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

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 and Fred Donovan

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

Christopher Buchanan

Department of Physics and Astronomy,

Louisiana State University, Baton Rouge, LA 70803-4001, 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

Jan Harms

INFN, Sezione di Firenze, Sesto Fiorentino, 50019, Italy

Università degli Studi di Urbino "Carlo Bo", I-61029 Urbino, Italy

Matthew Perry

Planetary Science Institute, Lakewood, CO 80401, USA

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 [? ? ? ? ? ? ? ? ? ?]. 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 3 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 rapid magnitude and location estimates to the amplitudes observed. We find agreement within a factor of 3 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) [?], Virgo [?], and GEO600 [?] detectors. Any increase in duty cycle increases the sensitivity of gravitational-wave searches, including the observations of binary black hole mergers [? ?]. One source of ground motion that destabilizes the detectors are earthquakes [? ?], despite seismic isolation systems designed to minimize such effects [? ? ?].

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 [?]. The short warning times arise out of the physical processes that drive the earthquake rupture, where warning is given by seismometers measuring P-waves ($\approx 8\,\mathrm{km/s}$) and S-waves ($\approx 4\,\mathrm{km/s}$). 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 [?]. 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 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 [?]. These source-scaling laws are applied to early portions of seismograms to make predictions about the magnitude for EEW [?], important for hypocenter and magnitude computations in tsunamis [?], hazard computations in engineering seismology [?], and computation of the elastic response spectrum [?].

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 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. [?] 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. Real-time GPS seismic waveforms can be used to rapidly determine magnitude, typically within the first minute of rupture initiation and in many cases before the rupture is complete.

In previous work, [?] used advances in early earthquake warning to develop a low-latency earthquake early warning client named *Seismon*. 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. It was shown that about 90% of events had a measured ground velocity within a factor of 5 of the predicted value.

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 [?]. Machine learning algorithms are also used to differentiate earthquakes from other seismic transients [? ? ?]. In addition, they have been used to discriminate between deep and shallow microearthquakes [?]. It can also be used to add to undersampled or missing traces [?]. In addition, they have been used to make full-wave tomography images [?].

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. 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. 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 period.

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, 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 a machine learning algorithm. In particular, we use the two different machine learning approaches to to make the predictions: Tensorflow implementation of deep neural networks (DNN) and stacked ensemble classifier. The inputs to the algorithm are the earthquake magnitude, latitude, longitude, distance, depth, and azimuth. 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

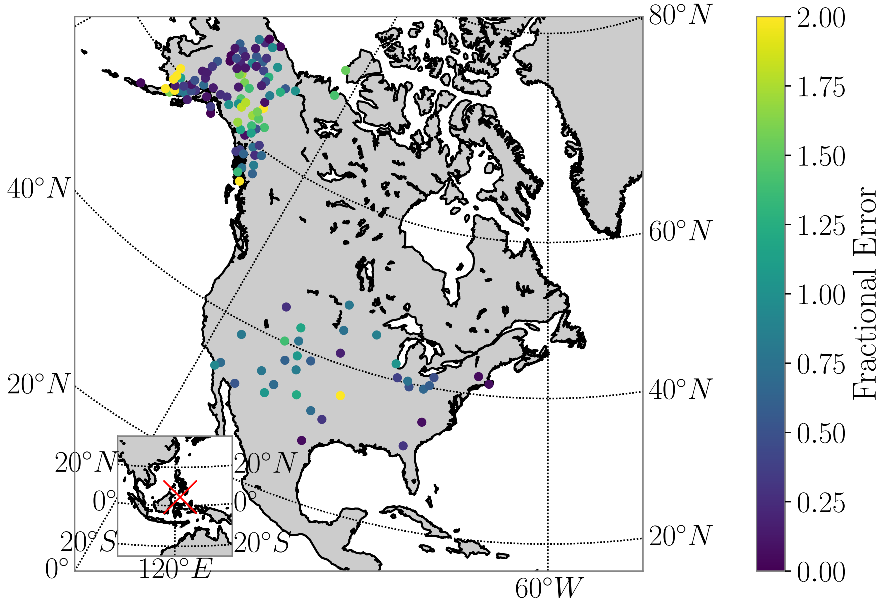

FIG. 1: The plot shows the location of the (?) seismic stations with available data from IRIS in the United States during the (?) earthquake. The color indicates the fractional error between the measured peak ground velocity and that predicted by the algorithm. The inset shows the location of the earthquake.

functional form. Second, it trivially includes more parameters, such as latitude, longitude, and earthquake azimuth relative to the detector. We observe a performance improvement in prediction accuracy from a factor of five to three for both the implementations. Inclusion of moment tensor parameters further improve the performance by a factor of (?)

We can use these parameter fits that go into the ground velocity prediction model to make predictions across the United States. Figure 1 shows the location of seismic stations that had data during the (?) earthquake. The colors indicate the fractional difference between the measured and predicted peak ground velocities.

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/or 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-1261681.

Methods.

Gravitational-wave detectors.

The Advanced LIGO [**?**] and the Advanced Virgo [**?**] detectors are multi-kilometer Michelson-based interferometers. 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.

In order 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 are physical environment 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 [**?**]. These are useful for measuring any source of ground motion that can couple into the interferometers. Gravitational-wave detectors have driven the development of both seismic [**?**] and rotation [**?**] sensors.

Earthquakes are one of the main sources of transient seismic motion for these gravitational-wave detectors. 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.

Seismic data.

We use multiple sources of seismic data. We perform an analysis of seismic timeseries that were made available through IRIS, covering the last 10 years, and LIGO, covering the last 2 years. These stations have time-series with response between 10 mHz to 10 Hz. The noise for these instruments are determined by a variety of sources including anthropogenic and atmospheric disturbances, earthquakes and ocean waves [**?**]. Seismic noise models are developed using global seismometer arrays [**? ? ?**]. 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 dominatex seismic ground spectra everywhere on Earth [**? ? ? ?**].

The peak ground velocities are calculated as follows. 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 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 so as to cover the range of earthquake magnitudes likely to significantly effect the gravitational-wave detectors.

Data Augmentation. To improve the learning and prevent early stopping, we augment the 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 ML algorithm to better learn and generalize to the the underlying smooth nonlinear function. New samples are generated from each of the original dataset by creating a gaussian jitter distribution centred around the parameter value followed by random draw of samples from these distributions.

Ground velocity predictions.

We use the earthquake parameters to make peak ground velocity predictions. We use historical earthquakes to train our machine learning algorithms. The parameters that enter

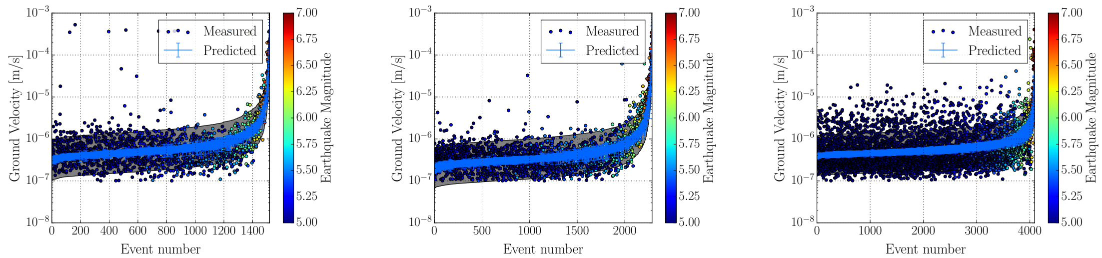

FIG. 2: Fit of peak velocities seen during O1-O2 at the interferometers (LHO, LLO, and Virgo) using Gaussian Process Regression. The events have been ordered by their measured peak ground velocity (in blue) and grey error bar corresponds to a factor of 3 within the predicted value. About 90% of events are within a factor of 3 of the predicted value.

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, 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}. We take the peak ground velocities measured for each earthquake and use a machine learning algorithm to make predictions of that velocity.

MLA description

In a previous work, [**?**] 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. In this work, we employ a deep neural networks (DNN) to carry out the nonlinear regression. The network with topology is inspired from generalized regression neural networks but we back-propagate the errors and update the weights by training it through several epochs. Results of applying the trained network to unseen data (Fig. 3)show the improvement in prediction accuracy as compared to traditional regression techniques. The model accurately predicts 90% of the observational data within a factor of 3. DNNs require larger data sets to learn the underlying function without overfitting the data. As the observed data could possibly 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 is able to 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. First level consists of 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(()superLearner 2007).

TABLE I: Rf amplitude prediction performance of different ML algorithms in different scenarios

	Deep Neural Nets	Stacked Ensemble	Model 3	Model 4
LIGO Livingston	89% 85%	93% 89%	−− −−	−− −−
LIGO Hanford	86% 84%	91% 88%	−− −−	−− −−

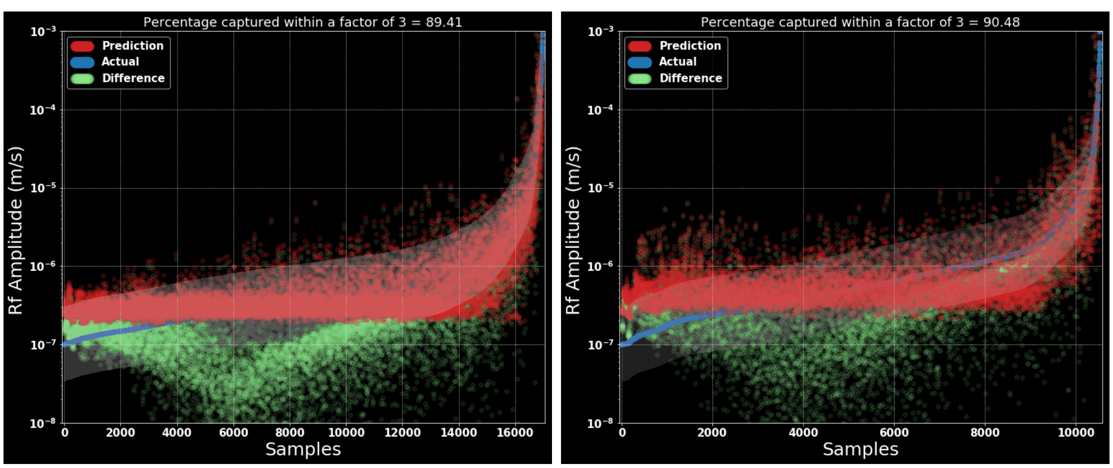

FIG. 3