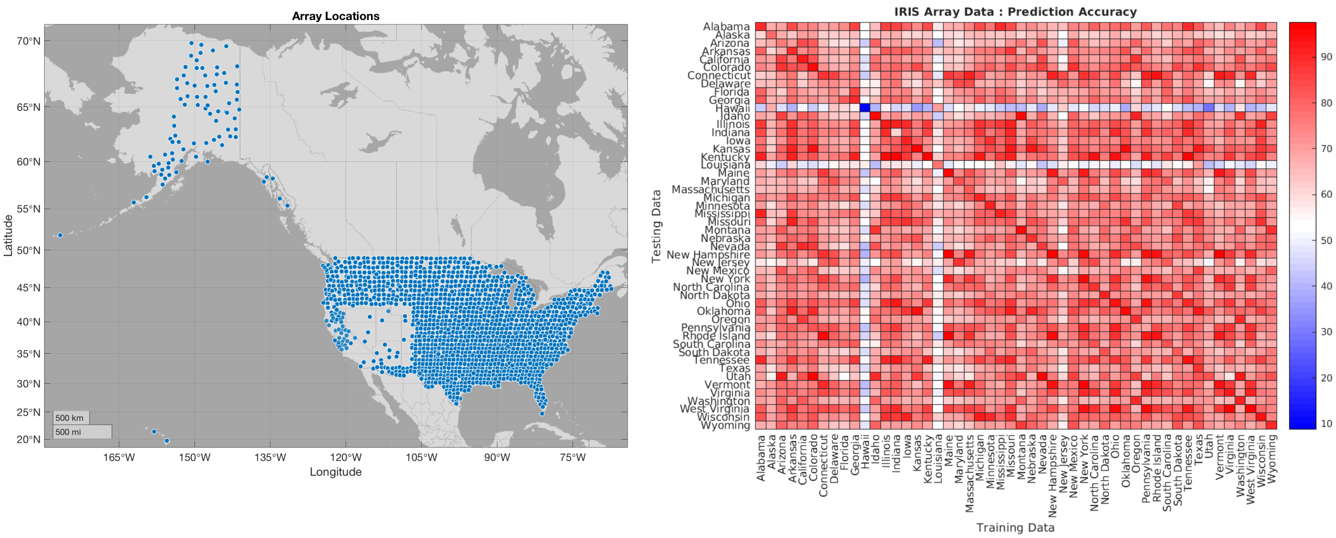

FIG. 6: Figure to the left shows the locations of seismometer array used to collect the earthquake data. Heatmap on the right shows the ground motion prediction accuracy within each state making use of archival data from the same as well as every other state.

stantly 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.

V. RESULTS

Figure 5 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 5 to 2.5 can be attributed to increased availability of data, the inclusion of more earthquake parameters and the usage of robust algorithms. As for the outliers, they do not 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 non-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. The possibility of further improvements using moment tensor solutions is left to future work.

We also demonstrate the usefulness 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 from the same state as well as the other states and compare the prediction accuracy as shown in Figure 6. For each row representing a state, we show how well we can make ground velocity predictions based on archival events from the state 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 teleseismic events could mean corresponding similarity within their geological properties such as shear velocity profiles, elasticity, and local soil density. Future state-wise seismic modeling could use this information to the 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 ex-

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

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

tended beyond the realm of gravitational-wave detector sites for hazard-based early warning alerts.

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. Figure 2 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. 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 %. 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.

VI. CONCLUSIONS

In conclusion, we have used MLAs to predict peak ground velocities from teleseismic 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 3. Alert system based on this scheme has been implemented at the GW observatories and would be used in the near future 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.

This paper has discussed the predictions of ground velocities from earthquakes and their effects on ground-based detectors. While these predictions are useful, they are only part of the story, as accurate earthquake arrival times are required for a useful response. For this reason, in the future, we will improve the existing infrastructure to use the machine learning algorithms described here to improve arrival time predictions. Further investigations into how to best broaden the data resources and link into the vast global network of seismic stations to improve ground velocity and arrival time predictions will be explored. In addition, we would explore the possibility of using historical seismometer data along with CMT parameters to predict radiation pattern 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. 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 Science Foundation under Cooperative Agreement EAR-126168

[1] J Aasi et al 2015 *Classical and Quantum Gravity* **32** 074001 URL http://stacks.iop.org/0264-9381/32/i=7/a=074001

[2] F Acernese et al 2015 *Classical and Quantum Gravity* **32** 024001 URL http://stacks.iop.org/0264-9381/32/i=2/a=024001

[3] Grote H for the LIGO Scientific Collaboration 2010 *Class. Quantum Grav.* **27** 084003

[4] Abbott, B P et al (LIGO Scientific Collaboration and Virgo Collaboration) 2016 *Phys. Rev. Lett.* **116**(6) 061102 URL http://link.aps.org/doi/10.1103/PhysRevLett.116.061102

[5] Abbott, B P et al (LIGO Scientific Collaboration and Virgo Collaboration) 2016 *Phys. Rev. Lett.* **116**(24) 241103 URL http://link.aps.org/doi/10.1103/PhysRevLett.116.241103

[6] Abbott, B P et al (LIGO Scientific and Virgo Collaboration) 2017 *Phys. Rev. Lett.* **118**(22) 221101 URL https://link.aps.org/doi/10.1103/PhysRevLett.118.221101

[7] Abbott, B P et al (LIGO Scientific Collaboration and Virgo Collaboration) 2017 *Phys. Rev. Lett.* **119**(14) 141101 URL https://link.aps.org/doi/10.1103/PhysRevLett.119.141101

[8] Abbott, B P et al (LIGO Scientific Collaboration and Virgo Collaboration) 2017 *Phys. Rev. Lett.* **119**(16) 161101 URL https://link.aps.org/doi/10.1103/PhysRevLett.119.161101

[9] The LIGO Scientific Collaboration and Virgo Collaboration 2016 *Classical and Quantum Gravity* **33** 134001 URL http://stacks.iop.org/0264-9381/33/i=13/a=134001

[10] 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 *The Leading Edge* **35** 590–593

[11] Venkateswara K, Hagedorn C A, Turner M D, Arp T and Gundlach J H 2014 *Review of Scientific Instruments* **85** 015005

[12] Coughlin M, Stubbs C, Barrientos S, Claver C, Harms J, Chris Smith R and Warner M 2015 *Experimental Astronomy* **39** 387–404 ISSN 0922-6435 URL http://dx.doi.org/10.1007/s10686-015-9454-1

[13] Coughlin M, Earle P, Harms J, Biscans S, Buchanan C, Coughlin E, Donovan F, Fee J, Gabbard H, Guy M, Mukund N and Perry M 2017 *Classical and Quantum Gravity* **34** 044004 URL http://stacks.iop.org/0264-9381/34/i=4/a=044004

[14] Abbott R, Adhikari R *et al.* 2002 *Classical and Quantum Gravity* **19** 1591 URL http://stacks.iop.org/0264-9381/19/i=7/a=349

[15] Stochino A, Abbot B, Aso Y, Barton M, Bertolini A, Boschi V, Coyne D, DeSalvo R, Galli C, Huang Y, Ivanov A, Marka S, Ottaway D, Sannibale V, Vanni C, Yamamoto H and Yoshida S 2009 *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **598** 737 – 753 ISSN 0168-9002

[16] Matichard F, Lantz B, Mittleman R, Mason K, Kissel J, Abbott B, Biscans S, McIver J, Abbott R, Abbott S, Allwine E, Barnum S, Birch J, Celerier C, Clark D, Coyne D, DeBra D, DeRosa R, Evans M, Foley S, Fritschel P, Giaime J A, Gray C, Grabeel G, Hanson J, Hardham C, Hillard M, Hua W, Kucharczyk C, Landry M, Roux A L, Lhuillier V, Macleod D, Macinnis M, Mitchell R, O'Reilly B, Ottaway D, Paris H, Pele A, Puma M, Radkins H, Ramet C, Robinson M, Ruet L, Sarin P, Shoemaker D, Stein A, Thomas J, Vargas M, Venkateswara K, Warner J and Wen S 2015 *Classical and Quantum Gravity* **32** 185003 URL http://stacks.iop.org/0264-9381/32/i=18/a=185003

[17] RM Allen 2012 *Science* **335** 297–298

[18] Kuyuk, HS and RM Allen 2013 *Geophysical Research Letters* **40**

[19] Kuyuk, HS and RM Allen 2013 *Seismological Research Letters* **84** 946–954

[20] Kuyuk, HS et al 2014 *Bulletin of Seismological Society of America* **104**

[21] Cochran E, Lawrence J, Christensen C, and Chung A 2009 *IEEE Inst and Meas* **12** 8–15

[22] Cochran E, Lawrence J, Christensen C, and Jakka R 2009 *Seismological Research Letters* **80** 26–30

[23] Bose M, Allen R, Brown H, Gua G, Fischer M, Hauksson E, Heaten T, Hellweg M, Liukis M, Neuhauser D, Maechling P, Solanki K, Vinci M, Henson I, Khainovski O, Kuyuk S, Carpio M, Meier M A and Jordan T 2014 CISN ShakeAlert: An Earthquake Early Warning Demonstration System for California *Early Warning for Geological Disasters* Advanced Technologies in Earth Sciences ed Wenzel F and Zschau J (Springer Berlin Heidelberg) pp 49–69 ISBN 978-3-642-12232-3 URL http://dx.doi.org/10.1007/978-3-642-12233-0_3

[24] Hoshiba M, Kamigaichi O, Saito M, Tsukada S and Hamada N 2008 *Eos, Transactions American Geophysical Union* **89** 73–74 ISSN 2324-9250 URL http://dx.doi.org/10.1029/2008EO080001

[25] Hoshiba M, Iwakiri K, Hayashimoto N, Shimoyama T, Hirano K, Yamada Y, Ishigaki Y and Kikuta H 2011 *Earth, Planets and Space* **63** 7 ISSN 1880-5981 URL http://dx.doi.org/10.5047/eps.2011.05.031

[26] Strauss J A and Allen R M 2016 *Seismological Research Letters* **87** 765–772 ISSN 0895-0695 (*Preprint* http://srl.geoscienceworld.org/content/87/3/765.full.pdf) URL http://srl.geoscienceworld.org/content/87/3/765

[27] Kuyuk H S, Colombelli S, Zollo A, Allen R M and Erdik M O 2015 *Geophysical Research Letters* **42** 5266–5273 ISSN 1944-8007 2015GL063881 URL http://dx.doi.org/10.1002/2015GL063881

[28] Kong Q, Allen R M, Schreier L and Kwon Y W 2016 *Science Advances* **2** (*Preprint* http://advances.sciencemag.org/content/2/2/e1501055.full.pdf) URL http://advances.sciencemag.org/content/2/2/e1501055

[29] Kuyuk H S, Yildirim E, Dogan E and Horasan G 2011 *Natural Hazards and Earth System Sciences* **11** 93–100

[30] Kortström J, Uski M and Tiira T 2016 *Computers & Geosciences* **87** 22 – 30 ISSN 0098-3004 URL http://www.sciencedirect.com/science/article/pii/S0098300415300832

[31] Perol T, Gharbi M and Denolle M 2017 *ArXiv e-prints* (*Preprint* 1702.02073)

[32] Mousavi S M, Horton S P, Langston C A and Samei B 2016 *Geophysical Journal International* **207** 29–46

[33] Jia Y and Ma J 2017 *GEOPHYSICS* **82** V163–V177

[34] Diersen S, Lee E J, Spears D, Chen P and Wang L 2011 *Procedia Computer Science* **4** 1572 – 1581 ISSN 1877-0509 proceedings of the International Conference on Computational Science, ICCS 2011 URL http://www.sciencedirect.com/science/article/pii/S1877050911002286

[35] Douglas J 2003 *Earth Science Reviews* **61** 43–104

[36] Melgar D, Crowell B W, Geng J, Allen R M, Bock Y, Riquelme S, Hill E M, Protti M and Ganas A 2015 *Geophysical Research Letters* **42** 5197–5205 ISSN 1944-8007 2015GL064278 URL http://dx.doi.org/10.1002/2015GL064278

[37] Biscans S, Warner J, Mittleman R, Buchanan C, Coughlin M W, Evans M, Gabbard H, Harms J, Lantz B, Mukund N, Pele A, Pezerat C, Picart P, Radkins H and Shaffer T 2018 *Classical and Quantum Gravity* URL http://iopscience.iop.org/10.1088/1361-6382/aaa4aa

[38] Bonnefoy-Claudet S, Cotton F and Bard P Y 2006 *Earth-Science Rev.* **79** 205

[39] Peterson J 1993 *Open-file report* **93-322**

[40] Berger J, Davis P and Ekström G 2004 *Journal of Geophysical Research* **109** B11307

[41] McNamara D E, Hutt C R, Gee L S, Benz H M and Buland R P 2009 *Seism. Res. Lett.* **80** 628

[42] Haubrich R A, Munk W H and Snodgrass F E 1963 *BSSA* **53** 27

[43] Toksöz M N and Lacoss R T 1968 *Science* **159** 872

[44] Cessaro R K 1994 *BSSA* **84** 142

[45] Friedrich A, Krüger F and Klinge K 1998 *J. Seism.* **2** 47

[46] Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, Devin M, Ghemawat S, Irving G, Isard M, Kudlur M, Levenberg J, Monga R, Moore S, Murray D G, Steiner B, Tucker P, Vasudevan V, Warden P, Wicke M, Yu Y and Zheng X 2016 Tensorflow: A system for large-scale machine learning *Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation* OSDI'16 (Berkeley, CA, USA: USENIX Association) pp 265–283 ISBN 978-1-931971-33-1 URL http://dl.acm.org/citation.cfm?id=3026877.3026899

[47] Wolpert D H 1992 *Neural networks* **5** 241–259

[48] Van der Laan M J, Polley E C and Hubbard A E 2007 *Statistical applications in genetics and molecular biology* **6**

[49] Rasmussen C E and Williams C K 2006 *The MIT Press, Cambridge, MA, USA* **38** 715–719

[50] Mahalanobis P C 1936 On the generalised distance in statistics *Proceedings of the National Institute of Sciences of India.* UAI'95 p 49–55

[51] Chawla N V, Bowyer K W, Hall L O and Kegelmeyer W P 2011 *ArXiv e-prints* (*Preprint* 1106.1813)

[52] Specht D F 1991 *IEEE transactions on neural networks* **2** 568–576

[53] Kingma D P and Ba J 2014 *arXiv preprint arXiv:1412.6980*

[54] Kaggle website https://www.kaggle.com accessed: 2018-06-15