Massachusetts Institute of Technology
Earth Resources Laboratory

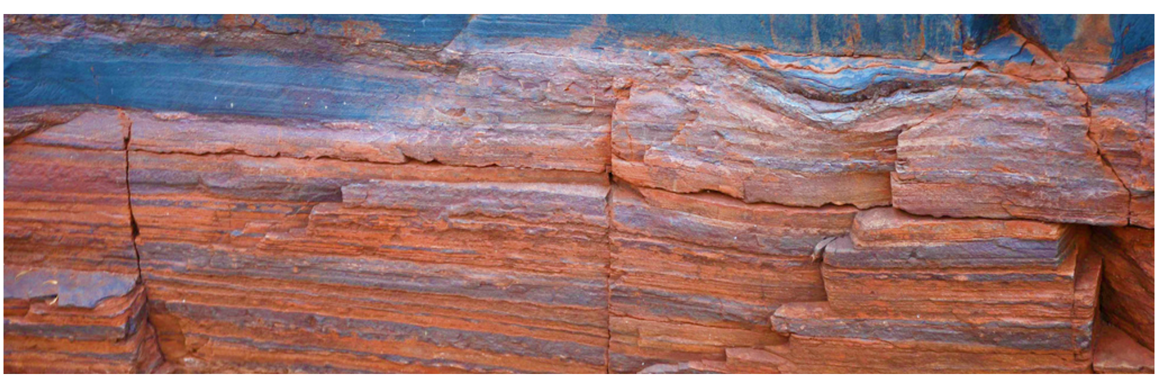
ERL in 2020

Research highlights and vision

Laurent Demanet, Associate Professor, Director

With input from:
Stephen Brown, Research Scientist
Daniel Burns, Research Scientist
Herbert Einstein, Professor
Michael Fehler, Senior Research Scientist
Aimé Fournier, Research Scientist
Ruben Juanes, Professor
Youssef Marzouk, Associate Professor
Nori Nakata, Principal Research Scientist
John Williams, Professor

Version: December 31, 2018



1. Personnel

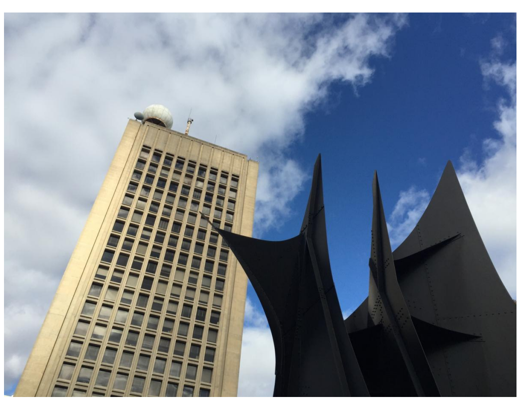
Faculty:

Name	Rank	Department(s)	Role in ERL
Laurent Demanet	Associate professor	Math and EAPS	Director
Herbert Einstein	Professor	CEE	
Brian Evans	Emeritus professor	EAPS	
Bradford Hager	Professor	EAPS	Associate director
Thomas Herring	Professor	EAPS	
Rob van der Hilst	Professor	EAPS	
Ruben Juanes	Professor	CEE and EAPS	
Youssef Marzouk	Associate professor	AeroAstro	
Dennis McLaughlin	Professor	CEE	
Dale Morgan	Professor	EAPS	Associate director
Shuhei Ono	Associate professor	EAPS	
Matej Pec	Assistant professor	EAPS	
Nafi Toksoz	Emeritus professor	EAPS	Founder
John Williams	Professor	CEE	

Research scientists with principal investigator status:

Name	Rank	Department(s)	Role in ERL
Stephen Brown	Research scientist	EAPS	
Daniel Burns	Research scientist	EAPS	
Michael Fehler	Senior RS	EAPS	Deputy director
Aimé Fournier	Research scientist	EAPS	
Nori Nakata	Principal RS	EAPS	
Sai Ravela	Principal RS	EAPS	

Research scientists (other): 8
Postdoctoral scholars: 15
Graduate students: 42
More information: https://erlweb.mit.edu/people



2. Introduction

The Earth Resources Laboratory (ERL) is MIT's home for geophysical research driven by technological questions. The laboratory is comprised of a dozen faculty members and their groups, active in areas ranging from **seismology to geomechanics, rock physics, flows in porous media,** and **methods of inversion, inference, and uncertainty quantification.**

As the "information revolution" is shaking up the research enterprise in many fields, ERL is embracing **scientific machine learning** as a main research objective. New tools lead to new questions, such as

- Are estimation and prediction still possible when the physical models are too coarse, or contain too much uncertainty, but when data are abundant?
- How can we bridge the "transfer learning" gap from synthetic to real data, or from labeled (rich) to unlabeled (poor) data?
- Uncertainty quantification in machine learning: what level of confidence should we give to the predictions that come from a neural network?
- Is it possible to automate tasks that otherwise require a human's ability to make generalizations?

Machine learning and artificial intelligence will only succeed in the sciences if their predictive power can outperform that of human-designed physical or statistical models. ERL has a long-term goal to identify the questions in geophysics, broadly understood, where machine learning genuinely extends the reach of traditional predictive models and data processing. You will find examples in this document.

ERL's research activity has an important role to play in addressing some of the environmental challenges of our time, for instance, via the design of next-generation capabilities in **geothermal engineering** and **carbon sequestration**.

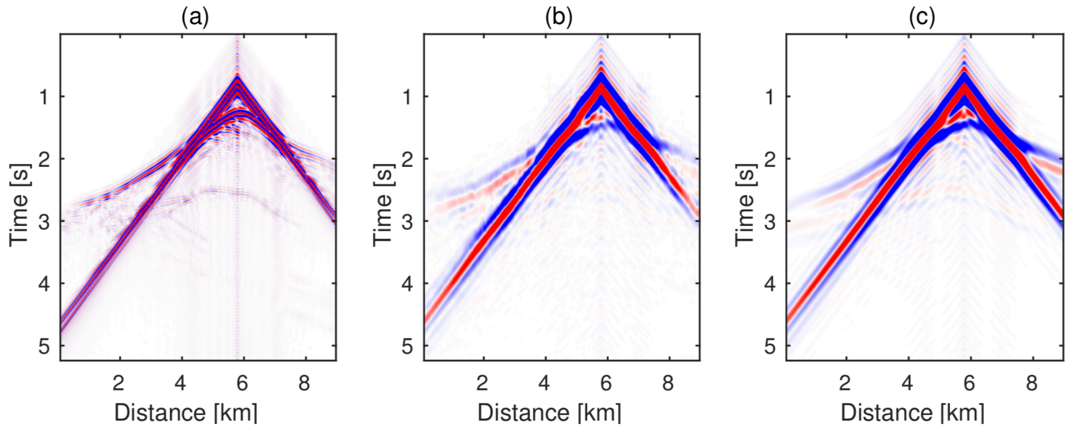
ERL's sponsored projects are often interdisciplinary, and bring together complementary expertise from all over MIT. The lab is unique in being able to integrate theory with physical evidence gathered from lab and field experiments. While ERL is primarily associated with the department of Earth, Atmospheric, and Planetary Sciences (EAPS), some of its faculty members are affiliated with the departments of Civil and Environmental Engineering; Mathematics; and AeroAstro.

Companies can join ERL as founding members to meet our students/postdocs, and get to know their research (see benefits of membership). We are continuously looking for talented people to join the lab (see our openings page). Stay in touch to learn about ERL's livestreams and other announcements: our social media are listed at https://erlweb.mit.edu/about.

We welcome your feedback, and hope to see you soon at one of our events!

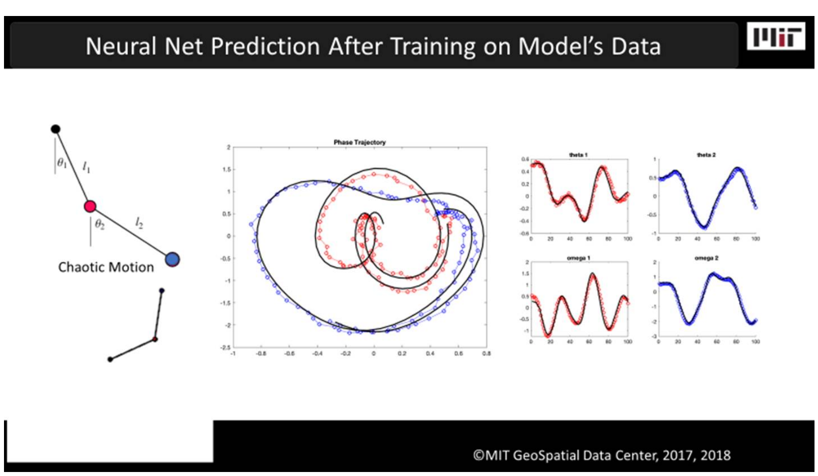
3. Research highlights

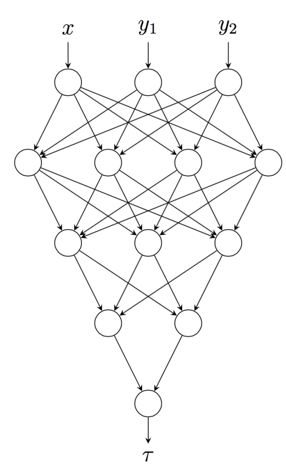
Deep learning for seismic bandwidth extension, in the group of **L. Demanet**



Over the past few years, L. Demanet and his group have explored computationally synthesizing the missing low frequencies in seismic data, for the purpose of addressing the nonconvexity challenge of full waveform seismic inversion (FWI). With Elita Li, they were the first to produce a method of signal processing, called the phase tracking method, to extrapolate the spectrum in unobserved bands. With Hongyu Sun, they designed a deep neural network to further improve and automate seismic bandwidth extension. They showed that such extrapolated data can usefully and robustly complement the observed data to alleviate the hardness of FWI, and converge to the correct solution in a range of useful geophysical scenarios. Figure above: a neural network trained on patches of the Marmousi community model can create the low frequencies of the BP 2004 model below 2Hz (center panel), and subsequently produce a good initial model for FWI to converge. (Left panel: input high frequencies. Right panel: true unknown low frequencies, for validation)

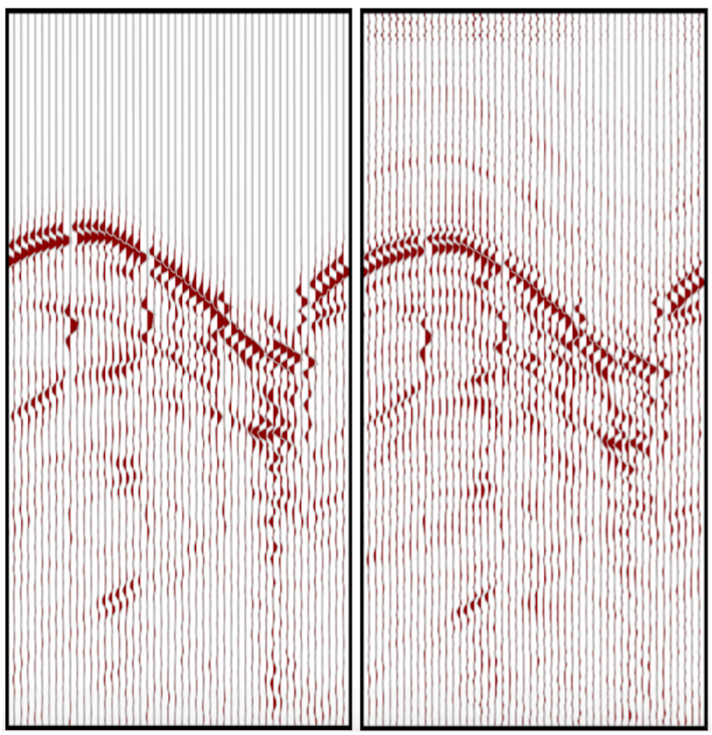
Surprises with machine learning, in the group of **J. Williams**





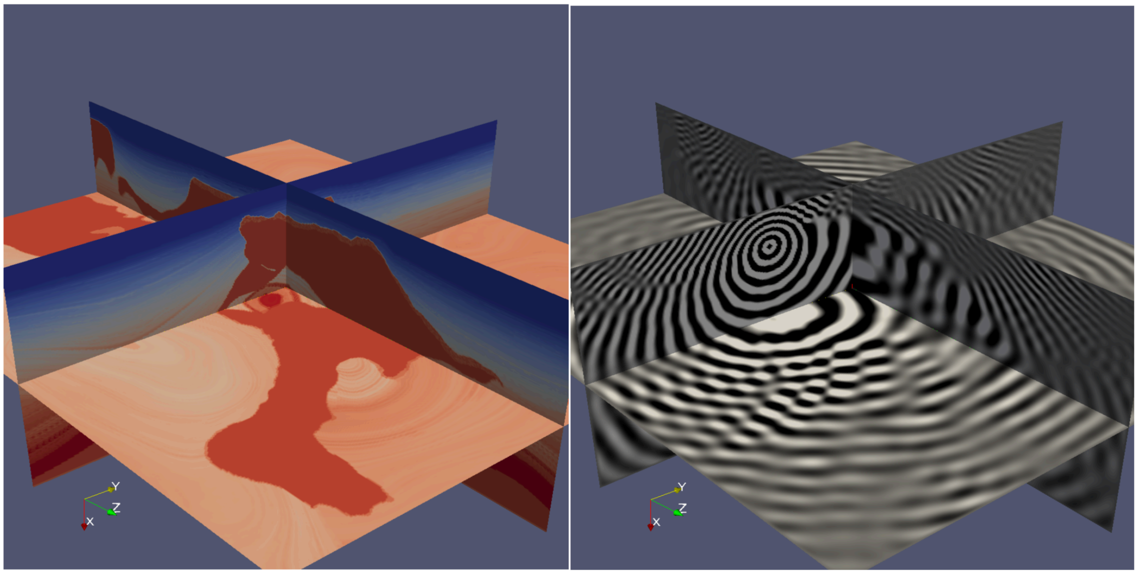
In December 2017, student Justin Montgomery co-authored an article on prediction of tight oil well productivity that was the first to properly disentangle the effects of new technology vs favorable well placement for predicting well production. They concluded that the US Department of Energy had vastly overstated the size of the US oil reserves, a finding for which they made the news. This is one example of how the group of John Williams is exploring how machine learning (ML) can be used to simulate physical systems. Most physical systems are highly non-linear, and they show how to use ML to "correct" for such non-linearities as time evolves. Once trained, the ML can make predictions faster than a traditional PDE simulator. Figure above: initial results predicting the behavior of a chaotic system.

Focused blind deconvolution, in the group of **A. Fournier and L. Demanet**



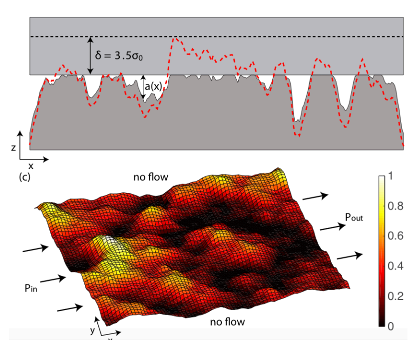
Aimé Fournier and collaborators have had two major research activities. With MITEI funding aiming at monitoring carbon sequestration, they developed new cement-integrity diagnostics using casing waveforms inferred from borehole data. With Equinor funding aiming at offshore drilling de-risking, they developed new data-analysis and inversion methods for MWD. "Focused blind deconvolution" (figure right) is a new method of machine learning to infer the true Green function (left) of a Marmousi impedance-model segment, using only drill-noise VSP records, without industry-standard (but dubious) assumptions about the model and source. They also extended seismo-electromagnetic data analysis and focusing to severely constrained borehole-acquisition geometries, with the aim that inferred resistivity structures constrain velocity inversion.

Scalable solvers for the 3D Helmholtz equation, in the group of **L. Demanet**



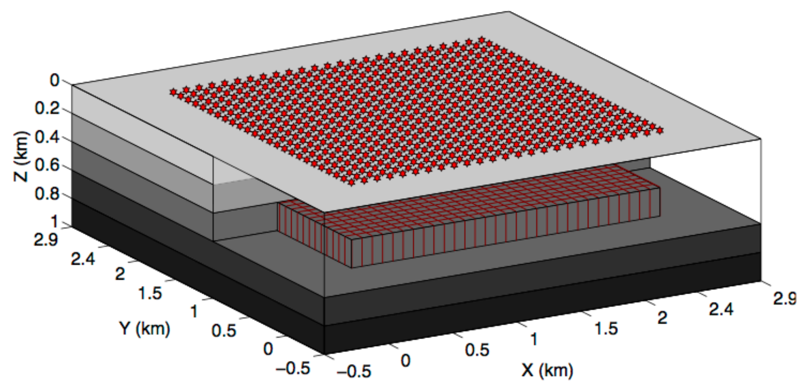
A line of work in L. Demanet's group concerns the design of truly scalable solvers for fixed-frequency computational wave propagation. The work with Leonardo Zepeda-Nunez, Matthias Taus, and others was the first to show that it is possible to achieve runtimes that are sublinear in the number N of grid points, on parallel clusters with P processors. The runtime is then proportional to N/P, with limitations on the size of P that were progressively loosened over the past few years. The algorithm that makes this possible is called the method of polarized traces. Extensions of the method also deal in a favorable way with the question of solving M > 1 problems; in that case, the runtime can grow more slowly than proportional to M. These computational developments should help address one practical bottleneck of inverse scattering for seismology and other applications. Figures: a computation with N = 4e8, on the SEAM model, which runs at 30s per right-hand side.

Emergence of anomalous transport in stressed rough fractures, in the group of **R. Juanes**



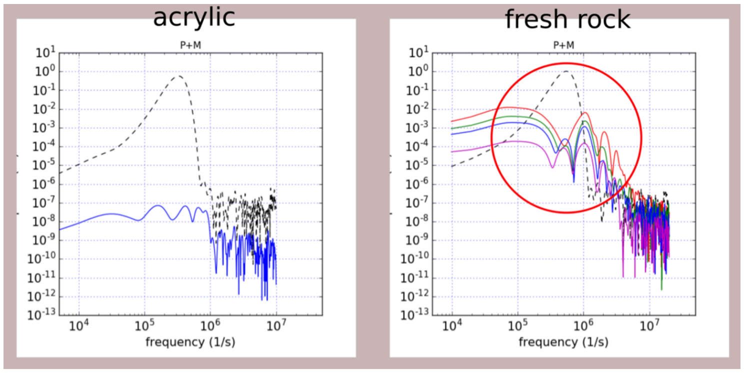
Fractures often serve as fast conduits for fluid flow and transport. The way in which the two rough surfaces of a fracture conform to each other depends critically on the level of confining stress. While the impact of normal stress on fluid flow through a rough fracture has been studied before, its impact on particle transport has not. Here, the authors show that particle transport on a rough fracture exhibits a transition from normal (Fickian) to anomalous (non-Fickian) as the level of confining stress increases, as a result of self-organization of the flow into preferential channels and stagnation regions. They propose a parsimonious stochastic transport model that accounts for the correlated nature of the flow velocity, and captures the transition to anomalous transport quantitatively. P. K. Kang, S. Brown, and R. Juanes, *Earth and Planetary Science Letters*, **454**, 46-54 (2016), doi:10.1016/j.epsl.2016.08.033. (pdf)

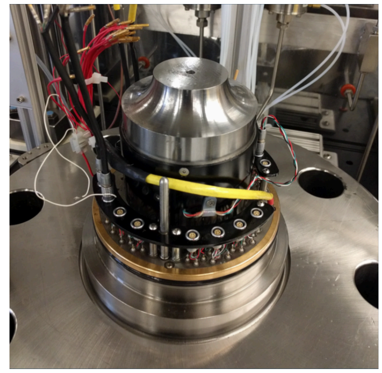
Sequential approach to joint flow-seismic inversion for improved characterization of fractured media, in the group of **R. Juanes**



Fracture characterization is a critical step for the design and risk assessment of many subsurface technologies, including nuclear waste disposal, geothermal energy production, and groundwater use. While seismic interpretation is essential to determine subsurface structures, it usually cannot constrain the medium's hydraulic properties, especially in challenging geologic environments like naturally fractured reservoirs. Here, the authors present a methodology for characterizing fractured geologic reservoirs by integrating flow and seismic data, which relies critically on the mechanistic relation between fracture compliance and fracture transmissivity. Their novel seismic inversion method provides the structural organization of the fracture compliance field, and inversion with flow data constrains the rock physics model and the seismic error model. By incorporating dynamic flow measurements into the seismic interpretation, they reduce the uncertainty in the seismic interpretation and dramatically improve the predictive ability of the reservoir flow models. P. K. Kang, Y. Zheng, X. Fang, R. Wojcik, D. McLaughlin, S. Brown, M. C. Fehler, D. R. Burns, and R. Juanes, *Water Resources Research*, **52**(2), 903–919 (2016), doi:10.1002/2015WR017412. (pdf)

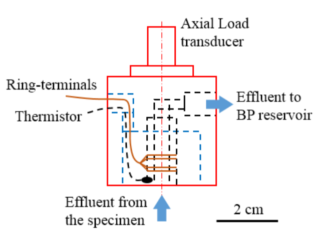
Quantitative nonlinearity in the subsurface, in the group of **S. Brown and D. Burns**



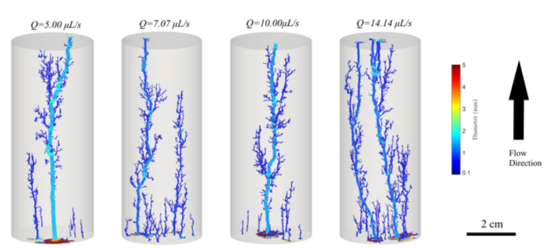


Like in medical imaging, multi-wave physics and nonlinear elasticity can provide new pathways for detection and monitoring of subtle features of importance in the subsurface. S. Brown and collaborators successfully adapted and applied principles from elastography and harmonic imaging to measure elastic nonlinearity parameters in rocks, which are sensitive to pore structure changes, variations in pore fluids and saturation levels, and to damage and stress. In work on sandstone samples, they have been able to quantify nonlinear behavior using time of flight transmission measurements. They now are looking at analysis methods using higher harmonics that would be more adapted to field scale reflection geometries. Applications include carbon capture and sequestration, hydrofracture operations and re-injection of hydrofracture fluids, enhanced/improved oil recovery, groundwater remediation, and geothermal field operations. Figure, left: the spectrum of higher harmonics generated by a single harmonic source. Right: the Autolab pressure chamber in B. Evans's laboratory, which can recreate the pressures and temperature present downhole.

Dissolution Processes in the Rock Matrix and in Rock Fractures, in the group of **H. Einstein**



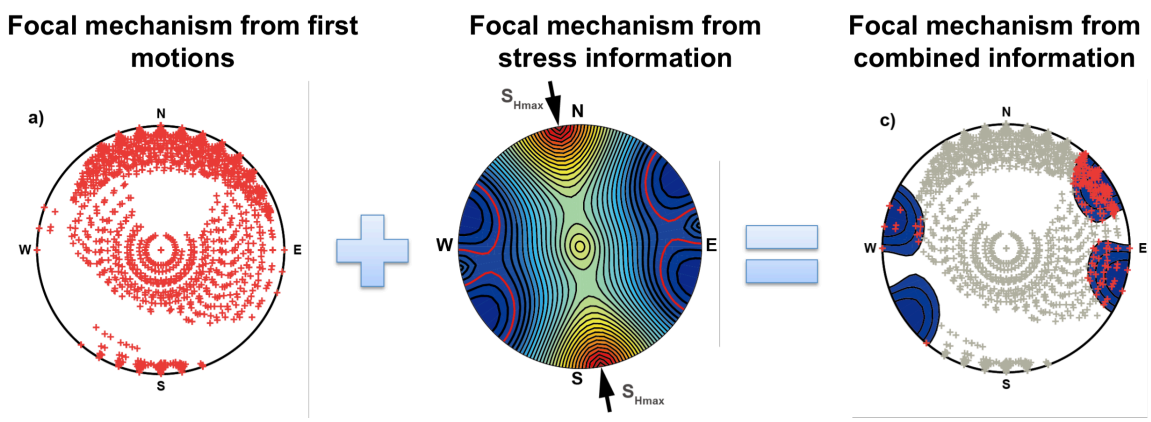
Device to measure effluent concentration



Wormhole development

Rock dissolution processes occur in nature (Karst) and in acid stimulation of hydrocarbon reservoirs as well as in CO_2 sequestration. Laboratory core flooding experiments with a new device that continuously measures effluent concentration are used to observe wormhole development and fracture enlargement in Gypsum-Water analogs. On the basis of these experiments models have been developed to predict the dissolution processes and their consequences on pore- and fracture geometry.

Moment Tensor Inversion of Induced Seismic Events, in the group of **M. Fehler**

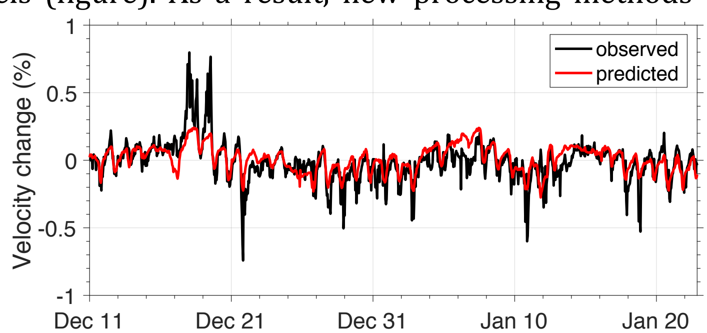


Most modern approaches for inferring the moment tensors involve an inversion where measured waveforms are compared with waveforms predicted from models simulated for a given set of fault parameters and velocity structure. This inversion is challenging due to the limited amount of data that are usually available, and uncertainties in the velocity model that can strongly impact the predicted waveforms. The authors have made progress by using a Bayesian formulation to obtain reliable estimates of the uncertainty of the results based on the input data and prior information, and the incorporation of additional data about the reservoir as a constraint on the inversion. Figure 1 shows an example of how they can use *in situ* stress information to constrain the focal mechanism determined from waveform first motions. The left panel shows a lower hemisphere projection of the range of focal mechanisms that are consistent with the first motion data from an induced seismic event. Each + represents the pole to a focal mechanism plane that is consistent with the data. There is a wide range of possible solutions. The middle panel shows schematically the pore pressure required for each possible plane to experience slip given the known *in situ* stress field. Blue represents regions of lower pore pressure. The red lines show the boundaries of the regions where planes may slip at a given known pore pressure determined from the pumping history. The right panel shows the range of possible slip planes that are consistent with the first motion data, the stress field and the injection pressure.

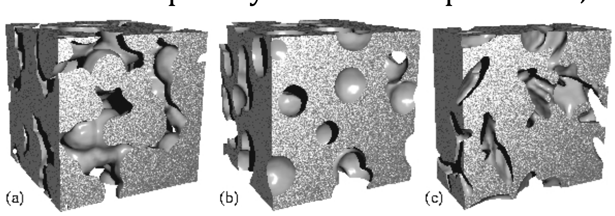
4. Vision

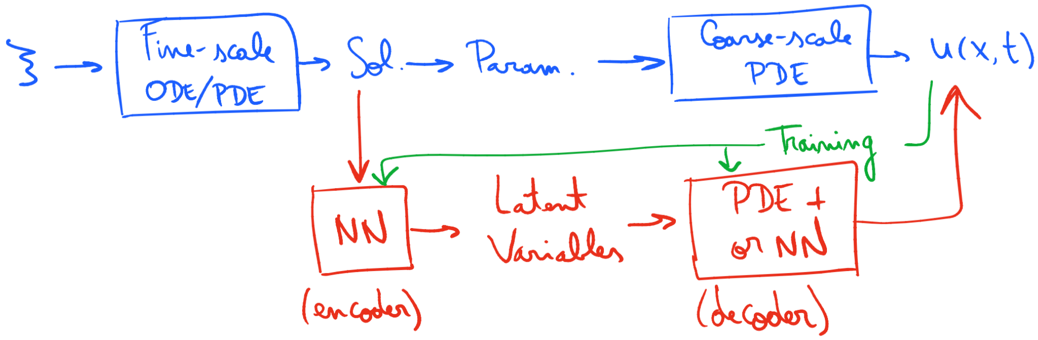
Machine learning (ML) is a common theme behind much of the proposed research activity in ERL. Possible topics for future investigation include:

- **How to make sense of ambient seismic noise.** The idea is to design new data processing methods for ambient seismic noise, in order to extract information that the current method of choice, cross-correlations of nearby stations, is unable to access. One limitation has been the inability to deal with non-isotropic random sources of waves. Demanet's group obtained early indications that deep neural networks can deal with this scenario, and moreover determine the directional dependence of the noise source in the far field itself. Another limitation has been the incomplete understanding of the underlying physics, which hinders the ability to make predictions of parameters like velocity changes in a reservoir. Nakata's group has obtained evidence that machine learning (SVM) can offer a meaningful forecast from environmental data, even in the absence of physical models (figure). As a result, new processing methods should unlock information that was not previously known to be present in noisy seismogram recordings. Providing starting models for full waveform inversion is one important end-goal of any new method of analyzing ambient seismic noise

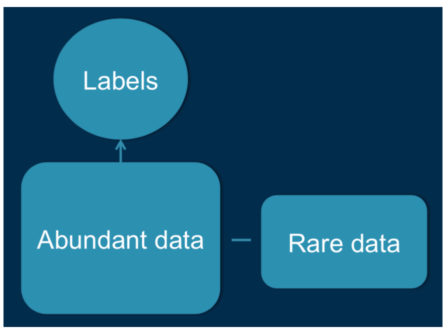


- **Deep learning upscaling**. Current upscaling practice consists in solving differential equations at the fine scale, in order to obtain effective parameters to be used in differential equations at the coarse scale. In this project, we plan to explore deep neural networks as an alternative point of view, with less reliance on the disciplinary scientist. In particular, we plan to consider two-phase flows in porous media. The traditional point of view is to use the Navier-Stokes equations at the mesoscale, determine permeability and viscosity for use in some form of Darcy's law, and set up homogenized models like Buckley-Leverett at the macro scale. The validity of these equations is dubious in the two-phase case, and does not account for hysteresis or capillary effects. Instead, we propose conditional autoencoders to replace the very notions of physical models and effective parameters with standard constructs from machine learning: encoder/decoder, and latent variables.





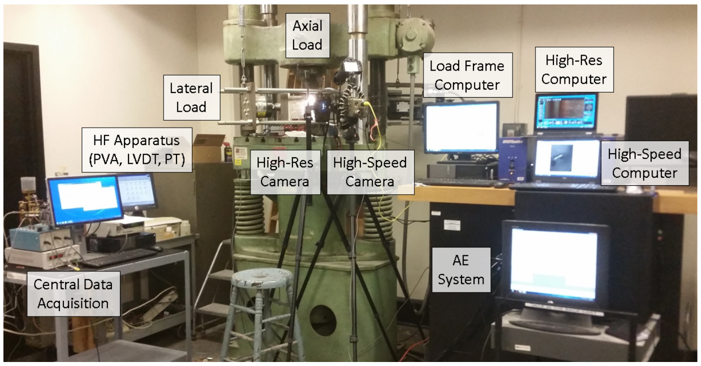
- **The transfer learning bridge.** The prospect of "scientific machine learning " as a meaningful discipline poses a range of fundamental questions. Prediction with neural networks can be risky because 1) real data may not be abundant enough for training, and 2) simulations are abundant, but may not be close enough to the real data to be useful for training. These issues arise for inversion of any kind of geophysical data, where deep nets are meant to replace physical laws. The proposed idea is to design methods of deep learning called transfer learning, that explicitly hardcode the similarity between real and simulation data in the feature space. It is related to the notion of semi-supervised learning. Figure: labels are only available for part of the data, yet are needed on the whole dataset.



- **Uncertainty quantification for ML.** The concept in the previous bullet point would be one facet of a broader community effort -- very much in its infancy -- to address uncertainty quantification FOR scientific machine learning. Just as we consider the validity of our physical models, and of inversions or predictions made with these models, we need to ascribe some level of confidence to any predictions or inferences built on ML models or architectures. With the proliferation of new techniques, the need for considering the statistical foundations and interpretations of any models we employ (whether they are misfit functions for inversion, discrepancy terms describing the mismatch between deep net predictions and data, new ensemble filtering techniques), is even greater. As practitioners blaze ahead, academia needs to ask the hard question: When or why are predictions from ML not merely accidentally accurate?

- **Machine learning for induced seismicity**. Machine learning has a lot of potential to replace current processing techniques for induced and triggered microseismic data, e.g., for forecasting future events. Migration-based imaging (e.g., Nakata and Beroza, 2016) can be used as an input for machine learning to simultaneously locate and detect microseismicity. Finding location and estimating characterization of events is considered an inverse problem, and machine learning, with careful parameter tuning and training, can surpass the limit of the conventional inversion resolution (Kim and Nakata, 2018). Denoising and/or big-data based on machine learning are also the future of wavefield analyses.

- **Laboratory and field experiments: rock fractures**. The labs of Herbert Einstein and Brian Evans provide a combination of expertise and capabilities in rock physics that is unique to the Earth Resources Laboratory. Disciplinary grounding and quality data are essential if one wishes to make progress in, e.g., any of the projects mentioned earlier. The labs study hydraulic fracturing with simultaneous visual and acoustic emission monitoring, dissolution processes with core flood experiments and CT scanning, fracture flow experiments, etc., with the ability to see what happens**.** Rock fracturing can be controlled in the lab, and it is intended to eventually apply these processes to control rock fracturing in the field as well. This will have applications in petroleum engineering, natural hazards and civil engineering.



This list is not an exhaustive representation of the vision of every group in ERL. Contacting individual PIs is the best way to obtain proposal-like materials. Your membership in the ERL consortium will go a long way toward maintaining ERL's ability to meet its research goals, although it is not a guarantee that any particular topic will be explored. For more information: https://erlweb.mit.edu/