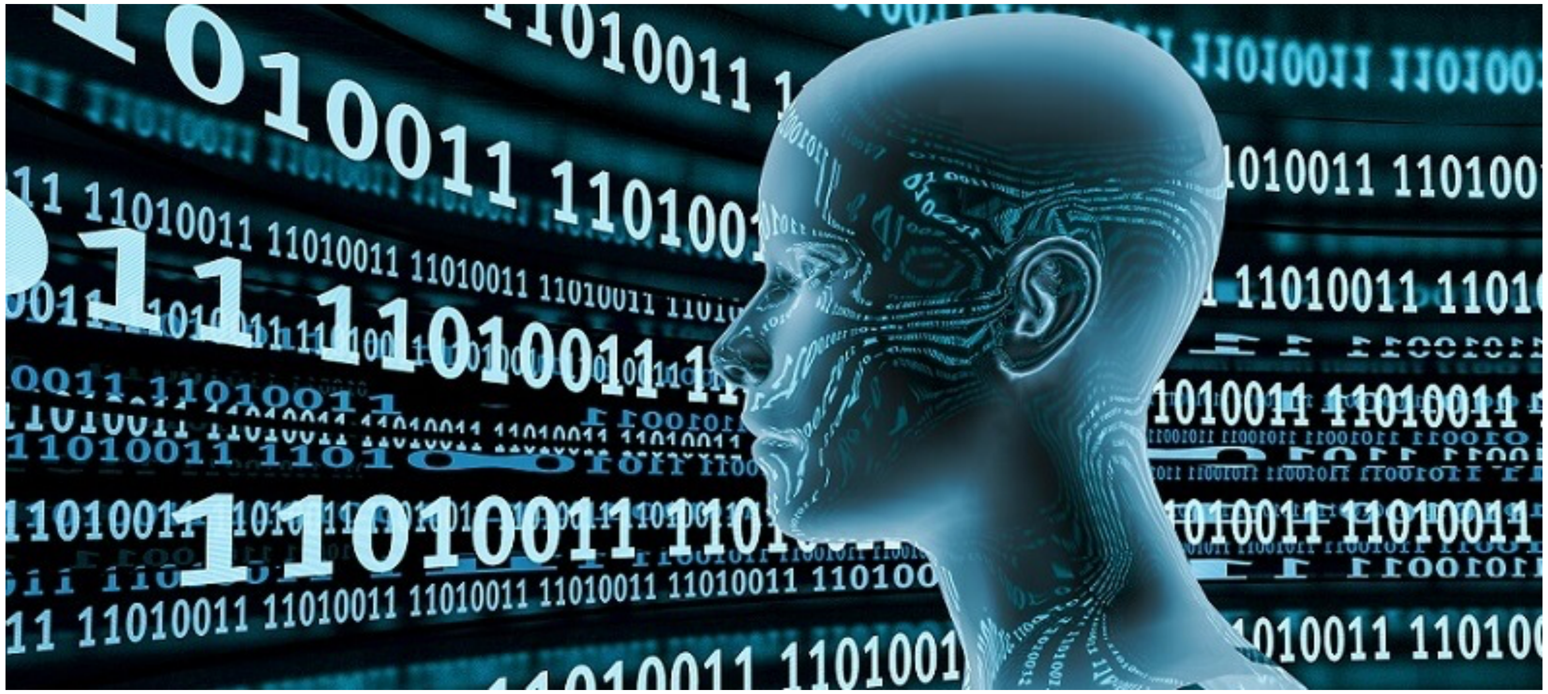
Machine Learning in Seismology

Alberto Michelini
with
Carlo Giunchi, Emanuele Casarotti, Dario Jozinovic

Istituto Nazionale di Geofisica e Vulcanologia

What is Machine Learning ?



- Machine learning is an application of artificial intelligence (AI) that provides systems the ability to *automatically learn* and *improve from experience* without being explicitly programmed.
- Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.
- The process of learning begins with *observations* or *data*, such as examples, direct experience, or instruction, in order to look for *patterns in data* and make better decisions in the future based on the examples that we provide.
- Data Mining: Tools for extracting *unknown* patterns or information from large data sets

From: https://www.expertsystem.com/machine-learning-definition/

Outline

- Important topics in seismology
- General problem
- Topics addressed through ML
- Activities at INGV

NB The presentation is based primarily on the recent review article by *Kong, Trugman, Ross, Bianco, Meade, Gerstoft (2018). Machine Learning in Seismology: Turning Data into Insights, SRL, DOI: https://doi.org/10.1785/0220180259*

Important topics in seismology

- Detection (*earthquakes and other phenomena generating seismic waves*)
- Picking of seismic phases (*seismic location using standard methodologies, tomography*)
- Earthquake Early Warning
- Inversion/tomography of seismic data for the Earth's interior
- Estimation/prediction of ground shaking
- Massive seismic waveform data sets mining, clustering and dimensionality reduction
- …

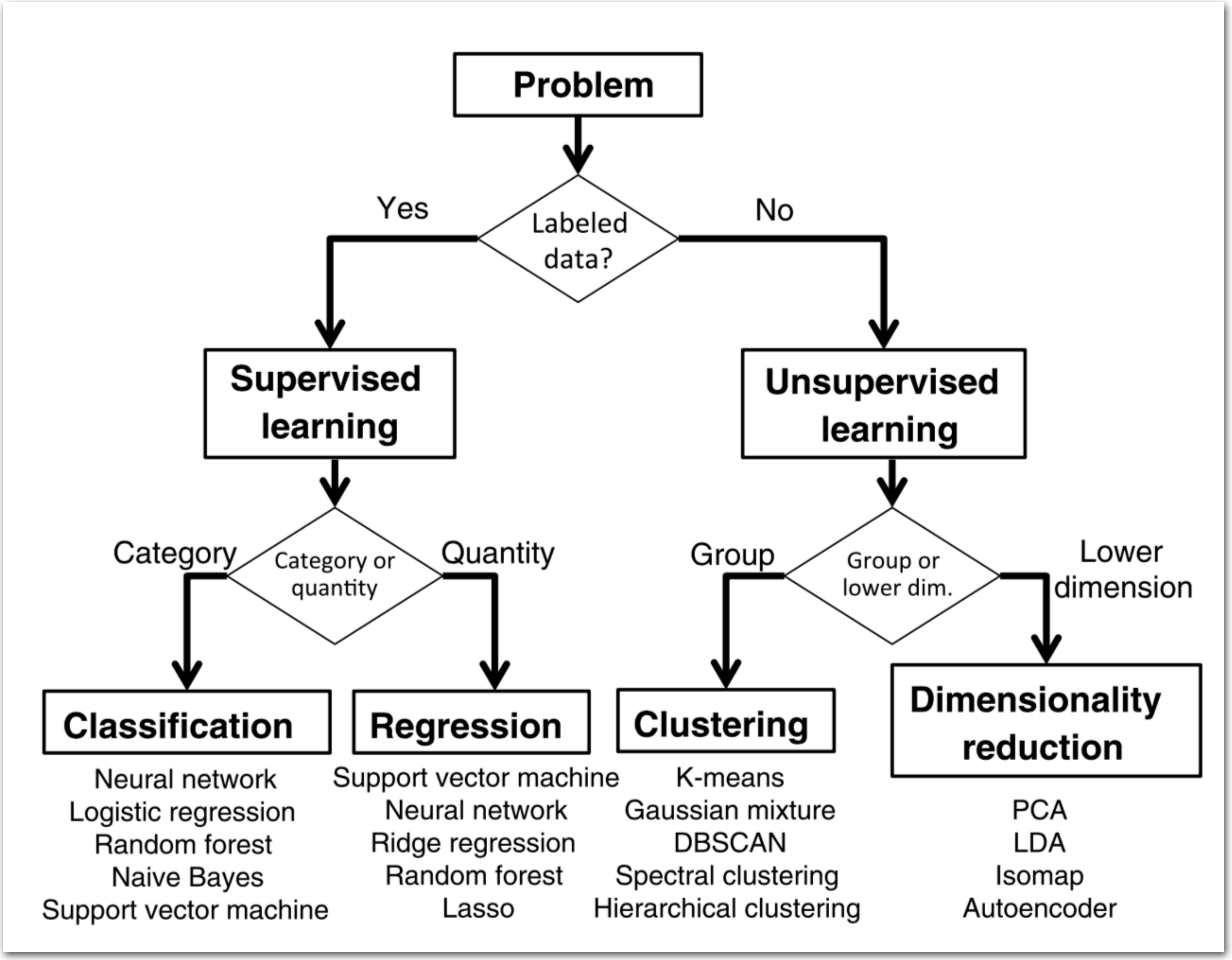
What seismology can provide

- **Raw data** (observed seismic waveforms, GNSS high accuracy positioning data)
- **Databases** of information resulting from the analysis of raw data (i.e., **labeled data**)
  - *earthquakes <u>direct measurements</u> (phases, ground motion amplitudes and durations)*
  - *earthquakes <u>"inverse modeling"</u> (location, magnitude(s), moment tensor, fault plane solutions, small number of finite fault solutions)*

INGV Data

- **Raw seismic data** (~150 TB)
- **Databases** of information resulting from the analysis of raw data (i.e., **labeled data**)
  - *~92,000 Earthquakes M≥2*
  - *~2M between P and S phases*
  - *>25,000 strong ground motion measurements M≥4*
  - *~4,000 ShakeMaps*
  - *~650 moment tensors in Italy (M≥3.6)*

ML general



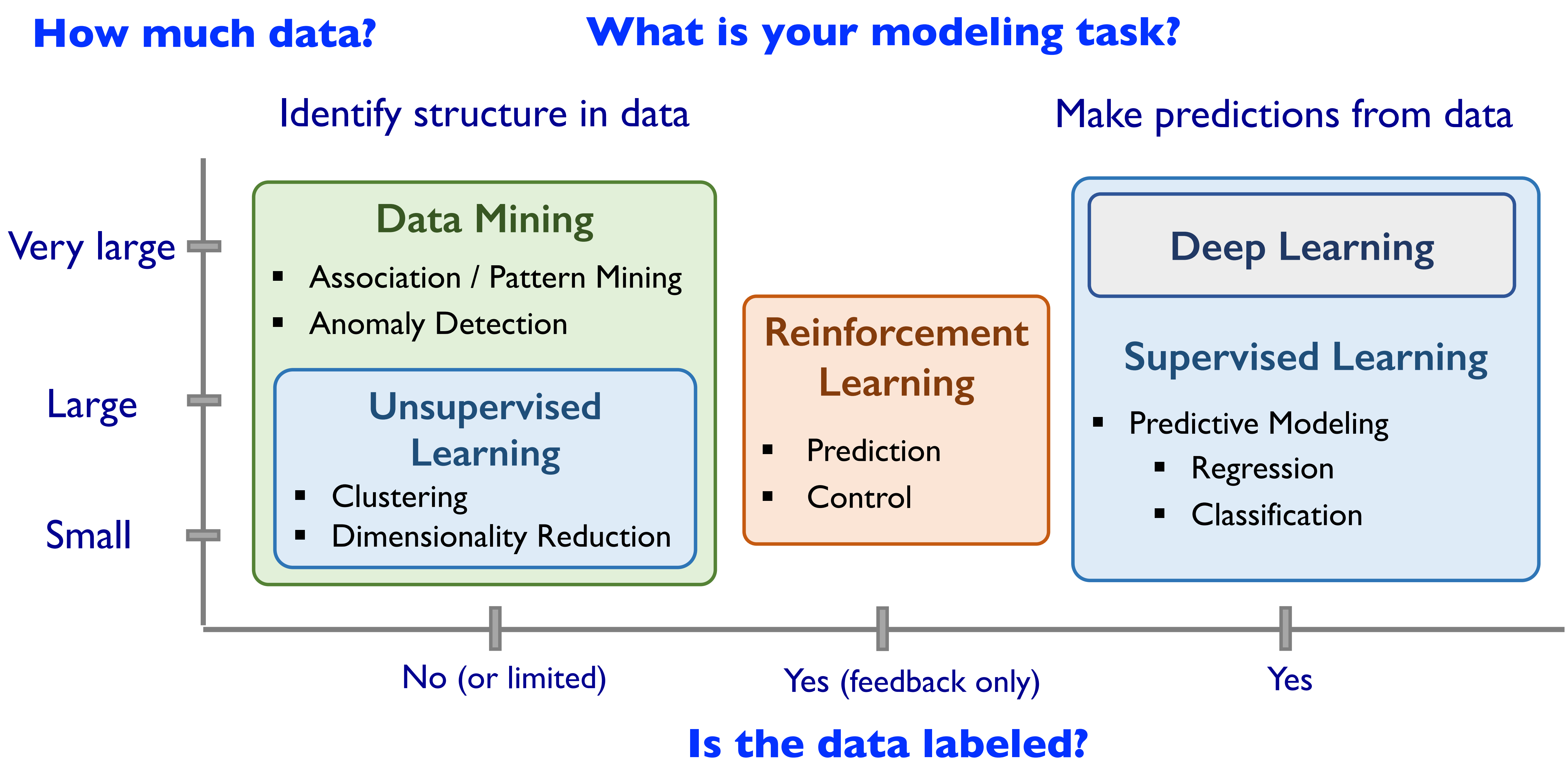
Kong et al. (2018), Machine Learning in Seismology: Turning Data into Insights, Seismological Research Letters, 90(1), 3–14, doi:10.1785/0220180259.

Selecting a machine learning approach

- How much data ?
- Is the data labeled ?
- What is the modeling goal ?
  - Identify structure in the data

    or
  - Make predictions

Selecting a machine learning approach

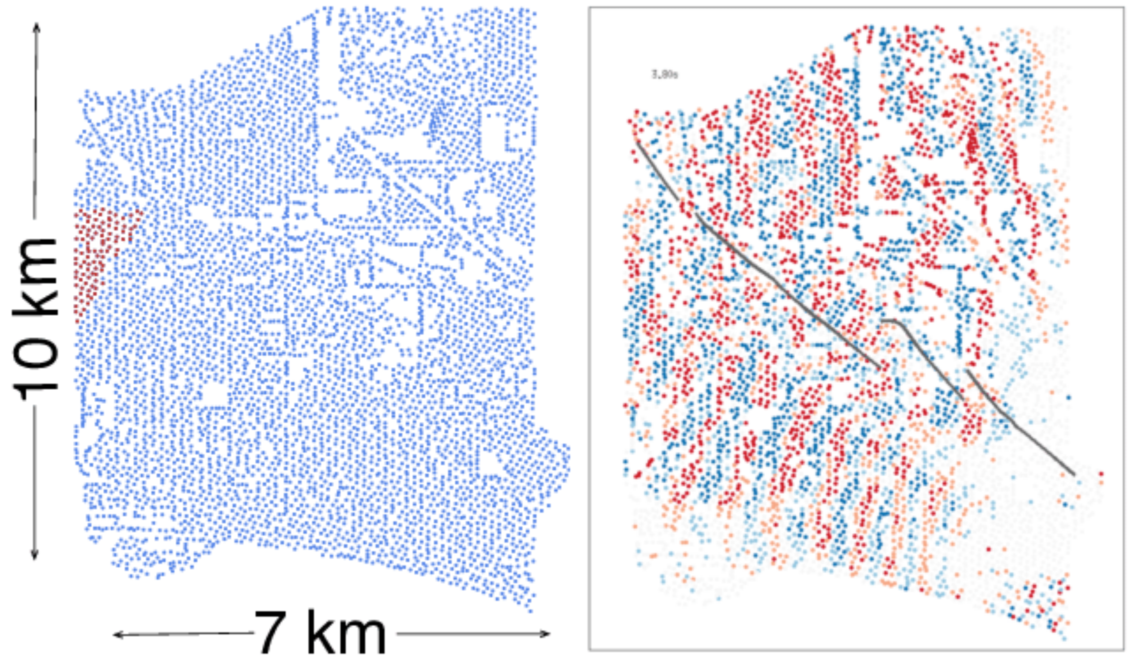


from Bergen (2018) *"Improving earthquake detection with data mining & machine learning "* presented at "IRIS 2018 workshop" held in Albuquerque, USA (June 12-14, 2018)

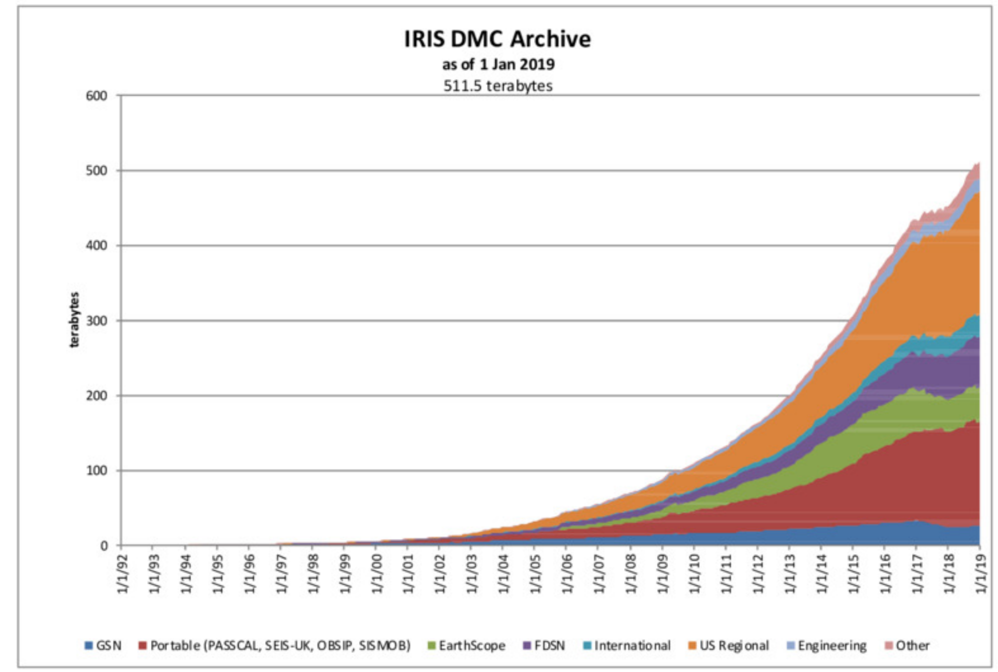
Why seismology can benefit from ML ?

- ✓ Massive seismic data sets
- ✓ New ML algorithms and models
- ✓ Improved technology both HW and SW

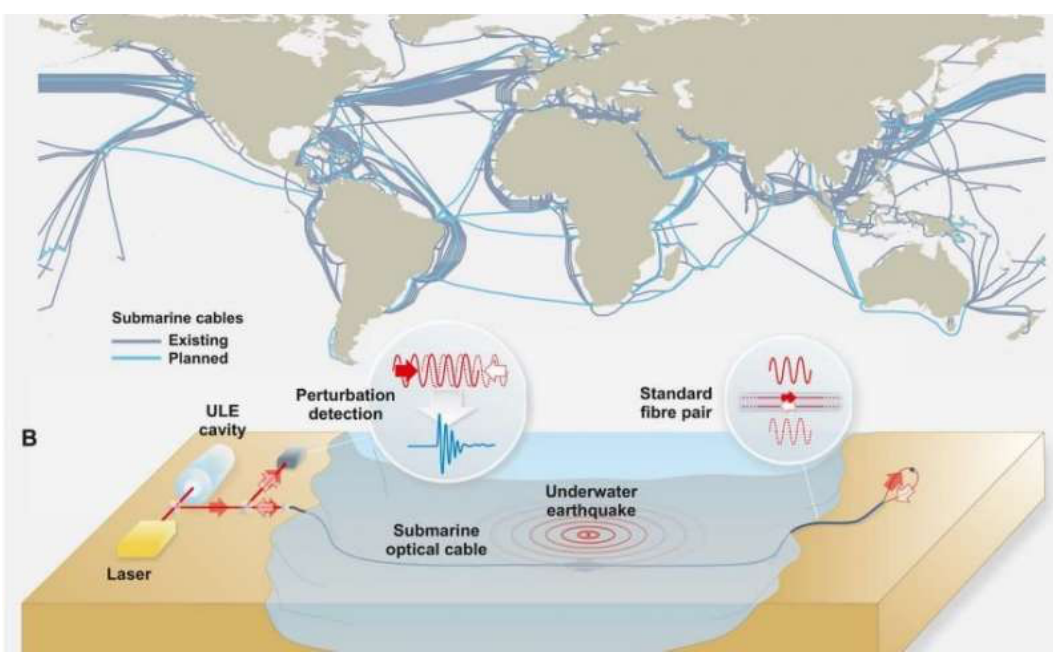
Massive seismic data sets



Li, Z., Z. Peng, X. Meng, A. Inbal, Y. Xie, D. Hollis, and J.-P. Ampuero (2015), Matched Filter Detection of Microseismicity in Long Beach with a 5200-station Dense Array, DOI: 10.1190/segam2015-5924260.1.



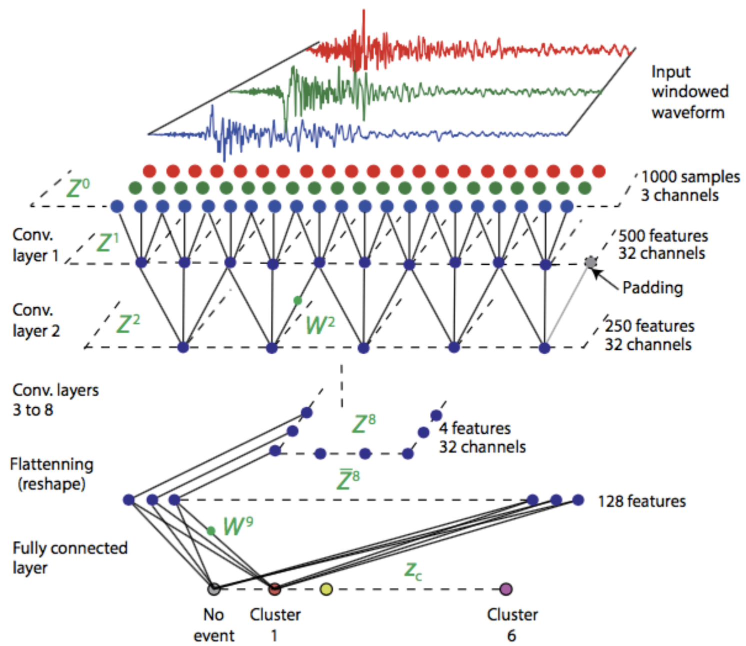
https://ds.iris.edu/files/stats/data/archive/Archive_Growth.jpg



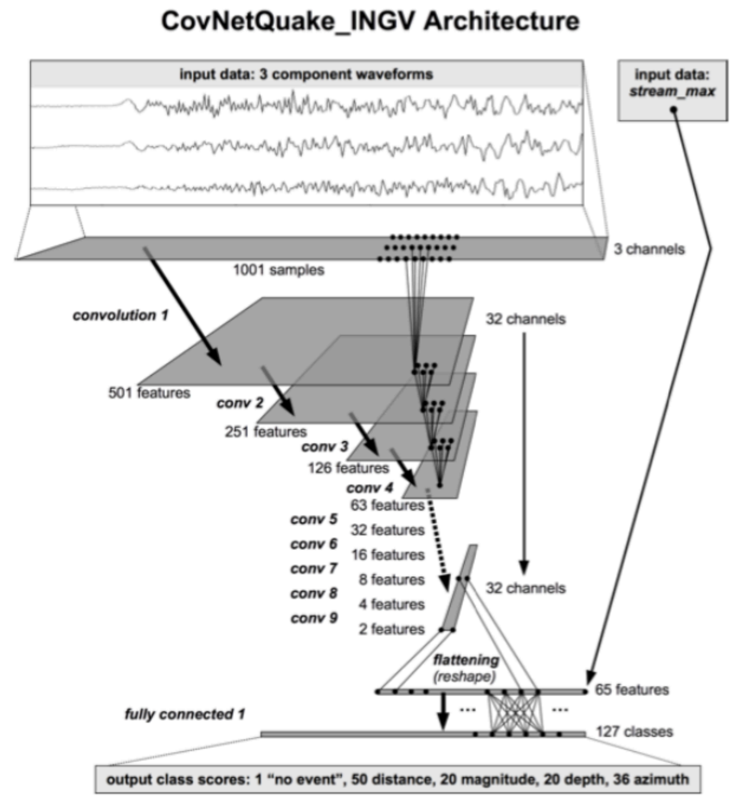
From Marra et al. (2018) Ultrastable laser interferometry for earthquake detection with terrestrial and submarine cables; Science 03 Aug 2018: Vol. 361, Issue 6401, pp. 486-490 DOI: 10.1126/science.aat4458

https://myshake.berkeley.edu

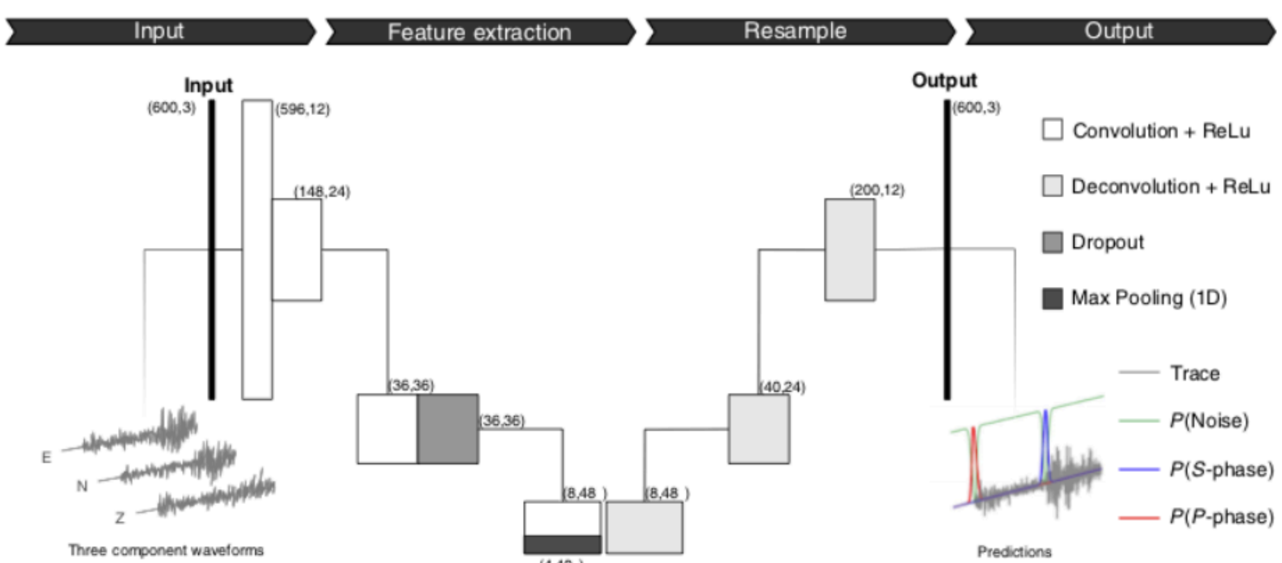
New ML algorithms and models



Perol, T., et al. (2018), Convolutional neural network for earthquake detection and location, Science Advances, 4(2), e1700578, doi: 10.1126/sciadv.1700578.

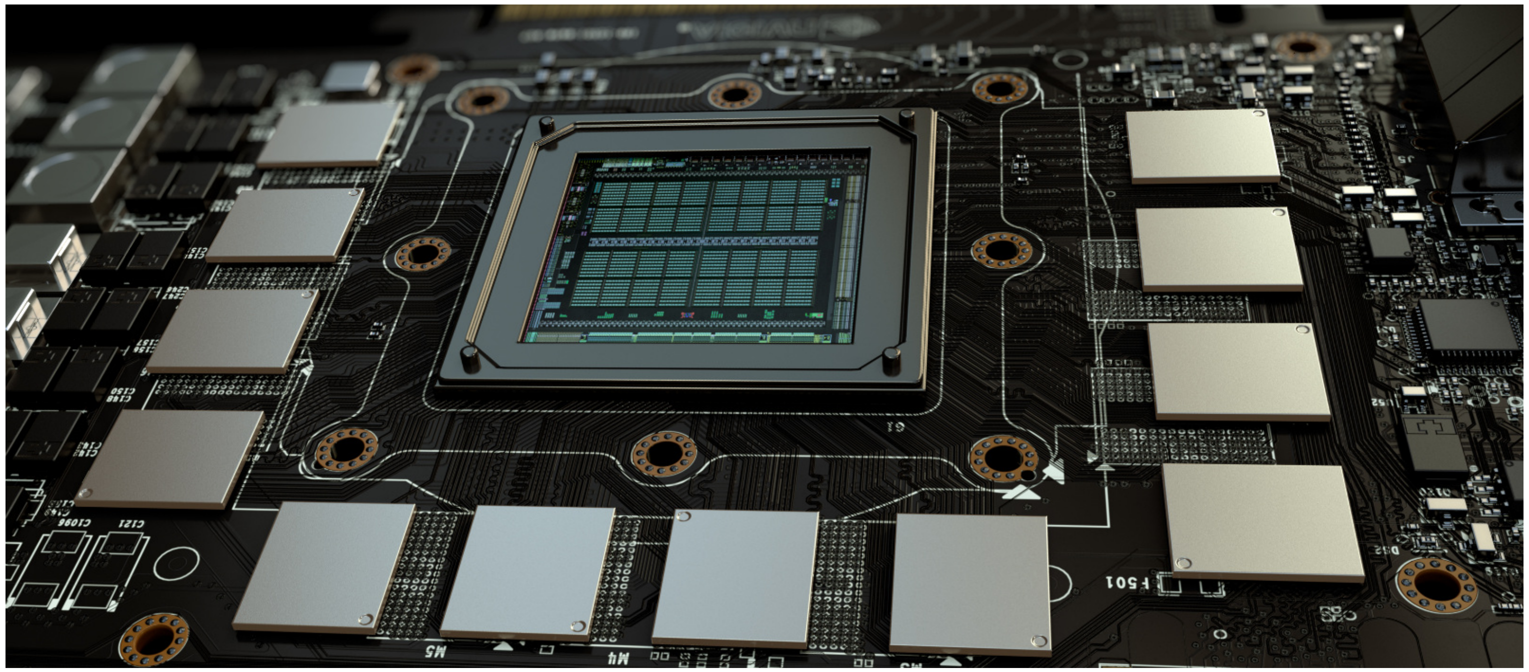


Lomax, et al. (2019). An investigation of rapid earthquake characterization using single-station waveforms and a convolutional neural network, in press in SRL



From Woollam et al. (2019). Convolutional Neural Network for Seismic Phase Classification, Performance Demonstration over a Local Seismic Network, in press SRL.

Improved technology both HW and SW



GPU hard-ware Technology

Open Source software ML Technology

Caffe2

theano

Microsoft | Cognitive Toolkit

Keras

scikit learn

Accord framework .NET v3.8

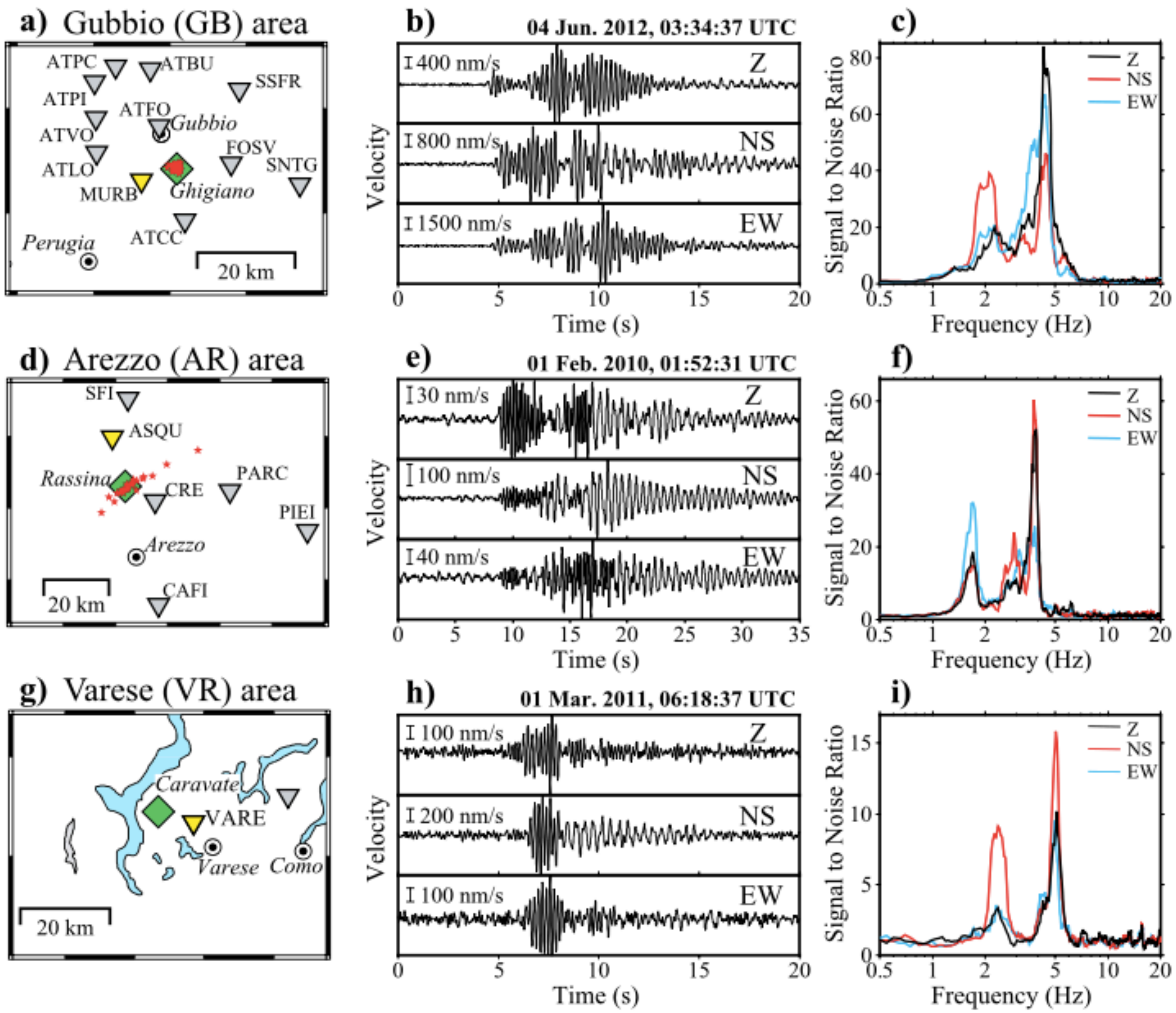
PyTorch

https://opensource.com/article/18/5/top-8-open-source-ai-technologies-machine-learning

General problem in seismology

- Extracting information from continuous waveforms
  - ➡ detect, locate and estimate the size of earthquakes
  - ➡ detect and locate and estimate the size of other phenomena (not earthquakes in the proper sense)
    - internal to the Earth (e.g., tremors, slow earthquakes, other phenomena induced by fast enough relative movements within the Earth)
    - outside the solid Earth (e.g., from the atmosphere)
    - anthropogenic

Examples of anthropogenic signal



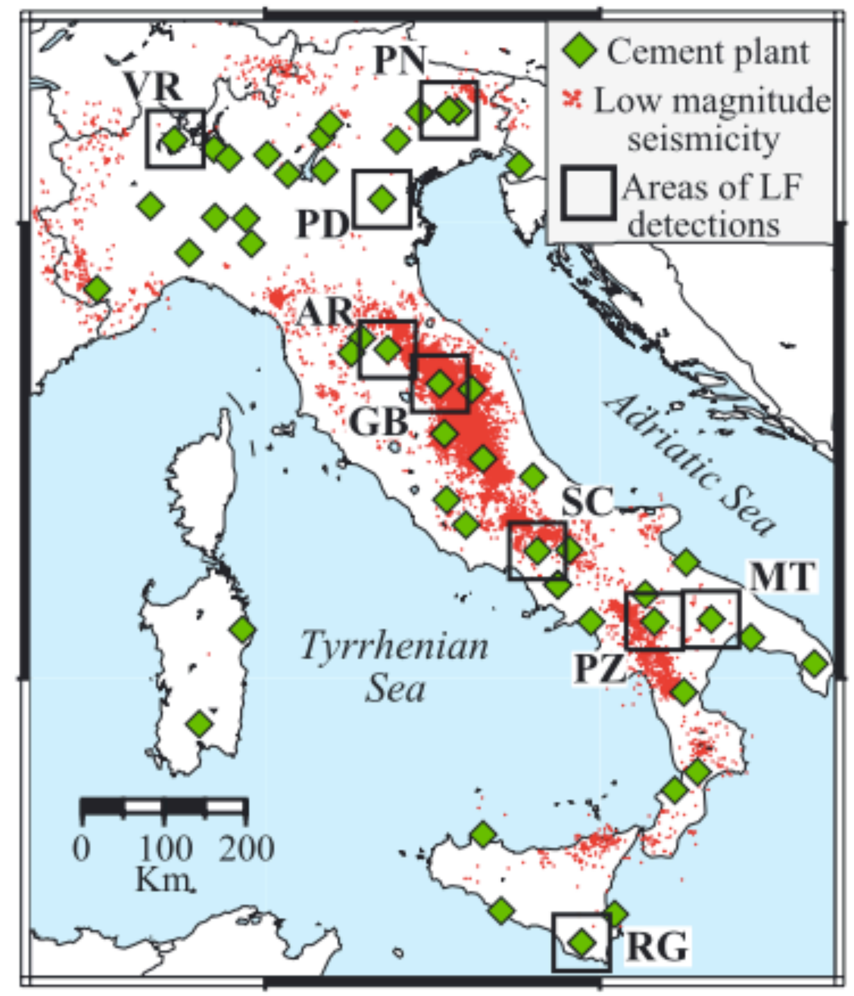


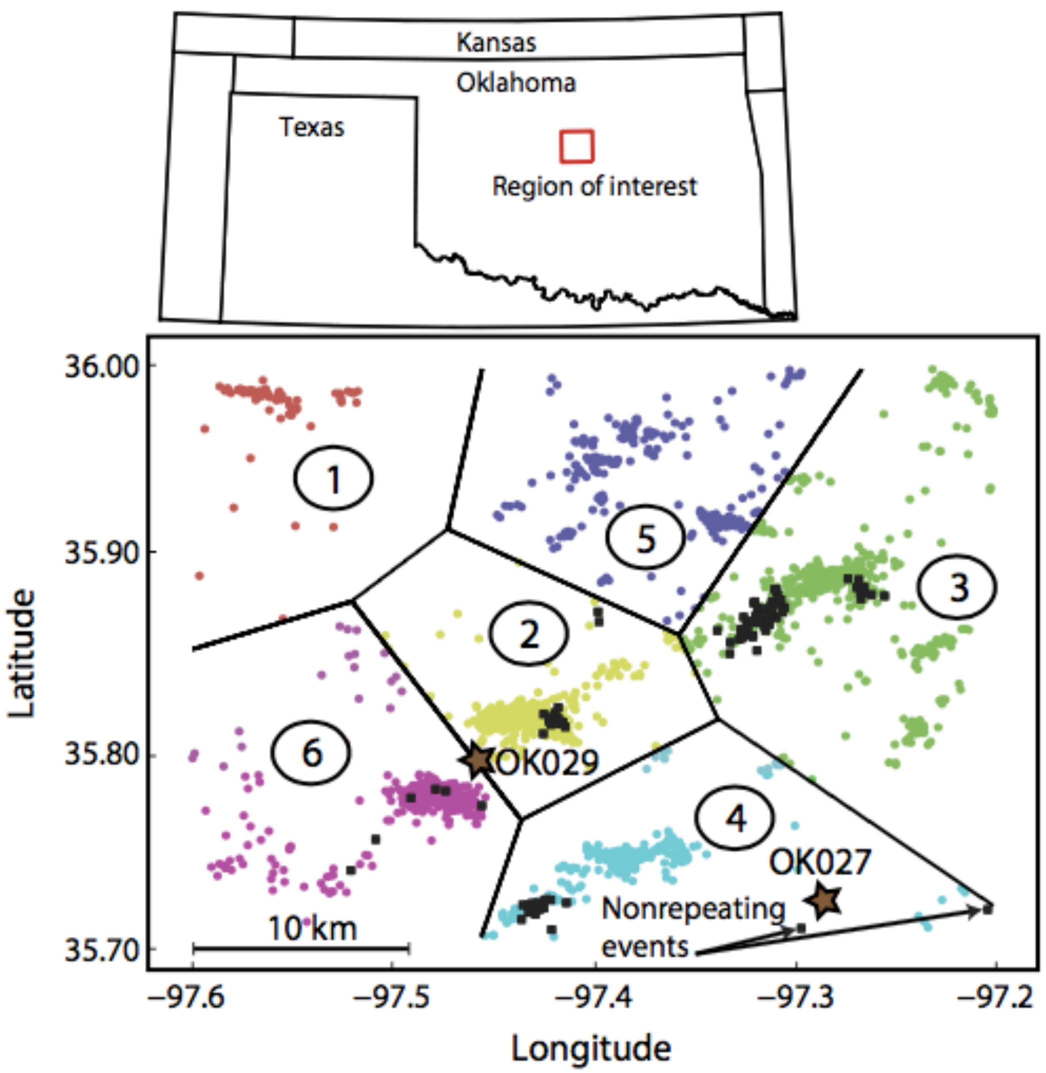
Figure 1. Epicentral map of earthquakes with magnitude lower than 1.0, recorded in Italy by the Istituto Nazionale di Geofisica e Vulcanologia (INGV) network from 2005 to date (red dots, 26,460 events). Black boxes indicate the location of the regions where we detected low-frequency (LF) events from 2010 to 2012: Varese (VR), Pordenone (PN), Padova (PD), Arezzo (AR), Gubbio (GB), Sesto Campano (SC), Potenza (PZ), Matera (MT), and Ragusa (RG). Green diamonds are the location of the 57 full-cycle cement factories active in Italy in 2010 [source AITEC, Associazione Italiana Tecnico Economica Cemento, http://www.aitecweb.com, last accessed August 2014].

Latorre et al. (2014), Man-induced low-frequency seismic events in Italy, Geophys. Res. Lett., doi:10.1002/(ISSN)1944-8007.

Earthquake Detection and Phase Picking

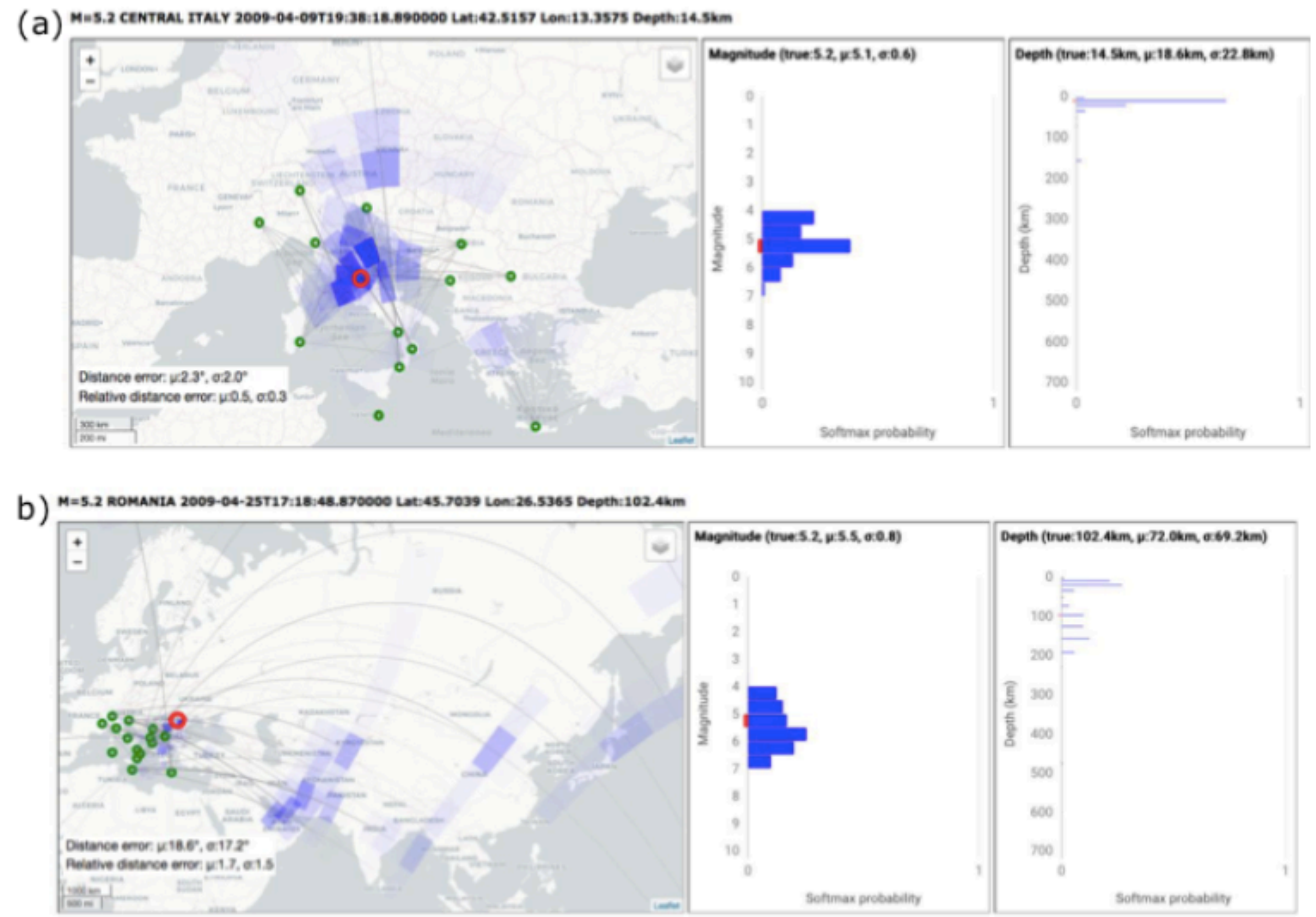
- Earthquake detection and identification of source region without picking
  - Convolutional Neural Network to detect and locate single-station (Perol et al., 2018, Lomax et al., 2018)
  - Convolutional Neural Network to locate clusters using multi-station (Kriegerowski et al., 2019)
  - FAST, Fingerprinting and Similarity Thresholding (FAST) algorithm (Yoon et al., 2015; Bergen and Beroza, 2018) is a data mining approach that converts an entire continuous waveform dataset into a database of binary fingerprints (*NOT ML*).
- Phase picking capabilities
  - early work on P- and S-wave picking using neural networks (e.g., Chiaruttini et al., 1989; Mousset et al., 1996; Dai and MacBeth, 1995, 1997; Zhao and Takano, 1999; Enescu, 1996; Wang and Teng, 1995; 1997; Gentili & Michelini, 2006)
  - recent work with
    - Generalized Phase Detection (Ross et al., 2018a,b)
    - PhaseNet (Zhu and Beroza, 2018)
    - Seismic Event and Phase Detection Using Time–Frequency Representation and CNN (Dokht et al., 2019).
    - CNN for phase identification (Wollam et al., 2019)

Convolutional Neural Network to detect and locate without any phase picking (only single station waveforms)



Detection and Classification to areas

Perol, T., et al. (2018), Convolutional neural network for earthquake detection and location, Science Advances, 4(2), e1700578, doi:10.1126/sciadv.1700578.



Detection
Classification according to:

- **distance**
- **azimuth**
- **magnitude**
- **depth**

Lomax, et al. (2019). An investigation of rapid earthquake characterization using single-station waveforms and a convolutional neural network, in press in SRL

Phase picking capabilities: Very early work -> 1989

228 *C. Chiaruttini* et al.

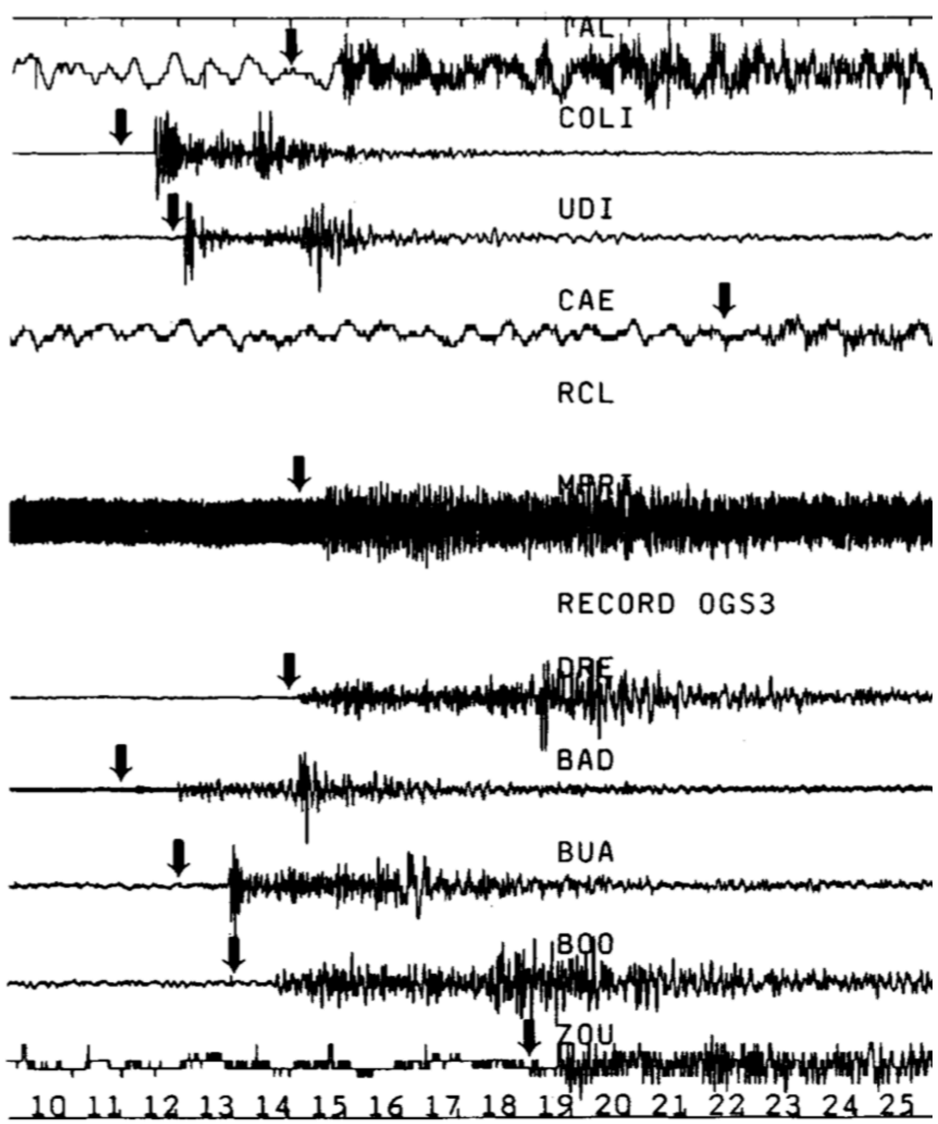
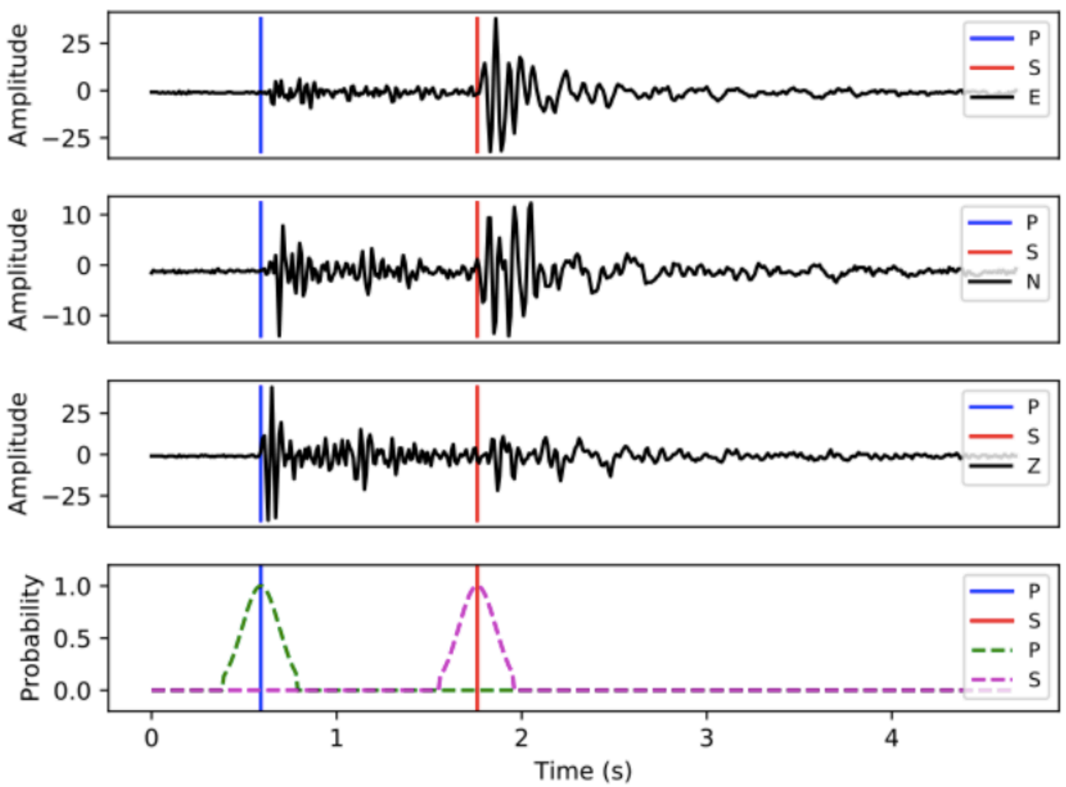


Figure 3. A clear local event recorded by the North-Eastern Italy Seismomet
noise are not shown; the black vertical arrows indicate the arrival times of *Pg*

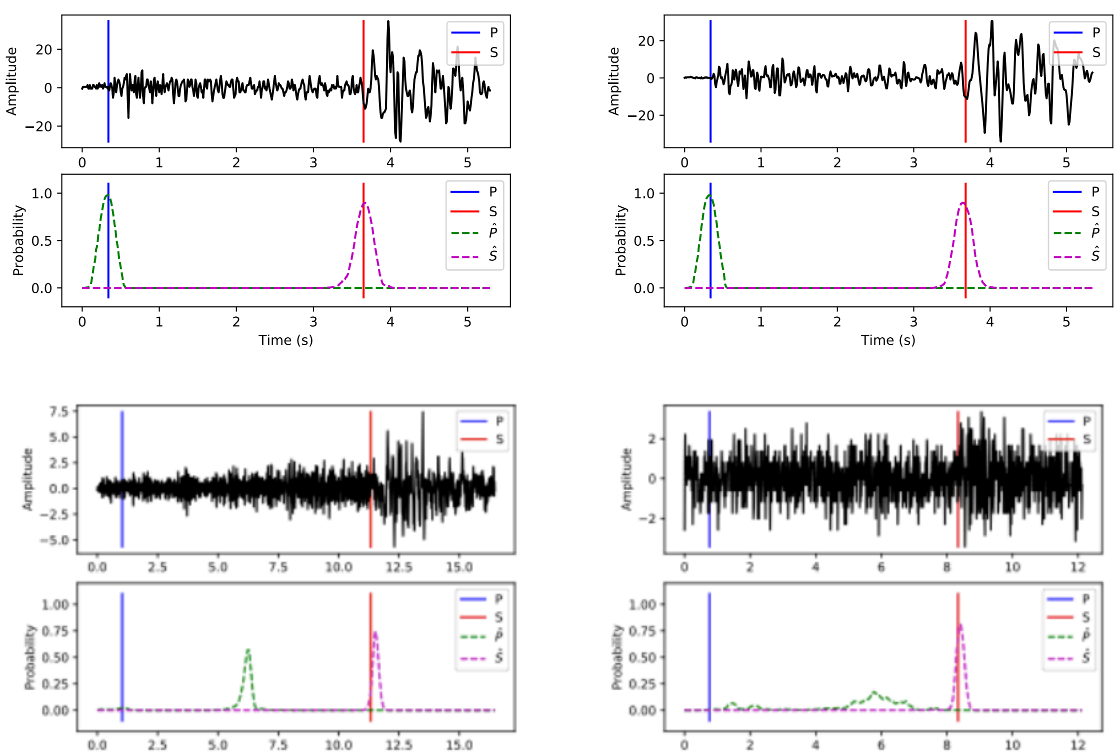
Chiaruttini, C., V. Roberto, and F. Saitta (1989), Artificial intelligence techniques in seismic signal interpretation, *Geophys. J. Int*, 223–232.

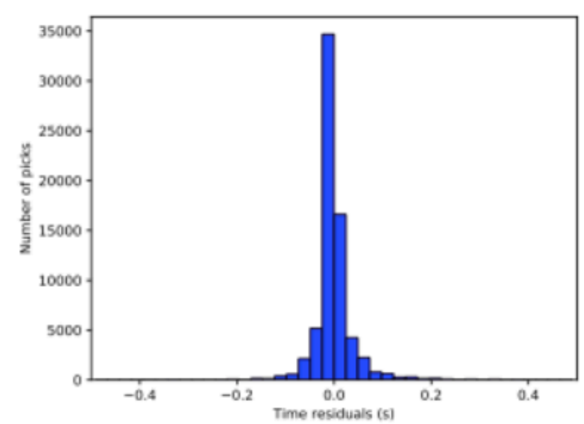
Phase picking capabilities: PhaseNet



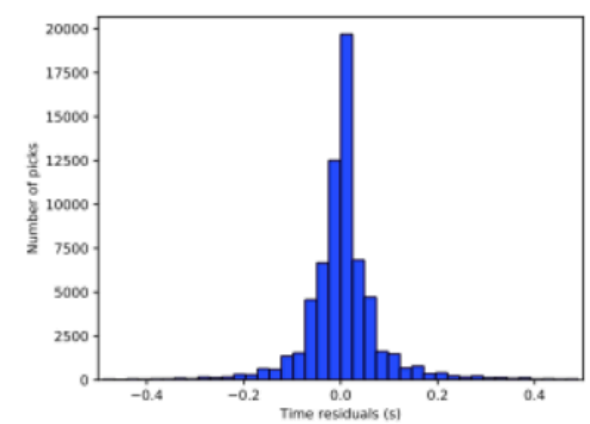
PhaseNet is trained on the prodigious available data set provided by analyst-labelled P and S arrival times from the Northern California Earthquake Data Center. The data set we use contains more than 700,000 waveform samples extracted from over 30 yr of earthquake recordings.

…PhaseNet achieves much higher picking accuracy and recall rate than existing methods when applied to the waveforms of known earthquakes, which has the potential to increase the number of S-wave observations dramatically over what is currently available. This will enable both improved locations and improved shear wave velocity models.

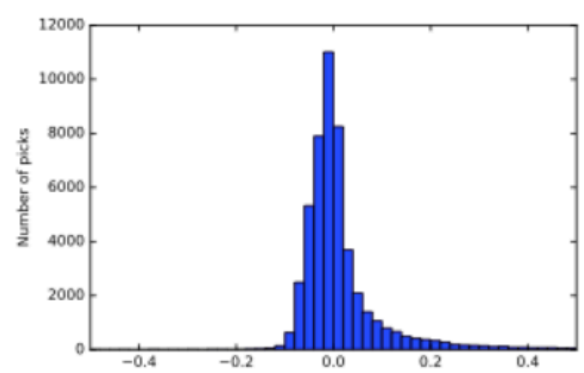




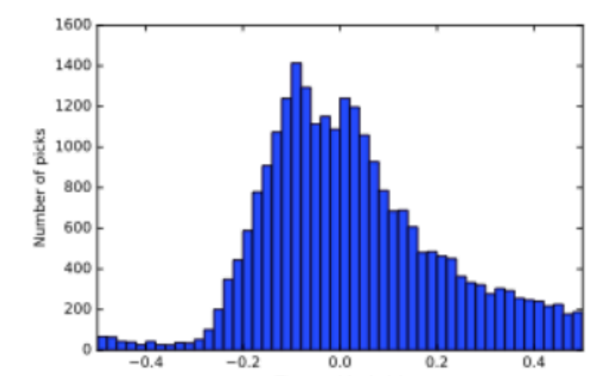
(a) P picks of PhaseNet



(b) S picks of PhaseNet



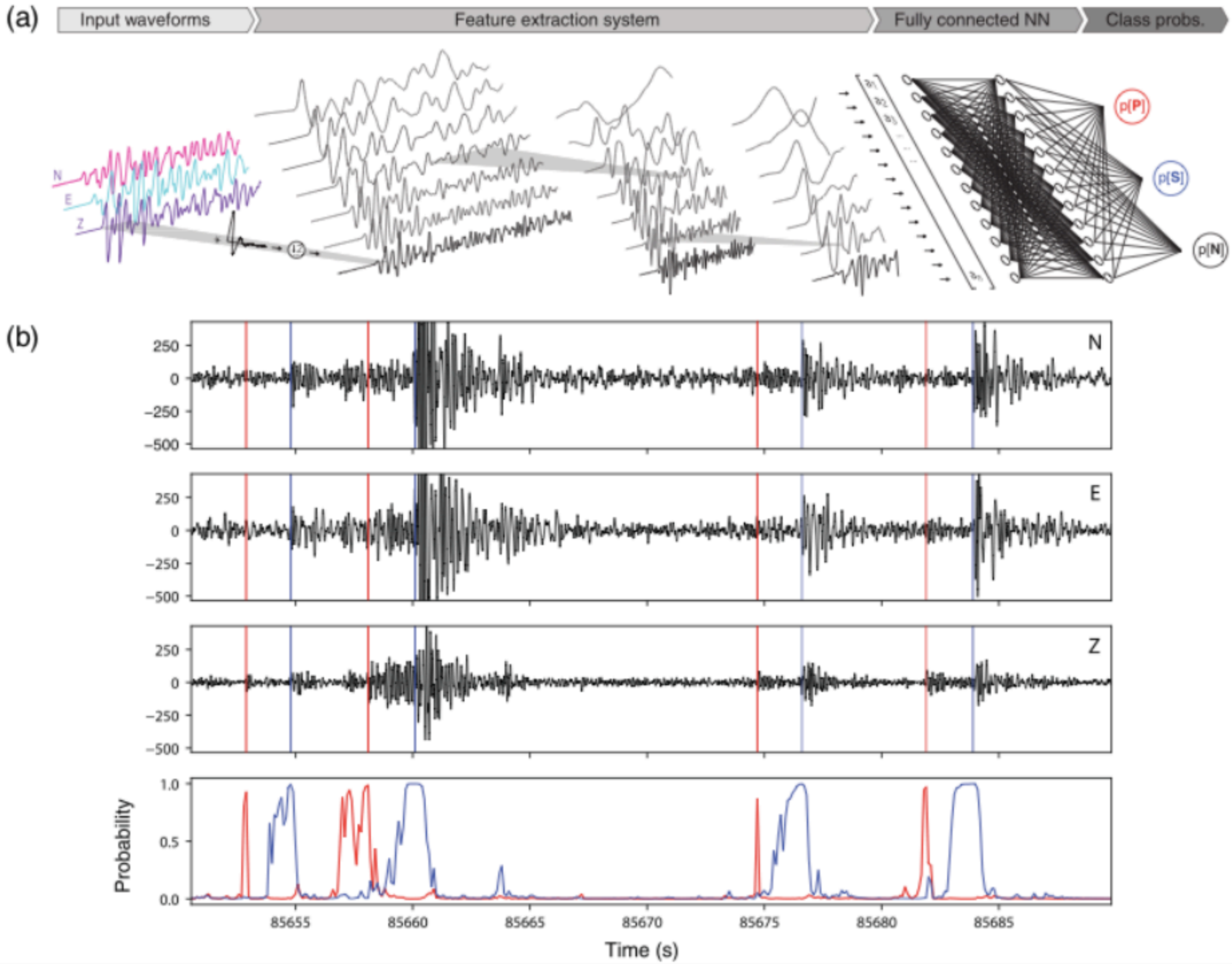
(c) P picks of AR picker



(d) S picks of AR picker

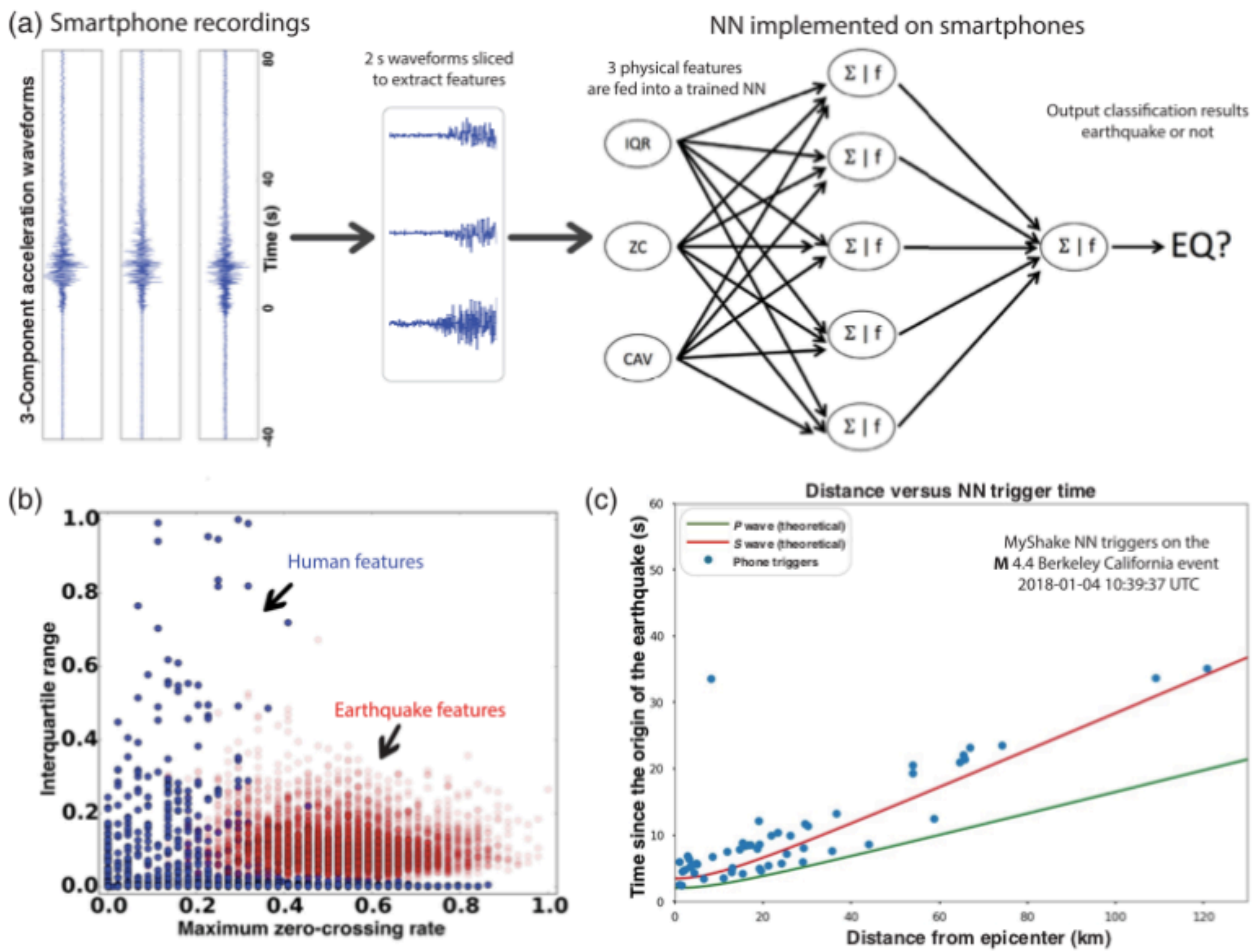
Zhu and Beroza (2019). PhaseNet: A Deep-Neural-Network-Based Seismic Arrival Time Picking Method. *Geophys. J. Int.* 216, 261-273.

Phase picking capabilities: Generalized Phase Detection



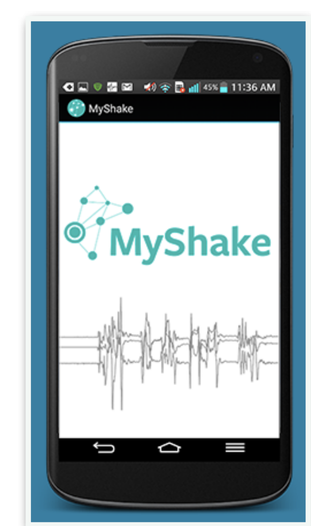
Ross, Z. E., M.-A. Meier, E. Hauksson, and T. H. Heaton (2018). Generalized seismic phase detection with deep learning, Bull. Seismol. Soc. Am. doi: 10.1785/0120180080.

EEW and Real-Time ML

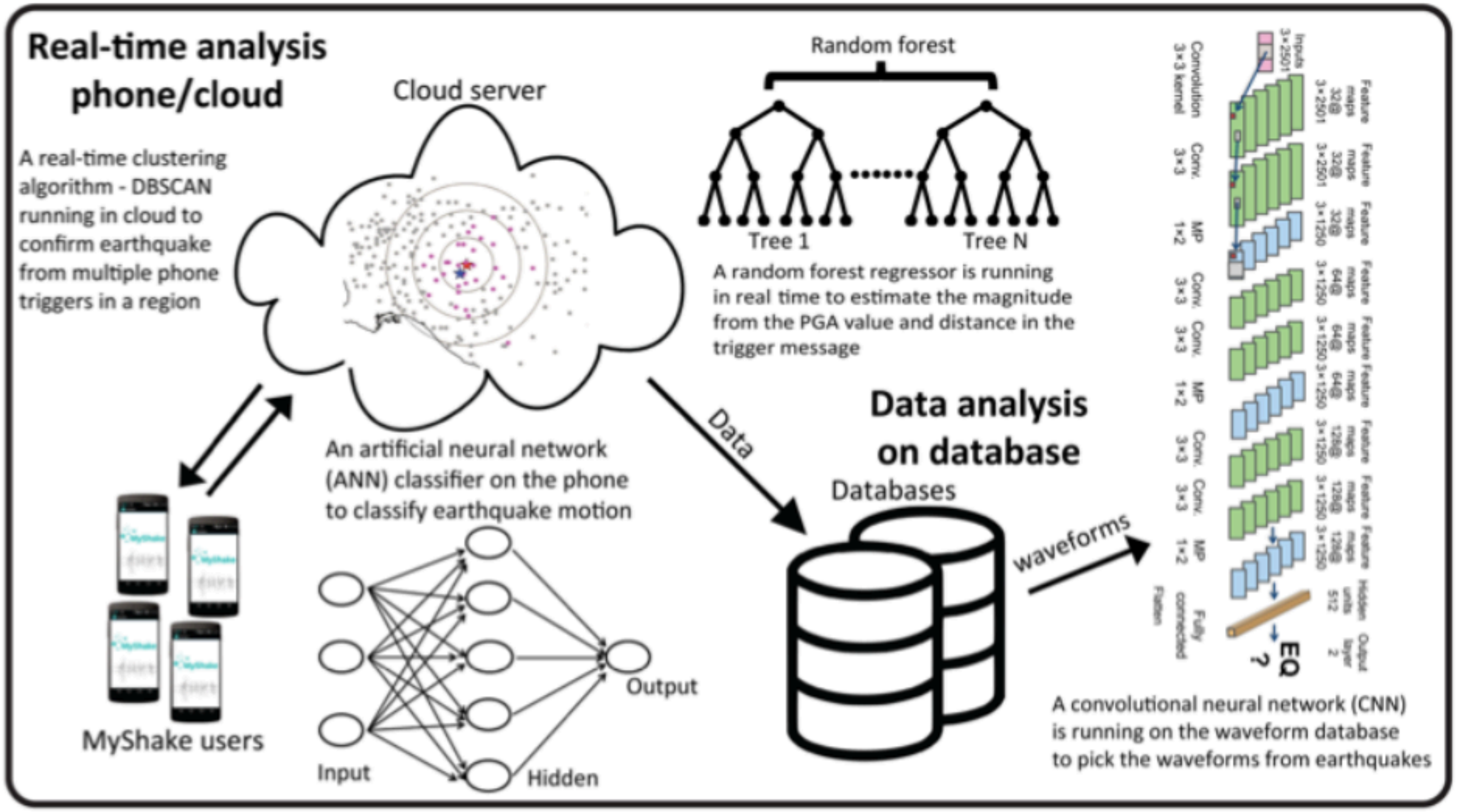


▲ **Figure 5.** The NN used in the MyShake earthquake early warning (EEW) phone application. (a) The workflow of the NN algorithm on the phone, including extraction of features from recorded phone motion and implementation of an NN classifier to distinguish between motions from humans and earthquakes. (b) The interquartile range and maximum zero crossing rate are two important features for distinguishing between earthquake and nonearthquake motions (modified from Kong, Allen, Schreier, *et al.*, 2016). (c) Example application of MyShake at the network level to an **M** 4.4 earthquake that occurred in January 2018. NN triggers from individual users are compared against theoretical *P* and *S* arrivals.

from Kong et al., 2018 SRL, 90(1), 3–14, doi: 10.1785/0220180259.



EEW and Real-Time ML (2)



▲ **Figure 1.** Sketch overview of the MyShake system and the machine learning (ML) algorithms that are currently used or under testing in the system both in real time and offline modes. DBSCAN, density-based spatial clustering of applications with noise; PGA, peak ground acceleration. The color version of this figure is available only in the electronic edition.

from: Kong, Q., A. Inbal, R. M. Allen, Q. Lv, and A. Puder (2018), Machine Learning Aspects of the MyShake Global Smartphone Seismic Network, Seismological Research Letters, 1–7, doi:10.1785/0220180309.

Ground-Motion Prediction Using Supervised Learning

The **classical approach** to ground-motion prediction **uses linear regression** to model the first-order aspects of these effects (Campbell and Bozorgnia, 2008).

- input to the NN are *M, Vs,30, resonant frequency, source depth, Rhypo or Repi*
- trained with natural log of the *recorded amplitude* [PGA, PGV, SA(T)]

- Earlier work done by
  - Alavi and Gandomi (2011)
  - Derras et al. (2012)

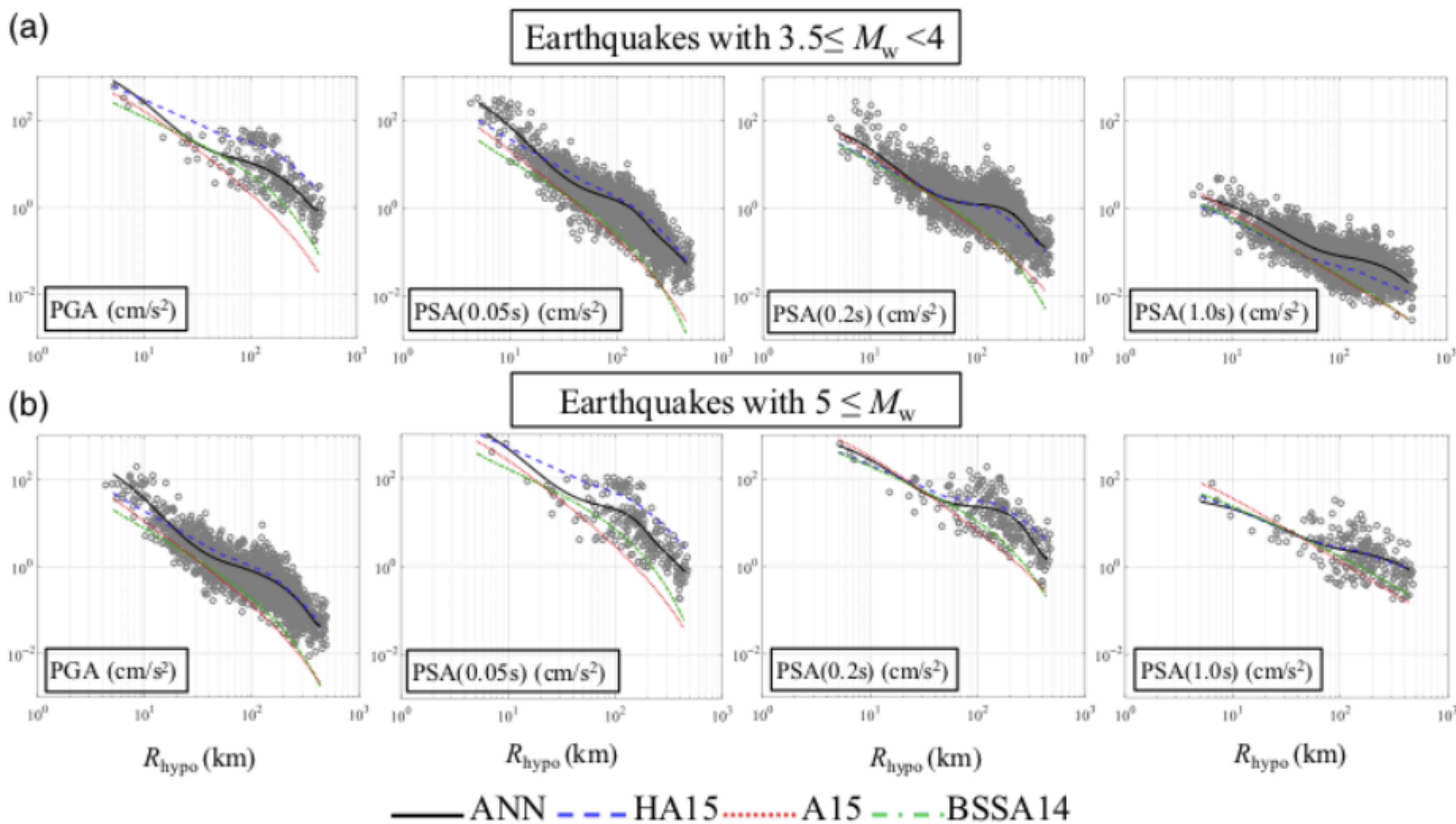
Ground-Motion Prediction Using Supervised Learning (2)

- Recent work
  - Derras et al. (2014 and 2016) NN with a single hidden layer to predict peak ground acceleration, velocity and pseudospectral accelerations at periods of interest for structural design.

  - Alimoradi and Beck (2014) developed a technique to synthesize realistic strong-motion records by applying Gaussian process regression to a sparse, orthonormal set of basis vectors called eigenquakes, which represent characteristic earthquake records.
  - **Khosravikia et al. (2018)**

- **Trugman and Shearer (2018)** used a generalization of the random forest supervised learning algorithm to relate earthquake stress drop and PGA (5297 earthquakes M1-4). The event residual terms learned by the random forest GMPE have a physical basis in the variability in earthquake stress drop, highlighting the utility of ML techniques in ground-motion modeling.
- Khoshnevis and Taborda (2019) ANN to estimate peak ground velocity

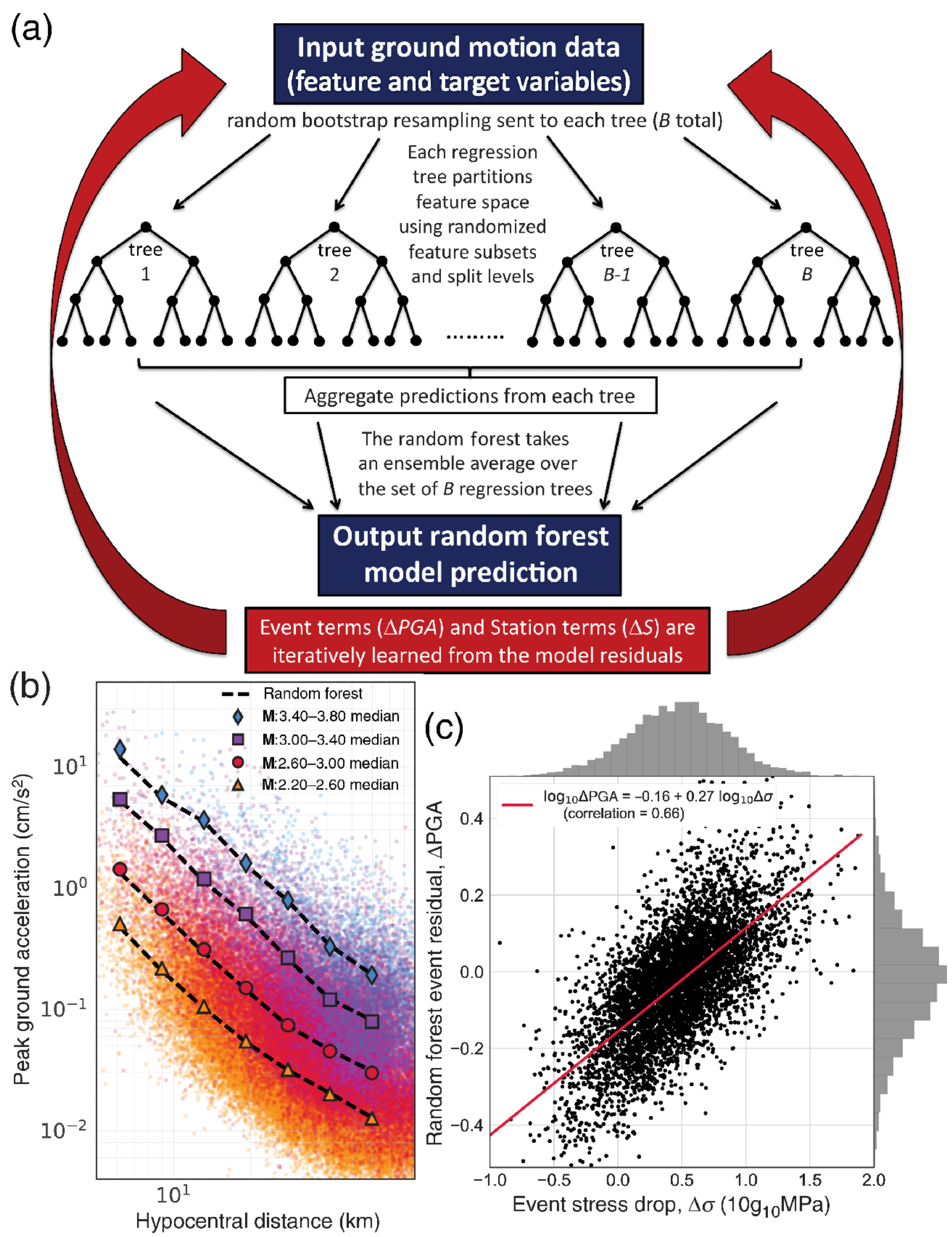
Ground-Motion Prediction Using Supervised Learning



▲ **Figure 7.** Intensity measure to distance relations of the GMMs determined in this study in comparison with Hassani and Atkinson (2015; hereafter, HA15) GMMs developed for central and eastern North America, Atkinson (2015; hereafter, A15) GMMs developed for small to moderate events at short hypocentral distances with applicability to induced seismicity, and Boore *et al.* (2014; hereafter, BSSA14) developed as part of the Next Generation Attenuation-West2 project. All GMMs are plotted for V_{S30} = 760 m/s as well as M_w = 3.7 for (a) and M_w = 5.3 for (b). The color version of this figure is available only in the electronic edition.

from: Khosravikia et al (2019) Artificial Neural Network–Based Framework for Developing Ground-Motion Models for Natural and Induced Earthquakes in Oklahoma, Kansas, and Texas

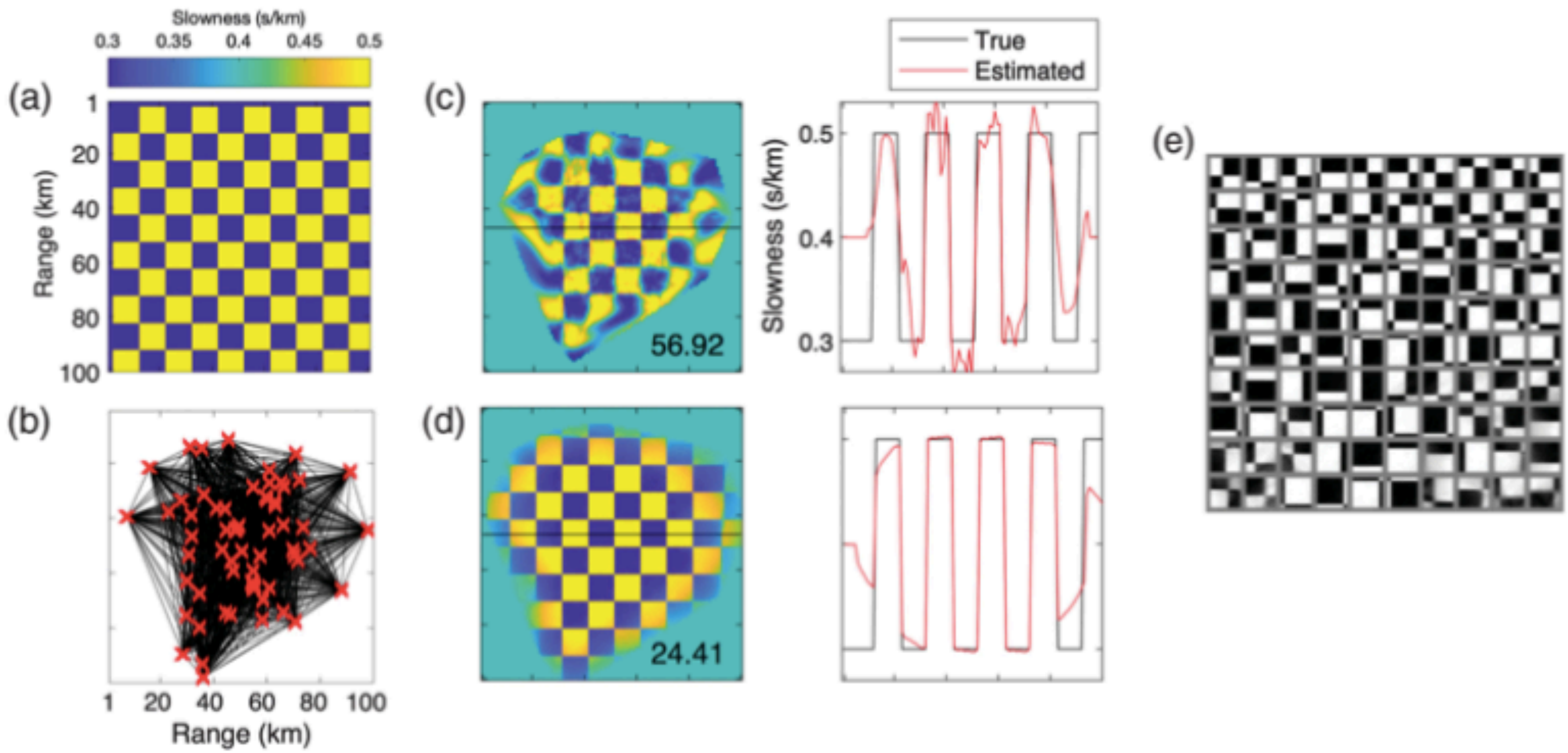
Ground-Motion Prediction Using Supervised Learning



Random forest ground-motion prediction equation (GMPE; Trugman and Shearer, 2018) and earthquake stress drop versus peak ground acceleration (PGA). (a) Schematic workflow for training the random forest GMPE. (b) PGA versus hypocentral distance for seismicity in the San Francisco Bay Area. Each point represents a site-corrected PGA measurement from an earthquake at a single station. Also shown is the median value in equally spaced magnitude–distance bins (large markers) and predicted values from the random forest GMPE (dashed lines). (c) Event PGA residuals learned from random forest GMPE versus earthquake stress drop. The least-squares linear fit and correlation coefficients are marked for reference.

from Trugman, D. T., and P. M. Shearer (2018). Strong correlation between stress drop and peak ground acceleration for eecent M 1-4 earth- quakes in the San Francisco bay area, Bull. Seismol. Soc. Am. 108, no. 2, 929–945

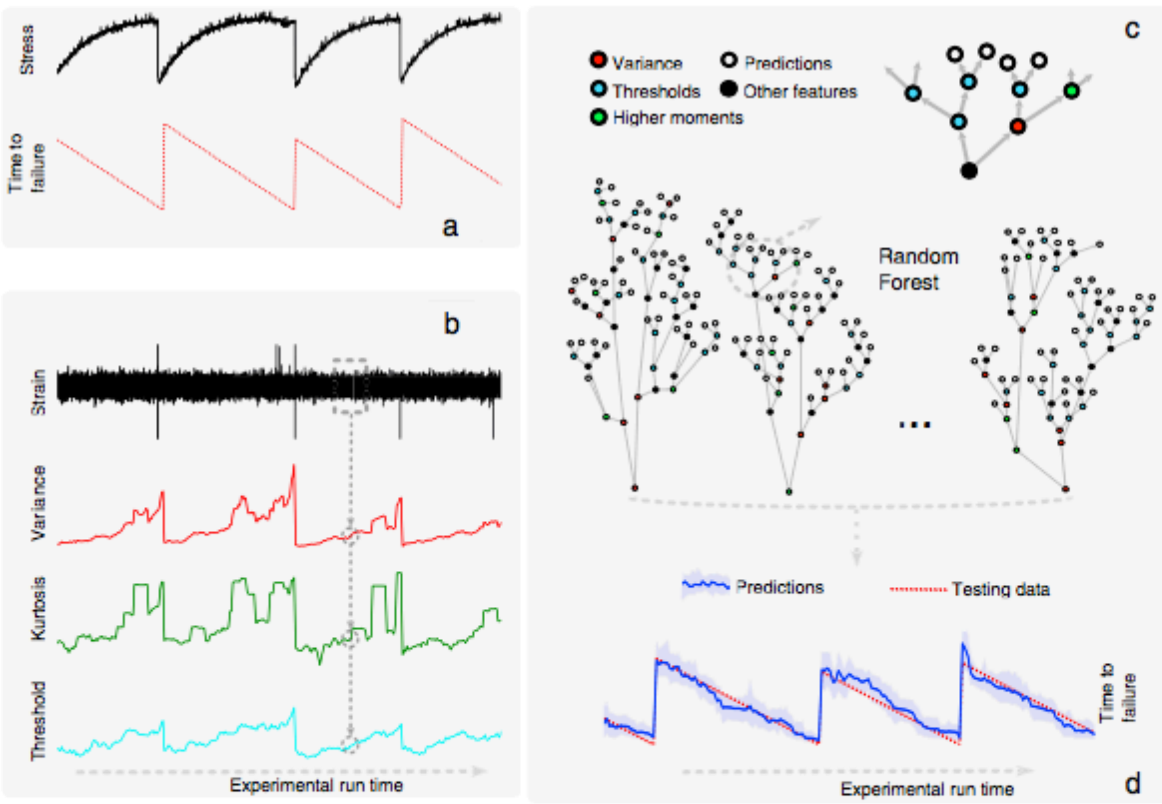
Model inversion/tomography with ML



▲ **Figure 7.** Locally sparse travel-time tomography (LST; Bianco and Gerstoft, 2018) of checkerboard slowness. (a) Synthetic checkerboard slowness patterns with 100 × 100 pixel grid (km) are sampled by (b) 2016 straight rays from 64 seismic stations. (c) Conventional inversion using damping and smoothing regularization (Aster *et al.*, 2011) and (d) LST. Profiles from the 2D inversion are shown with true and estimated slownesses. The root mean square error (ms/km) estimated relative to the true slowness is printed on the 2D estimates. (e) Dictionary learned from LST contains checkerboard-like atom (100 atoms shown). Each atom (patch) is 10 × 10 pixels.

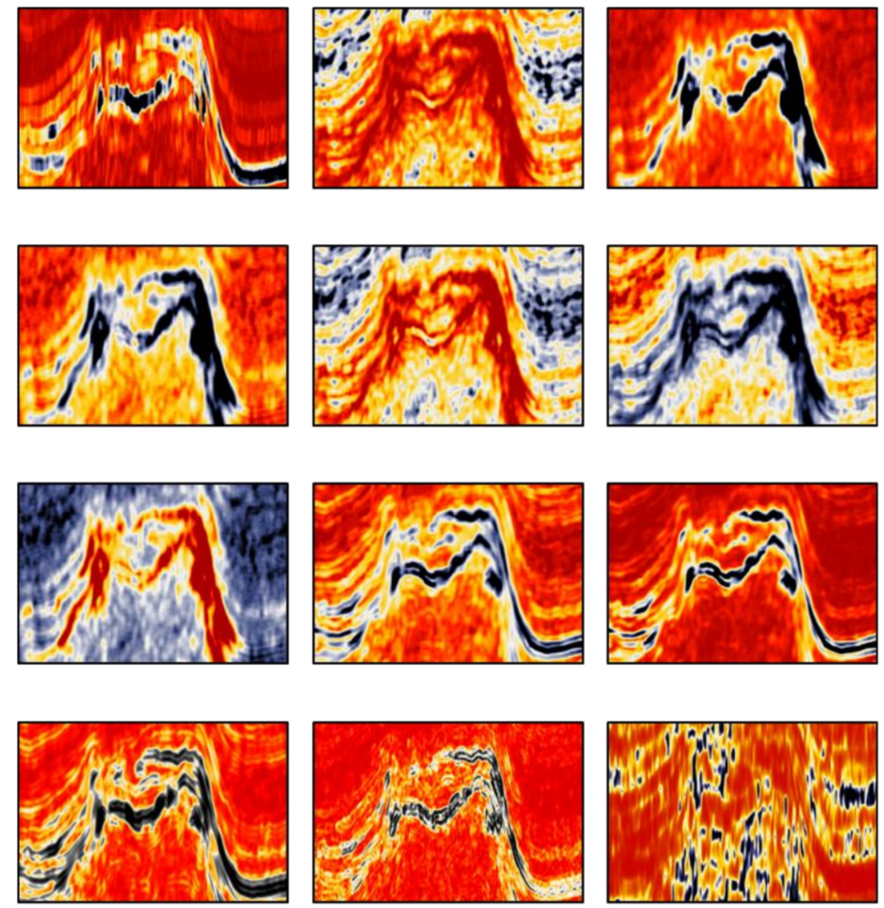
Other applications

Machine Learning Predicts Laboratory Earthquakes



Rouet-Leduc, B., C. Hulbert, N. Lubbers, K. Barros, C. J. Humphreys, and P. A. Johnson (2017), Machine Learning Predicts Laboratory Earthquakes, GEOPHYSICAL RESEARCH LETTERS, 44(18), 9276–9282, doi: 10.1002/2017GL074677.

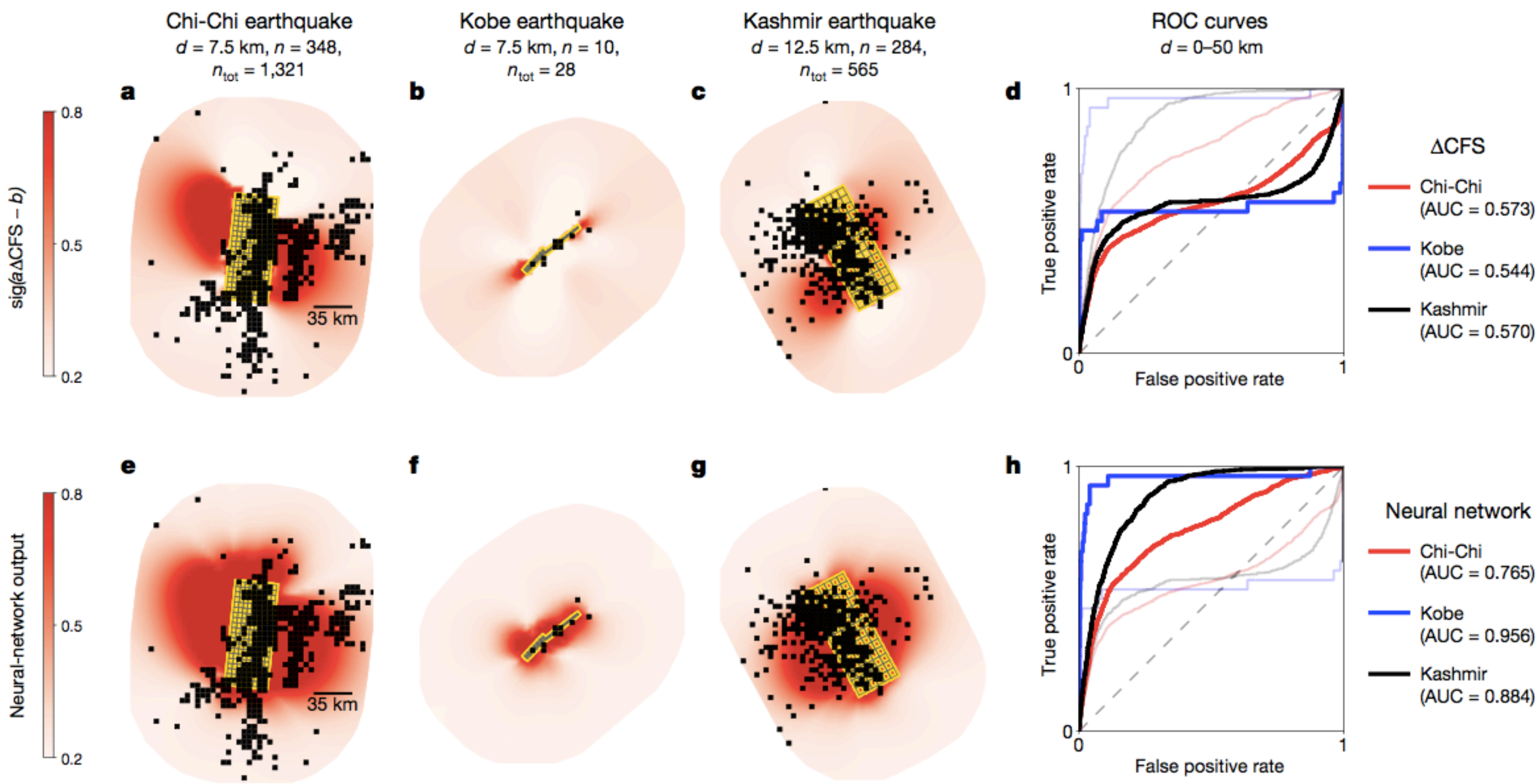
Machine learning meets seismic interpretation



https://agilescientific.com/blog/2017/6/20/machine-learning-meets-seismic-interpretation

Other applications (2)

Machine learning to forecast aftershock locations



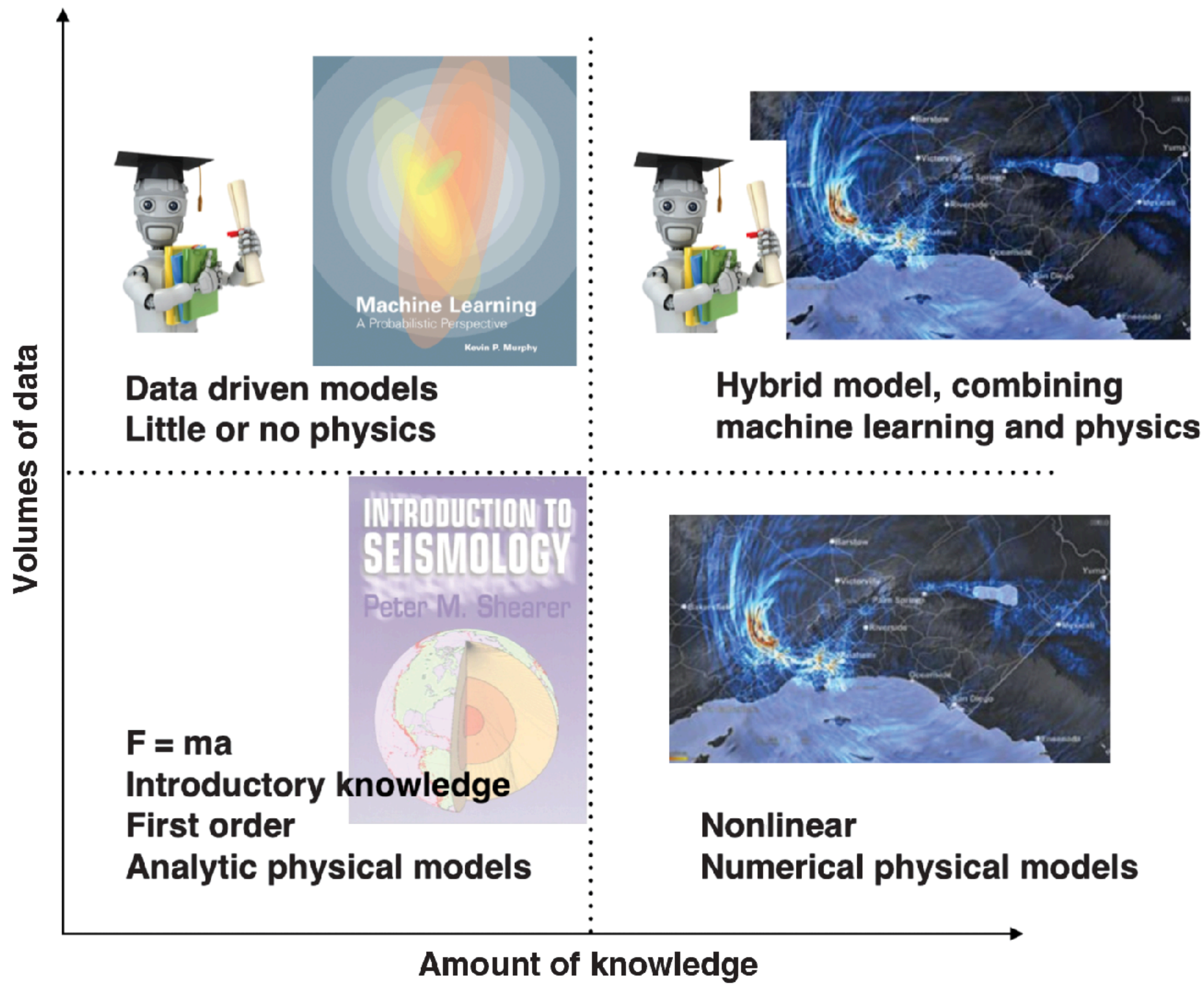
"…we use a **deep-learning** approach to identify a static-stress-based criterion that **forecasts aftershock locations** without prior assumptions about fault orientation. We show that **a neural network trained on more than 131,000 mainshock–aftershock pairs can predict the locations of aftershocks** in **an independent test dataset of more than 30,000 mainshock–aftershock pairs** more accurately (area under curve of 0.849) than can classic Coulomb failure stress change (area under curve of 0.583). …"

DeVries, P. M. R., F. Viégas, M. Wattenberg, and B. J. Meade (2018), Deep learning of aftershock patterns following large earthquakes, Nature, 1–16, doi:10.1038/s41586-018-0438-y.

INGV ML work

- Esposito, A. M., F. Giudicepietro, L. D'Auria, S. Scarpetta, M. Martini, M. Coltelli, and M. Marinaro (2008), Unsupervised Neural Analysis of Very-Long-Period Events at Stromboli Volcano Using the Self-Organizing Maps, BULLETIN OF THE SEISMOLOGICAL SOCIETY OF AMERICA, 98(5), 2449–2459.
- Esposito, A. M., L. D'Auria, F. Giudicepietro, R. Peluso, and M. Martini (2012), Automatic Recognition of Landslides Based on Neural Network Analysis of Seismic Signals: An Application to the Monitoring of Stromboli Volcano (Southern Italy), Pure and Applied Geophysics PAGEOPH, 170(11), 1821–1832, doi:10.1007/s00024-012-0614-1.
- Esposito, A. M., L. D'Auria, F. Giudicepietro, T. Caputo, and M. Martini (2013), Neural analysis of seismic data: applications to the monitoring of Mt. Vesuvius, Annals of Geophysics, 56(4), 0446–9, doi:10.4401/ag-6452.
- Falsaperla, S., S. Graziani, G. Nunnari, and S. Spampinato (1996), Automatic classification of volcanic earthquakes by using Multi-Layered neural networks, Natural Hazards, 13(3), 205–228, doi:10.1007/BF00215816.
- Falsaperla, S., B. Behncke, H. Langer, M. Neri, G. G. Salerno, S. Giammanco, E. Pecora, and E. Biale (2013), “Failed” eruptions revealed by pattern classification analysis of gas emission and volcanic tremor data at Mt. Etna, Italy, International Journal of Earth Sciences, 103(1), 297–313, doi:10.1007/s00531-013-0964-7.
- Giudicepietro, F., A. M. Esposito, and P. Ricciolino (2017), Fast Discrimination of Local Earthquakes Using a Neural Approach, Seismological Research Letters, 88(4), 1089–1096, doi:10.1785/0220160222.
- Langer, H., S. Falsaperla, M. Masotti, R. Campanini, S. Spampinato, and A. Messina (2009), Synopsis of supervised and unsupervised pattern classification techniques applied to volcanic tremor data at Mt Etna, Italy, Geophys. J. Int, 178(2), 1132–1144, doi:10.1111/j.1365-246X.2009.04179.x.
- Langer, H., S. Falsaperla, A. Messina, S. Spampinato, and B. Behncke (2011), Detecting imminent eruptive activity at Mt Etna, Italy, in 2007–2008 through pattern classification of volcanic tremor data, Journal of Volcanology and Geothermal Research, 200(1-2), 1–17, doi:10.1016/j.jvolgeores.2010.11.019.
- Messina, A., and H. Langer (2011), Pattern recognition of volcanic tremor data on Mt. Etna (Italy) with KKAnalysis—A software program for unsupervised classification, Computers & Geosciences, 37(7), 953–961, doi:10.1016/j.cageo.2011.03.015.
- Gentili, S., and A. Michelini (2006), Automatic picking of P and S phases using a neural tree, Journal of Seismology, 10(1), 39–63, doi:10.1007/s10950-006-2296-6.

Conclusions



Thank you for the attention