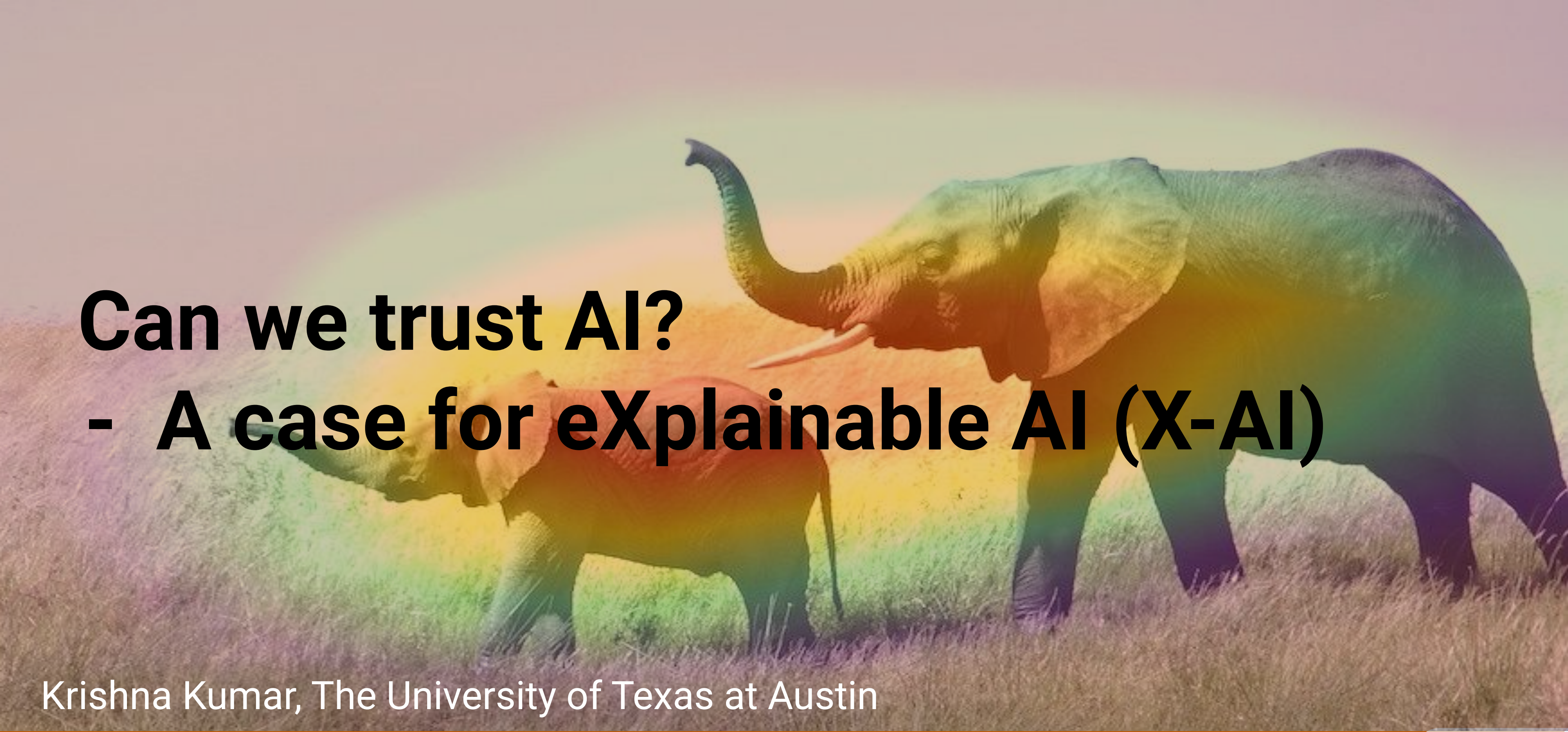


Can we trust AI?
- A case for eXplainable AI (X-AI)

Krishna Kumar, The University of Texas at Austin

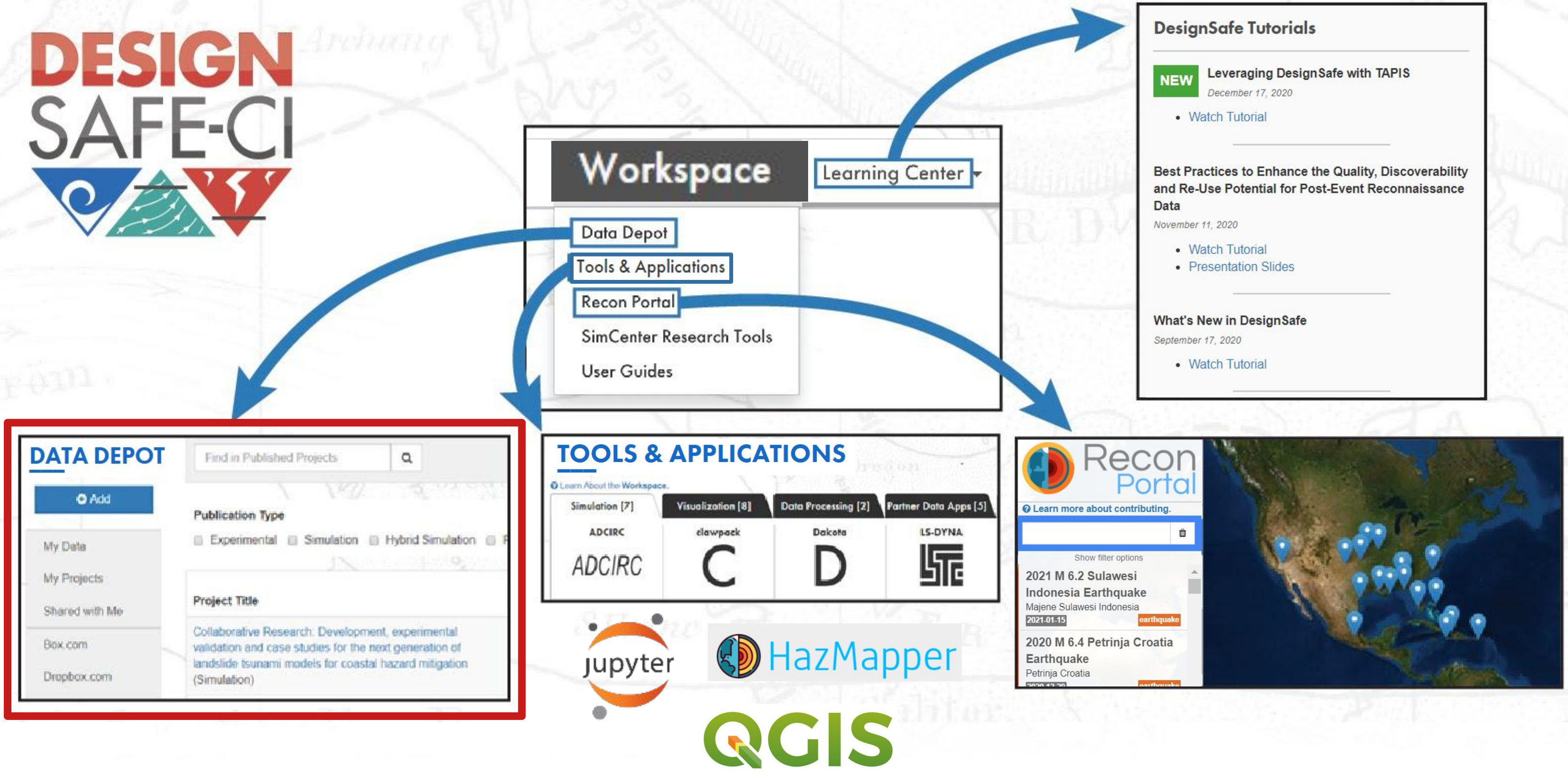
Machine Learning Training for Natural Hazards Engineering 2022

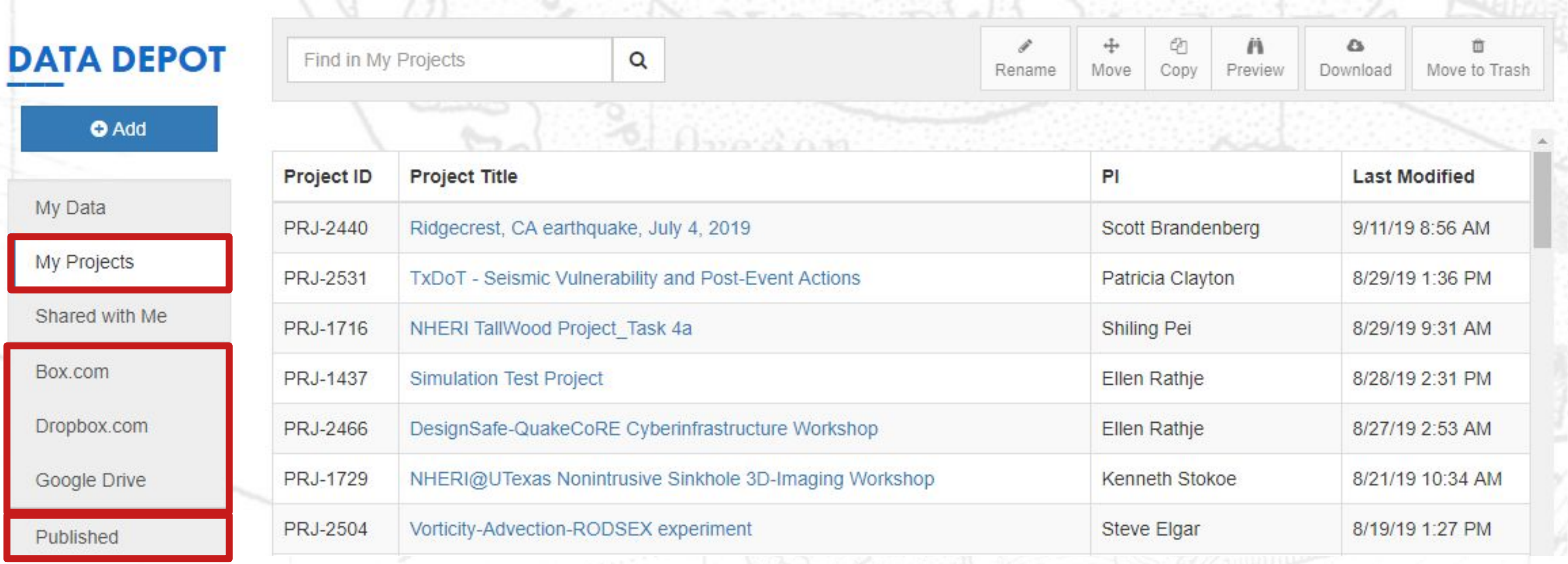
What is DesignSafe?

- A web-based research platform that enables transformative research to protect human life and reduce damage during natural hazard events

DesignSafe Vision

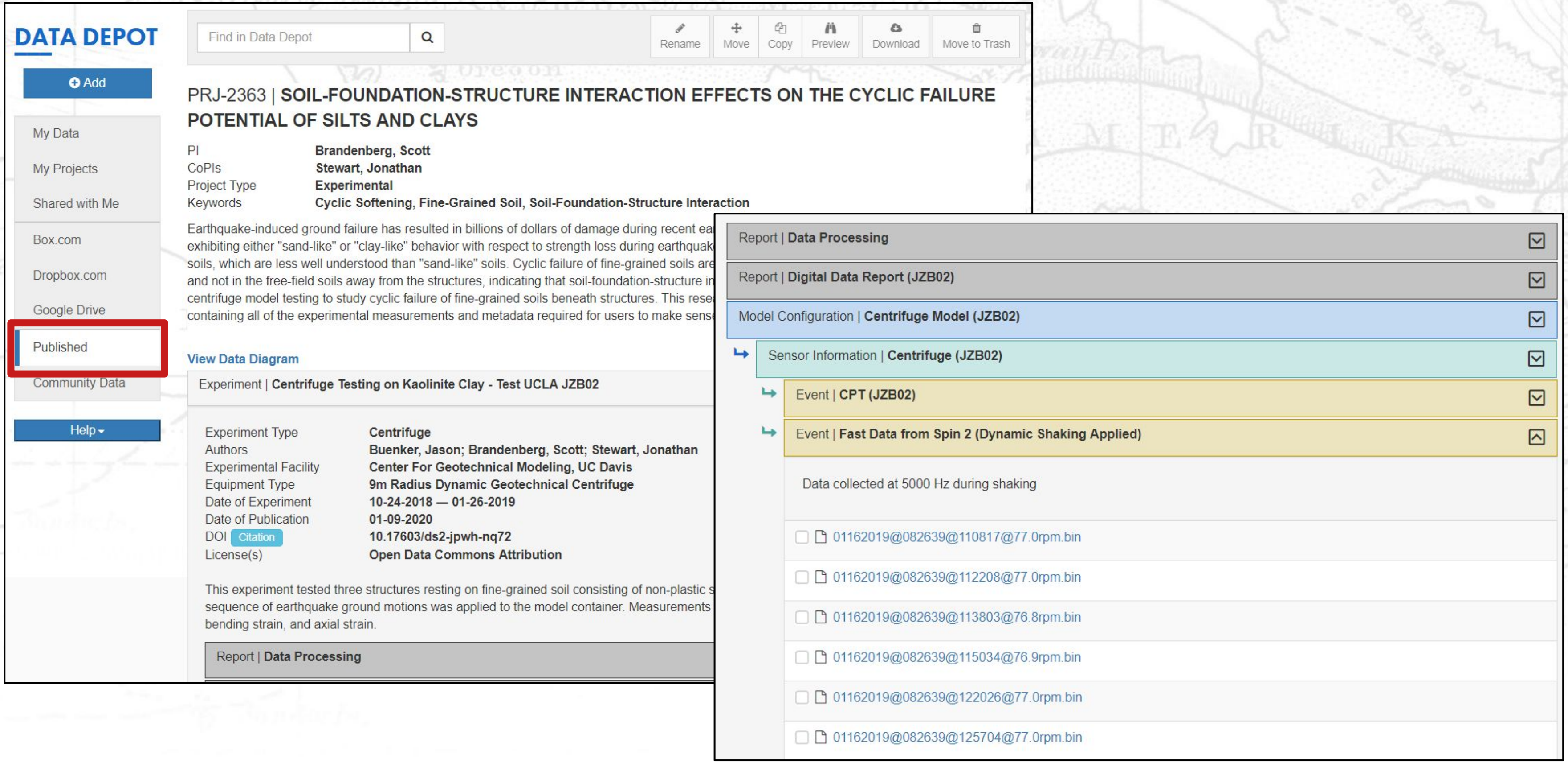
- A cyberinfrastructure (CI) that is an integral part of research discovery
  - Provide a platform for data sharing/publishing
  - Enable research workflows and access to high performance computing (HPC)
  - Deliver cloud-based tools that support the analysis, visualization, and integration of diverse data types
- Amplify and link the capabilities of natural hazards researchers in the US and abroad


DESIGN SAFE-CI
Workspace
Learning Center
Data Depot
Tools & Applications
Recon Portal
SimCenter Research Tools
User Guides
DesignSafe Tutorials
NEW
Leveraging DesignSafe with TAPIS
December 17, 2020
Watch Tutorial
Best Practices to Enhance the Quality, Discoverability and Re-Use Potential for Post-Event Reconnaissance Data
November 11, 2020
Watch Tutorial
Presentation Slides
What's New in DesignSafe
September 17, 2020
Watch Tutorial
DATA DEPOT
Find in Published Projects
Add
My Data
My Projects
Shared with Me
Box.com
Dropbox.com
Publication Type
Experimental
Simulation
Hybrid Simulation
Project Title
Collaborative Research: Development, experimental validation and case studies for the next generation of landslide tsunami models for coastal hazard mitigation (Simulation)
TOOLS & APPLICATIONS
Learn About the Workspace.
Simulation [7]
Visualization [8]
Data Processing [2]
Partner Data Apps [5]
ADCIRC
clawpack
Dakota
LS-DYNA
jupyter
HazMapper
QGIS
Recon Portal
Learn more about contributing.
Show filter options
2021 M 6.2 Sulawesi Indonesia Earthquake
Majene Sulawesi Indonesia
2021-01-15
earthquake
2020 M 6.4 Petrinja Croatia Earthquake
Petrinja Croatia



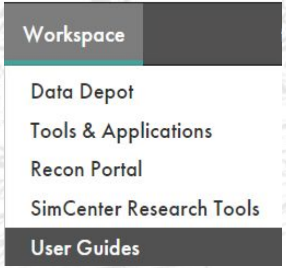
Project ID	Project Title	PI	Last Modified
PRJ-2440	Ridgecrest, CA earthquake, July 4, 2019	Scott Brandenberg	9/11/19 8:56 AM
PRJ-2531	TxDoT - Seismic Vulnerability and Post-Event Actions	Patricia Clayton	8/29/19 1:36 PM
PRJ-1716	NHERI TallWood Project_Task 4a	Shiling Pei	8/29/19 9:31 AM
PRJ-1437	Simulation Test Project	Ellen Rathje	8/28/19 2:31 PM
PRJ-2466	DesignSafe-QuakeCoRE Cyberinfrastructure Workshop	Ellen Rathje	8/27/19 2:53 AM
PRJ-1729	NHERI@UTexas Nonintrusive Sinkhole 3D-Imaging Workshop	Kenneth Stokoe	8/21/19 10:34 AM
PRJ-2504	Vorticity-Advection-RODSEX experiment	Steve Elgar	8/19/19 1:27 PM

My Projects: A space to share files/data/results with collaborators and to eventually publish for public use


DATA DEPOT
Add
My Data
My Projects
Shared with Me
Box.com
Dropbox.com
Google Drive
Published
Community Data
Help
Find in Data Depot
Rename
Move
Copy
Preview
Download
Move to Trash
PRJ-2363 | SOIL-FOUNDATION-STRUCTURE INTERACTION EFFECTS ON THE CYCLIC FAILURE POTENTIAL OF SILTS AND CLAYS
PI Brandenberg, Scott
CoPIs Stewart, Jonathan
Project Type Experimental
Keywords Cyclic Softening, Fine-Grained Soil, Soil-Foundation-Structure Interaction
Earthquake-induced ground failure has resulted in billions of dollars of damage during recent ea
exhibiting either "sand-like" or "clay-like" behavior with respect to strength loss during earthquak
soils, which are less well understood than "sand-like" soils. Cyclic failure of fine-grained soils are
and not in the free-field soils away from the structures, indicating that soil-foundation-structure in
centrifuge model testing to study cyclic failure of fine-grained soils beneath structures. This rese
containing all of the experimental measurements and metadata required for users to make sens
View Data Diagram
Experiment | Centrifuge Testing on Kaolinite Clay - Test UCLA JZB02
Experiment Type Centrifuge
Authors Buenker, Jason; Brandenberg, Scott; Stewart, Jonathan
Experimental Facility Center For Geotechnical Modeling, UC Davis
Equipment Type 9m Radius Dynamic Geotechnical Centrifuge
Date of Experiment 10-24-2018 — 01-26-2019
Date of Publication 01-09-2020
DOI Citation 10.17603/ds2-jpwh-nq72
License(s) Open Data Commons Attribution
This experiment tested three structures resting on fine-grained soil consisting of non-plastic s
sequence of earthquake ground motions was applied to the model container. Measurements
bending strain, and axial strain.
Report | Data Processing
Report | Data Processing
Report | Digital Data Report (JZB02)
Model Configuration | Centrifuge Model (JZB02)
Sensor Information | Centrifuge (JZB02)
Event | CPT (JZB02)
Event | Fast Data from Spin 2 (Dynamic Shaking Applied)
Data collected at 5000 Hz during shaking
01162019@082639@110817@77.0rpm.bin
01162019@082639@112208@77.0rpm.bin
01162019@082639@113803@76.8rpm.bin
01162019@082639@115034@76.9rpm.bin
01162019@082639@122026@77.0rpm.bin
01162019@082639@125704@77.0rpm.bin

Curation Assistance

- Curation and publication guidelines under ***User Guides***
  – https://www.designsafe-ci.org/rw/user-guides/data-curation-publication/
  – Managing protected/identifiable data – talk to Maria Esteva
- Data transfer methods
  – https://www.designsafe-ci.org/rw/user-guides/data-transfer-guide/
  – Web browser/Dropbox/etc (smaller uploads), Globus, Cyberduck
- Virtual Curation Office Hours
  – DesignSafe Data Curators: Maria Esteva and Craig Jansen
  – Tuesday and Thursday at 1 pm Central (or by appt)
  – https://www.designsafe-ci.org/learning-center/training/

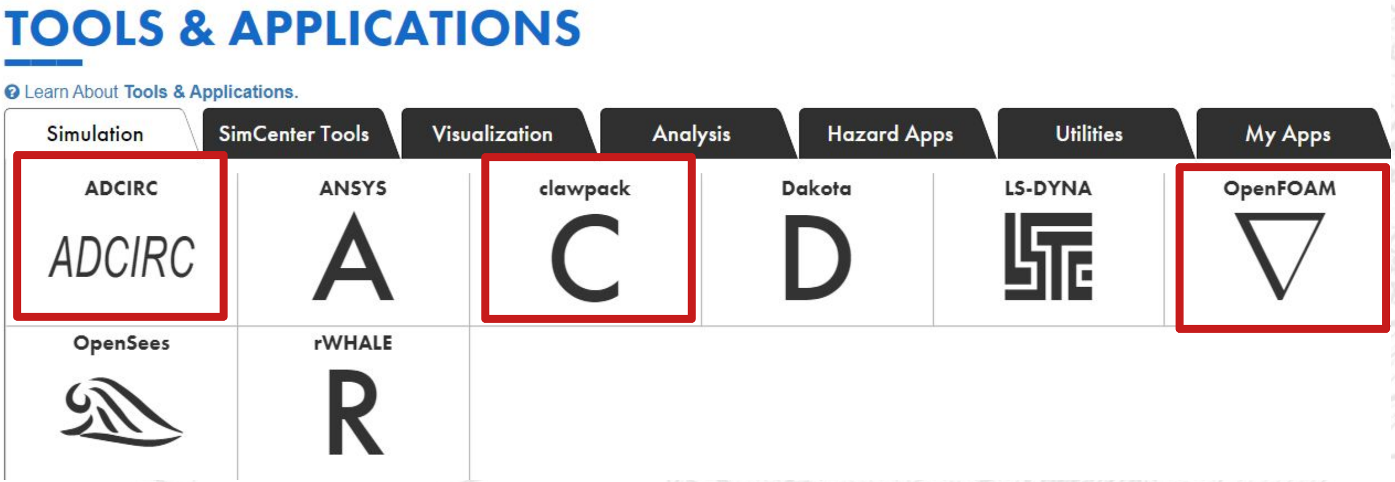


Reconnaissance Portal

Identifying Published Datasets from Recon Events

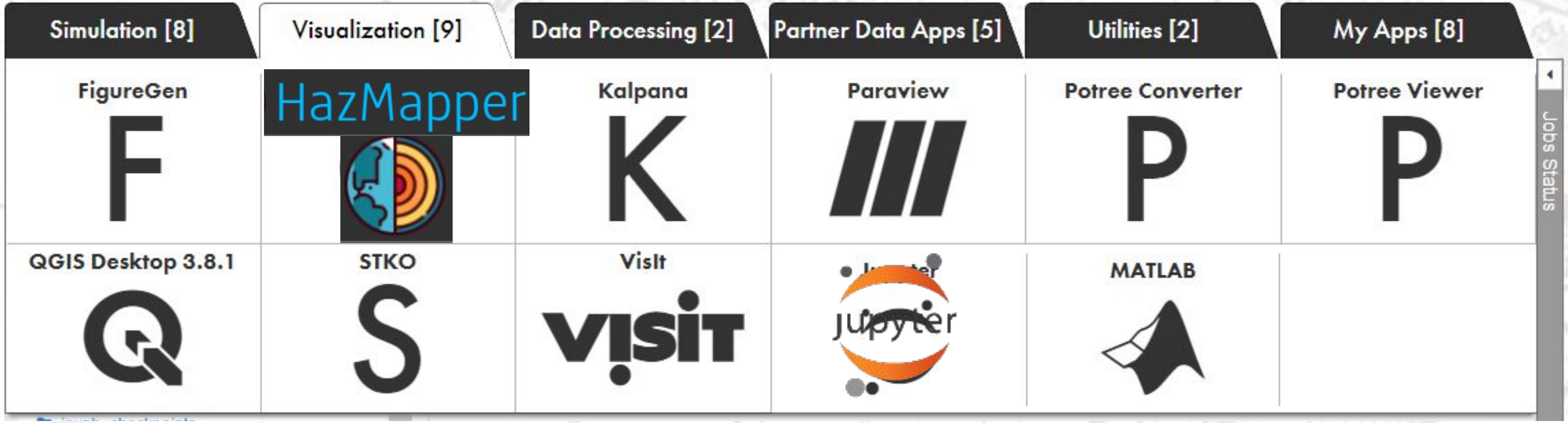
DESIGNSAFE-CI
NHERI: NATURAL HAZARDS ENGINEERING RESEARCH INFRASTRUCTURE
TEXAS The University of Texas at Austin
UCLA
TACC
RICE
Florida Tech

Tools & Apps: Simulation



- HPC-enabled simulation codes (Stampede2, Frontera)
- Available through portal or at the Command Line, easy access to HPC allocation (CPUs, GPUs) through DesignSafe

Tools & Apps: Data Analysis



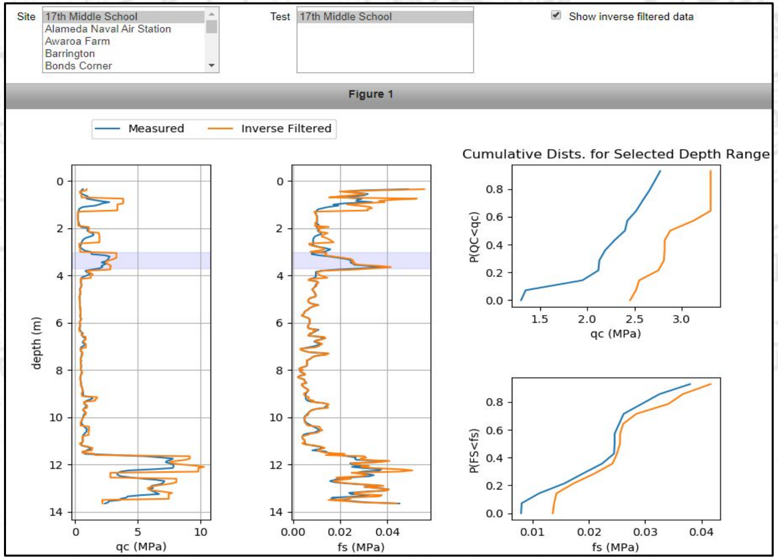
- Cloud-based tools for data analysis and visualization
- Access to files in Data Depot

Jupyter Notebooks

- Electronic notebooks in Python or R
- JupyterHub in DesignSafe
  – Access to Data Depot files
- Interactive data viewer
- Can write scripts for data processing, AI or machine learning
- Publish for use by others

jupyterhub

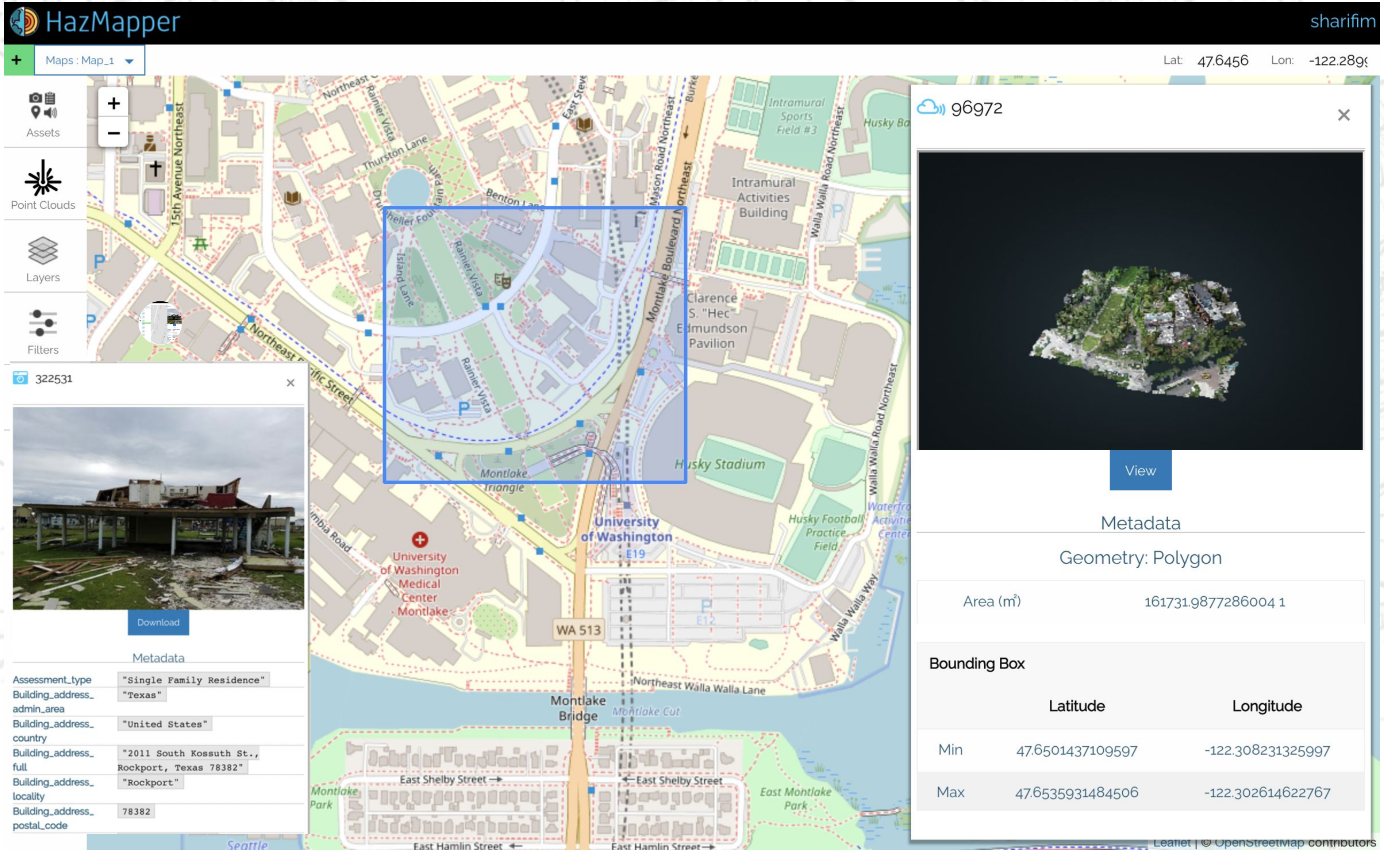
Next Generation Liquefaction



From Scott Brandenberg (UCLA)

HazMapper

- Easy access to images and point cloud data
- Location and preview exposed
- Link to Potree viewer provided
- Links to Streetview imagery (Mapillary)



Learning Center

DESIGNSAFE

Workspace | Learning Center

Education & Training

Archived DesignSafe Webinars

https://tinyurl.com/DesignSafe-Webinars

DesignSafe Tutorials

NEW **Development and utilization of a relational database to support post-earthquake building damage and recovery assessment**

March 12, 2021

- Watch Tutorial

Experimental Data Workflow for Real Time Decision Making Using Python and Jupyter Notebooks

February 3, 2021

- Watch Tutorial

Leveraging DesignSafe with TAPIS

December 17, 2020

- Watch Tutorial

Best Practices to Enhance the Quality, Discoverability and Re-Use Potential for Post-Event Reconnaissance Data

November 11, 2020

- Watch Tutorial
- Presentation Slides

SimCenter Webinars

NEW **Physical Modeling of Wave Attenuation & Wave Force Reduction by a Mangrove Forest**

December 4, 2020

- Watch Webinar

Computational Frameworks for the Implementation of Performance-Based Wind Engineering

November 23, 2020

- Watch Webinar

Partial Turbulence Simulation for Predicting Peak Wind Loads on Buildings

October 16, 2020

- Watch Webinar

Hurricane Loss Analysis for Single-Family Houses Considering Current and Changing Climate Conditions

October 6, 2020

- Watch Webinar

DesignSafe: We are here for you!

Available to the Global Natural Hazards Research Community

Analyze, Simulate & Share | Start Your Experiments | Expand Your Skills | Join the Community

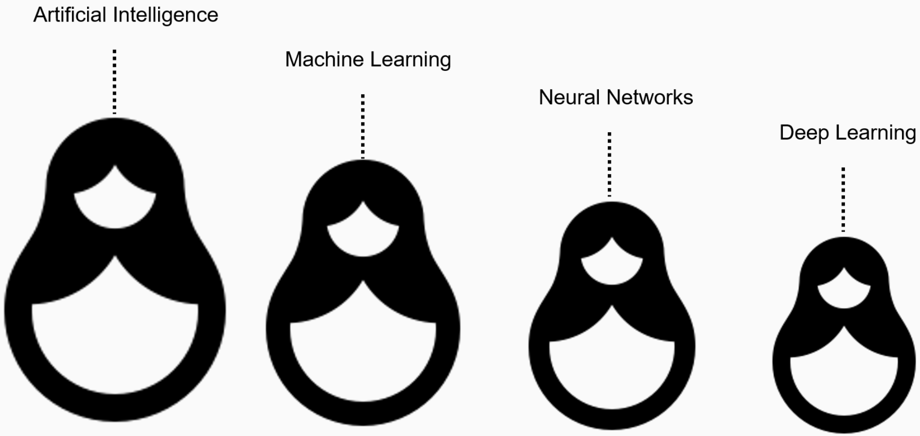
- Interact with us and the community using the DesignSafe Slack team
- Cite data using DOIs in your reference list!

MAKE DATA COUNT

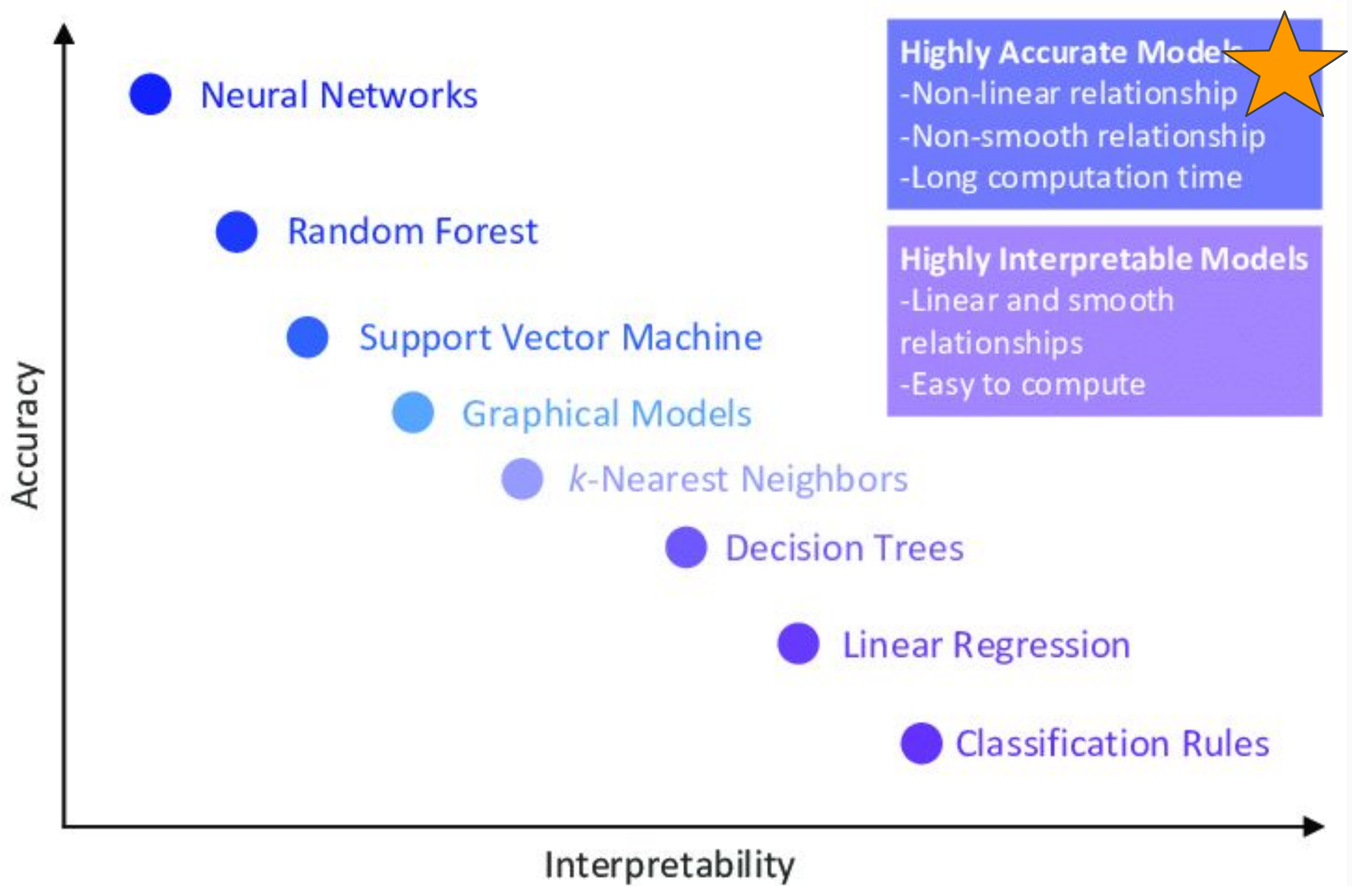
Please share your feedback, ideas, experiences!

AI vs Machine Learning vs Deep Learning

- **Artificial intelligence**: build intelligent programs and machines that can creatively solve problems, which has always been considered a human prerogative.
- **Machine learning** is a subset of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. In ML, there are different algorithms (e.g. neural networks) that help to solve problems.
- **Deep learning** is a subset of machine learning, which uses the neural networks to analyze different factors with a structure that is similar to the human neural system.



Interpretability vs accuracy



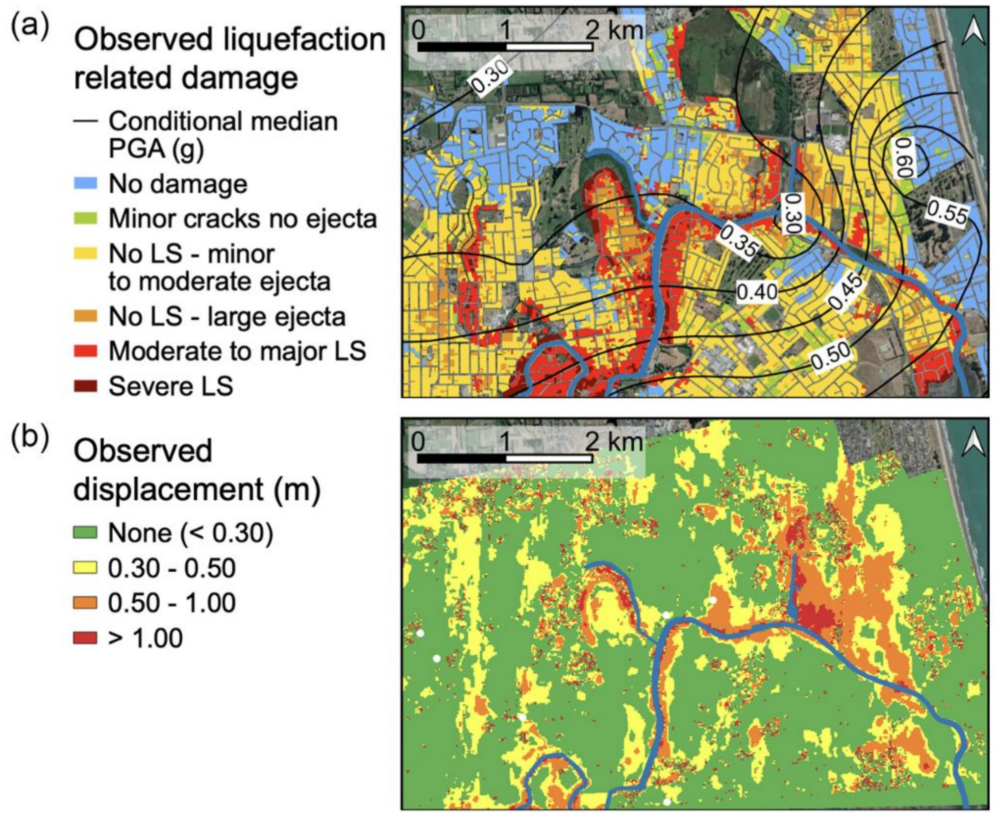
Lateral Spreading

Decision Trees

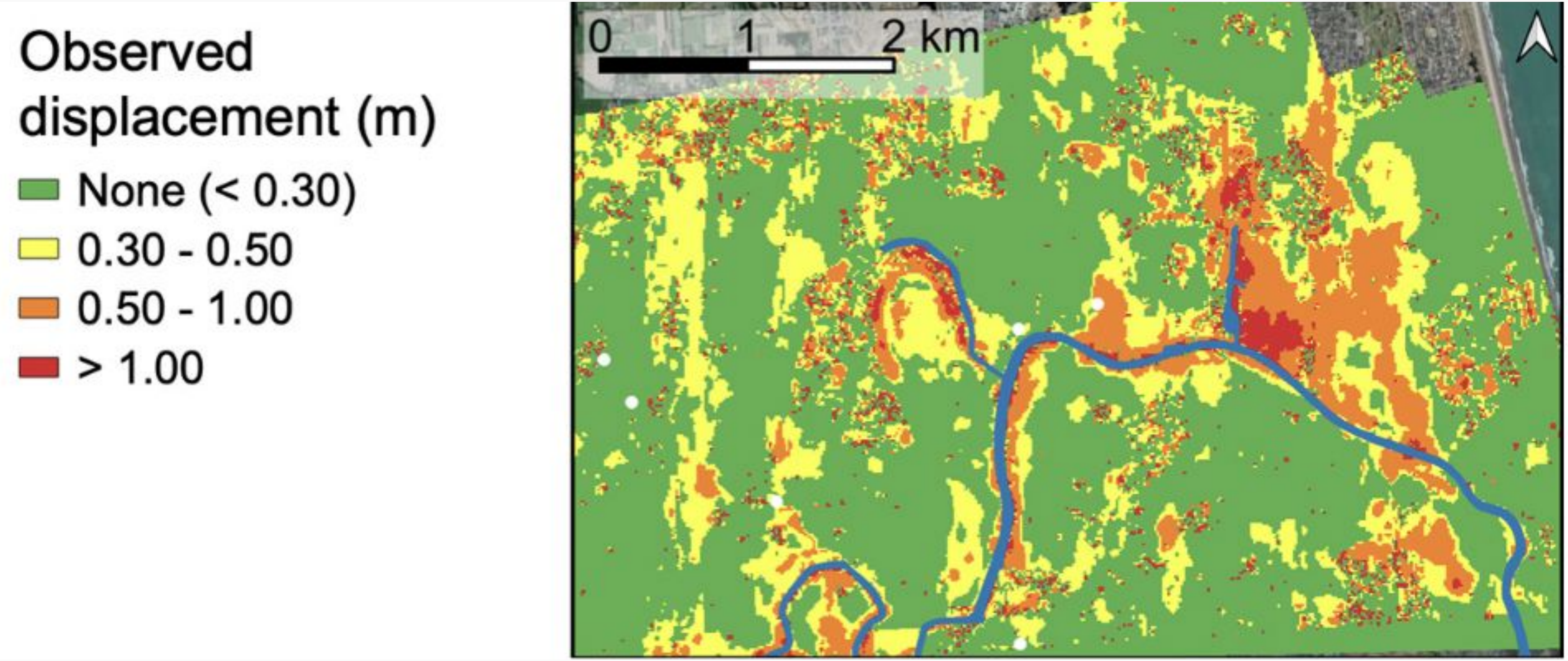
&

Random Forest

Liquefaction in NZ



Predicting lateral spreading in NZ



Durante and Rathje, 2021

Predicting lateral spreading in NZ

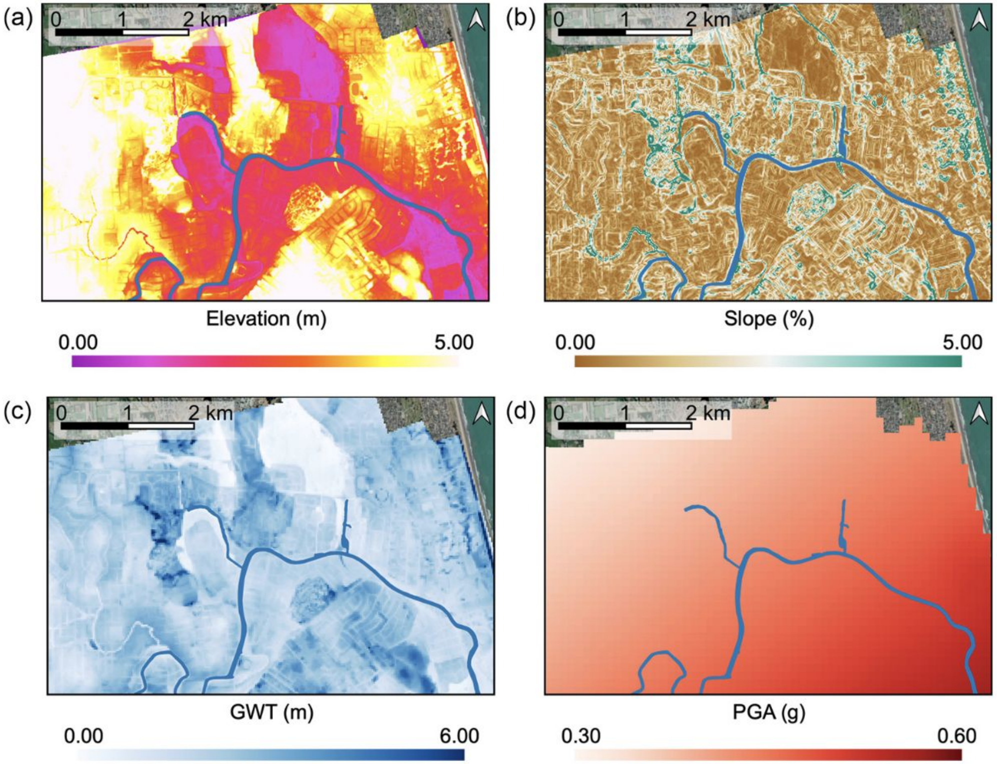
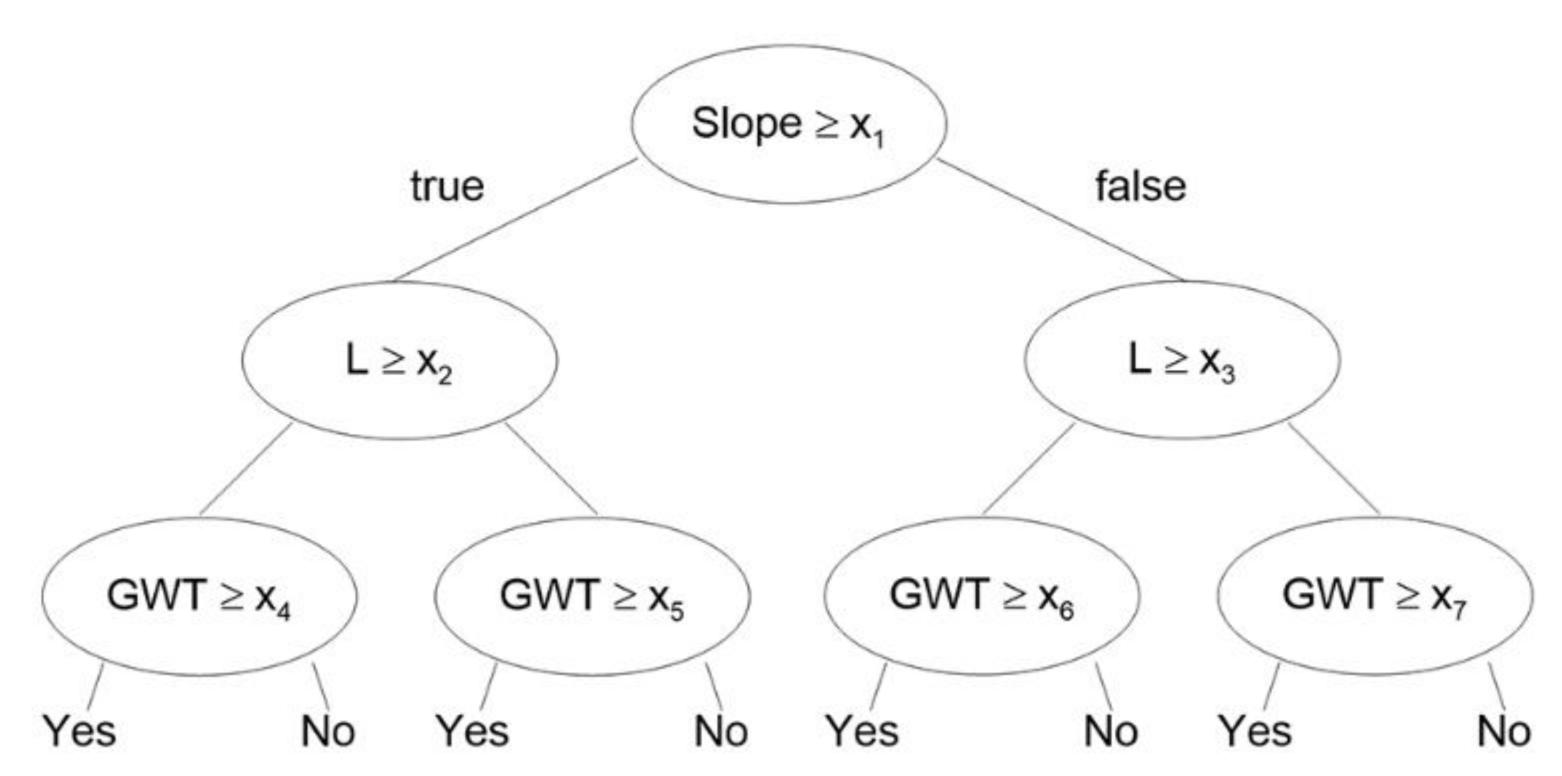


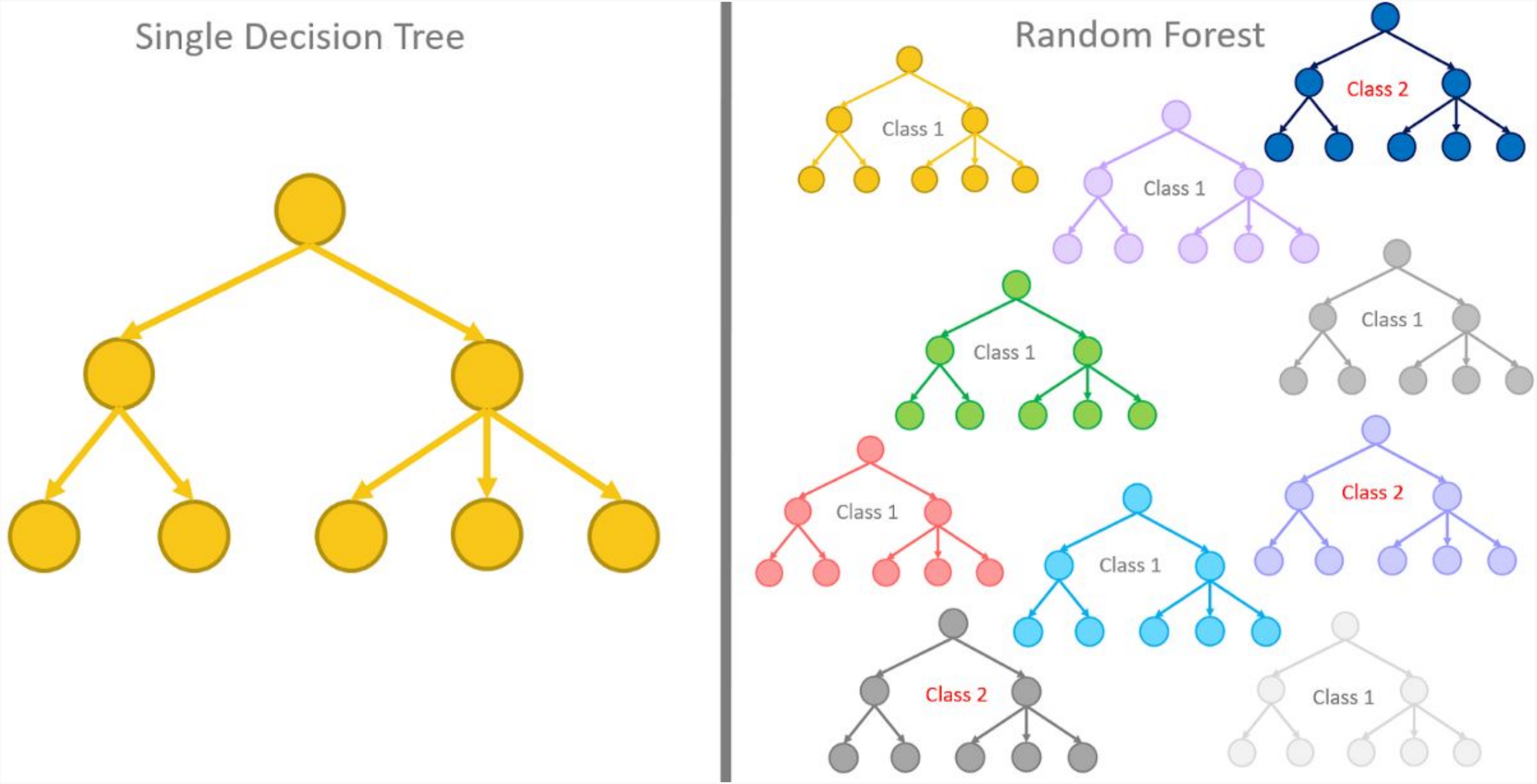
Figure 4. Spatial distribution of geometric and event-specific input features considered in ML models: (a) ground elevation, (b) ground slope, (c) GWT depth, and (d) PGA.

Durante and Rathje, 2021

Classification - Decision Tree



Random Forest



Random Forest model

Table 3. Summary of features used in each RF model analyzed

		Feature					
		L (m)	GWT (m)	Slope (%)	PGA (g)	Elevation (m)	CPT data
No CPT data	Model 0	✓	✓	✓	O	O	O
	Model 1	✓	✓	✓	✓	O	O
	Model 2	✓	✓	✓	O	✓	O
	Model 3	✓	✓	✓	✓	✓	O
CPT data	Model 4	✓	✓	✓	✓	O	✓
	Model 5	✓	✓	✓	✓	✓	✓

GWT: ground water table; PGA: peak ground acceleration; CPT: cone penetration tests.

Durante and Rathje, 2021

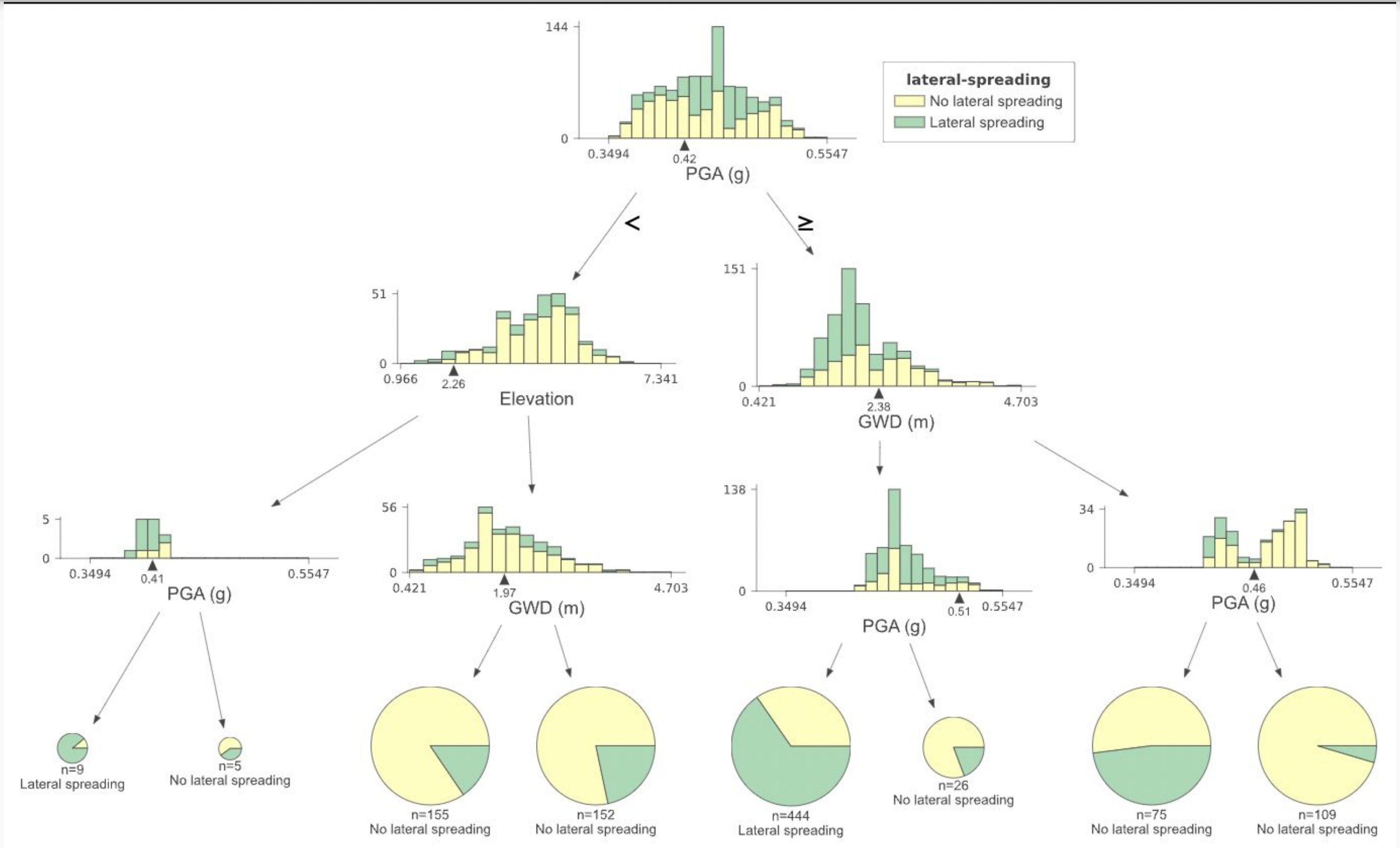
Random Forest - Performance

Table 4. Evaluation metrics for RF models for Yes/No and displacement classification problems

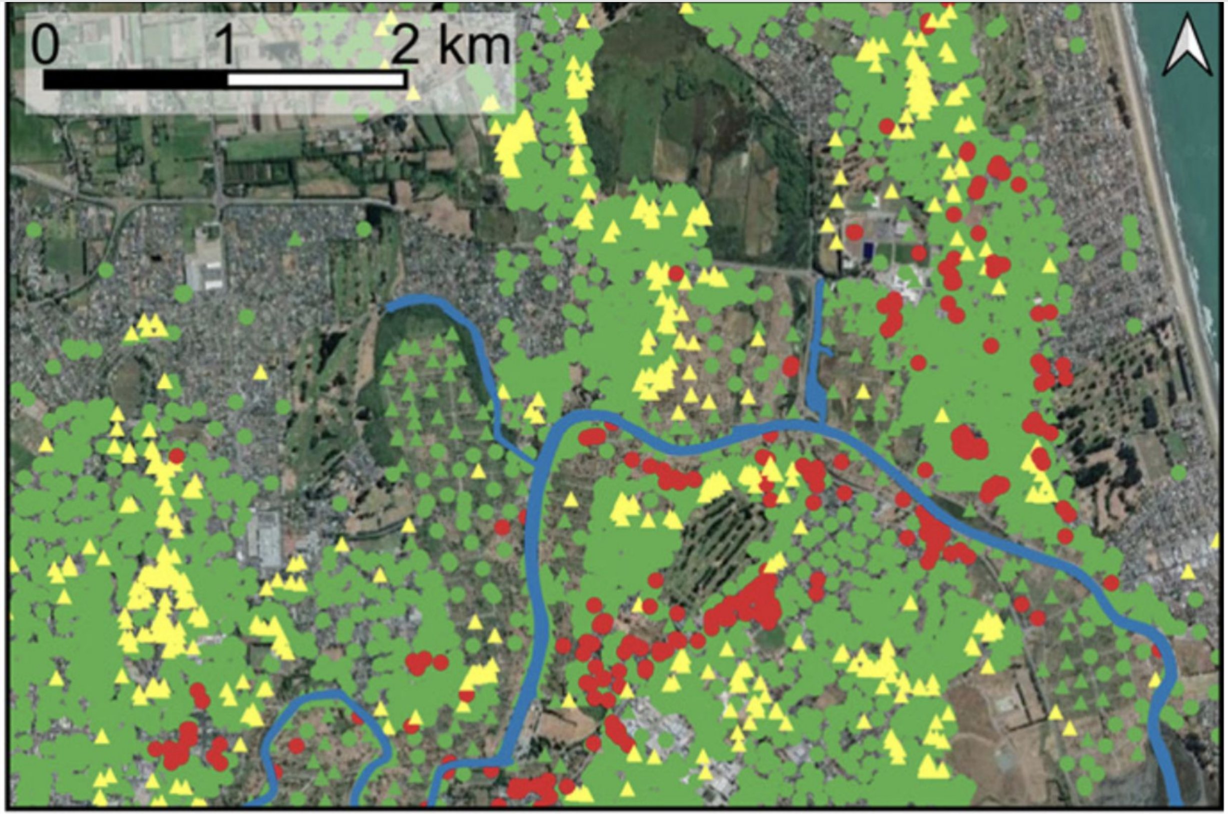
		No CPT data				CPT data	
		Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Y/N	Accuracy (Overall)	0.76	0.85	0.82	0.88	0.85	0.87
	Recall—Yes	0.58	0.75	0.74	0.82	0.76	0.78
	Recall—No	0.90	0.92	0.88	0.93	0.92	0.93
	Precision—Yes	0.80	0.88	0.82	0.89	0.87	0.89
	Precision—No	0.75	0.84	0.83	0.87	0.84	0.85
	ROC AUC	0.76	0.88	0.84	0.90	0.86	0.88
DISPL	Accuracy (Overall)	0.78	0.88	0.86	0.89	0.88	0.89
	Recall—Class 0	0.88	0.91	0.93	0.92	0.91	0.91
	Recall—Class 1	0.67	0.87	0.79	0.89	0.88	0.90
	Recall—Class 2	0.27	0.42	0.33	0.42	0.53	0.49
	Precision—Class 0	0.79	0.90	0.86	0.92	0.90	0.92
	Precision—Class 1	0.74	0.84	0.84	0.85	0.85	0.84
	Precision—Class 2	0.97	0.98	0.97	0.94	0.95	0.93
	ROC AUC	0.82	0.89	0.87	0.90	0.86	0.88

CPT: cone penetration tests; ROC: receiver operating characteristic; AUC: area under the ROC curve; Y/N: Yes/No Classification problem; DISP: Class Displacement Classification problem.

Visualizing the first tree in RF



Random forest prediction of lateral spreading



Durante and Rathje, 2021

▲ True Positive (TP) ● True Negative (TN) ● False Positive (FP) ▲ False Negative (FN)

SHAP Value: Contribution of each variable

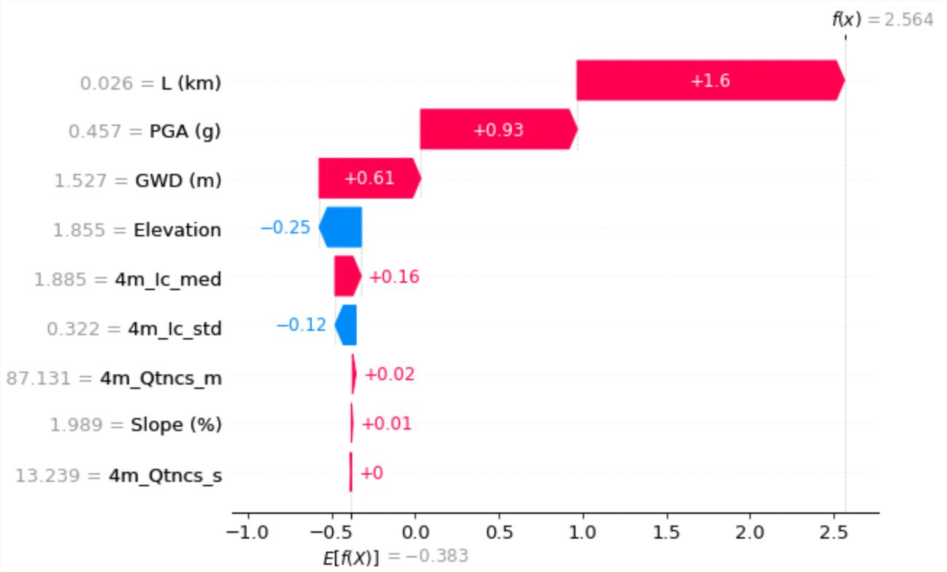
$$\text{A: } v(\{A\}) - v(\{\}) = 10 - 0 = 10$$
$$\text{B: } v(\{A,B\}) - v(\{A\}) = 60 - 10 = 50$$
$$\text{C: } v(\{A,B,C\}) - v(\{A,B\}) = 100 - 60 = 40$$

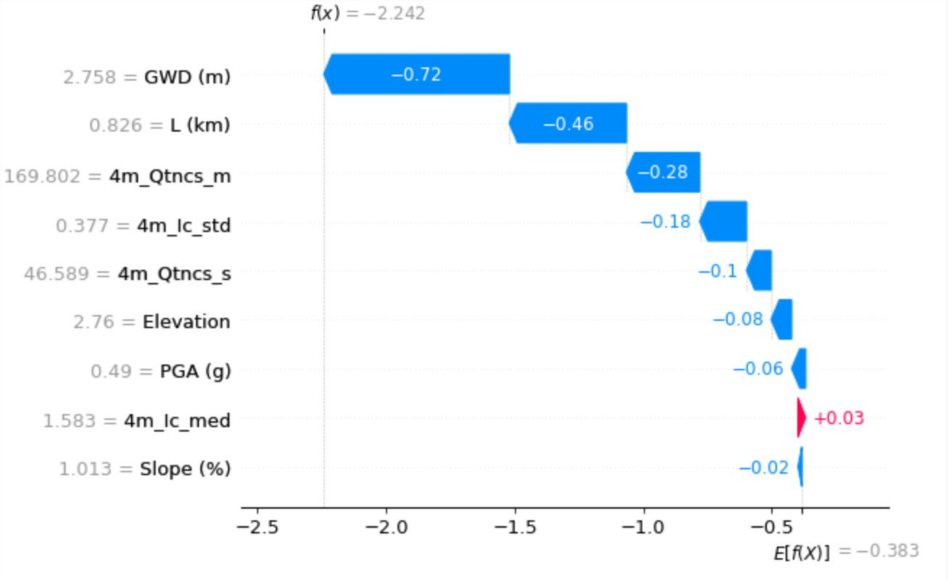
$$v(\{\}) = 0$$
$$v(\{A\}) = 10$$
$$v(\{B\}) = 20$$
$$v(\{C\}) = 30$$
$$v(\{A,B\}) = 60$$
$$v(\{B,C\}) = 70$$
$$v(\{A,C\}) = 90$$
$$v(\{A,B,C\}) = 100$$

$$\{\} \rightarrow \{A\} \rightarrow \{A,\ B\} \rightarrow \{A,\ B,\ C\} \,\|\, A = 10,\ B = 50,\ C = 40$$
$$\{\} \rightarrow \{A\} \rightarrow \{A,\ C\} \rightarrow \{A,\ B,\ C\} \,\|\, A = 10,\ B = 10,\ C = 80$$
$$\{\} \rightarrow \{B\} \rightarrow \{A,\ B\} \rightarrow \{A,\ B,\ C\} \,\|\, A = 40,\ B = 20,\ C = 40$$
$$\{\} \rightarrow \{B\} \rightarrow \{B,\ C\} \rightarrow \{A,\ B,\ C\} \,\|\, A = 30,\ B = 20,\ C = 50$$
$$\{\} \rightarrow \{C\} \rightarrow \{B,\ C\} \rightarrow \{A,\ B,\ C\} \,\|\, A = 30,\ B = 40,\ C = 30$$
$$\{\} \rightarrow \{C\} \rightarrow \{A,\ C\} \rightarrow \{A,\ B,\ C\} \,\|\, A = 60,\ B = 10,\ C = 30$$

$$A_{avg} = 30,\ B_{avg} = 25,\ C_{avg} = 45$$

Random Forest post-hoc explanation

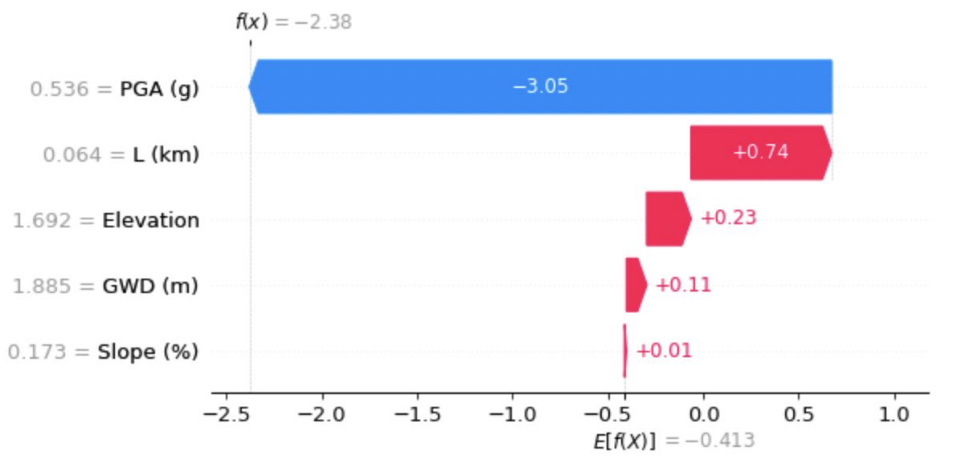


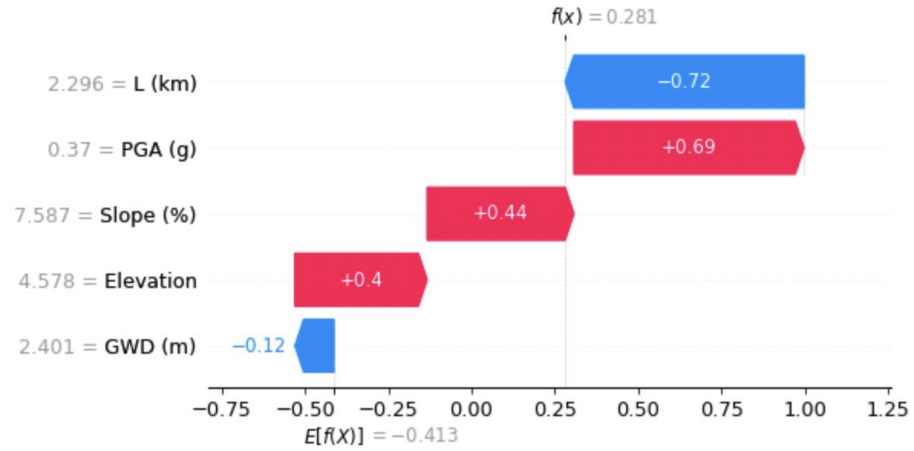


Lateral spreading - Prob (Lat spread) = 90%

No lateral spreading - Prob (Lat spread) = 7%

Incorrect learning

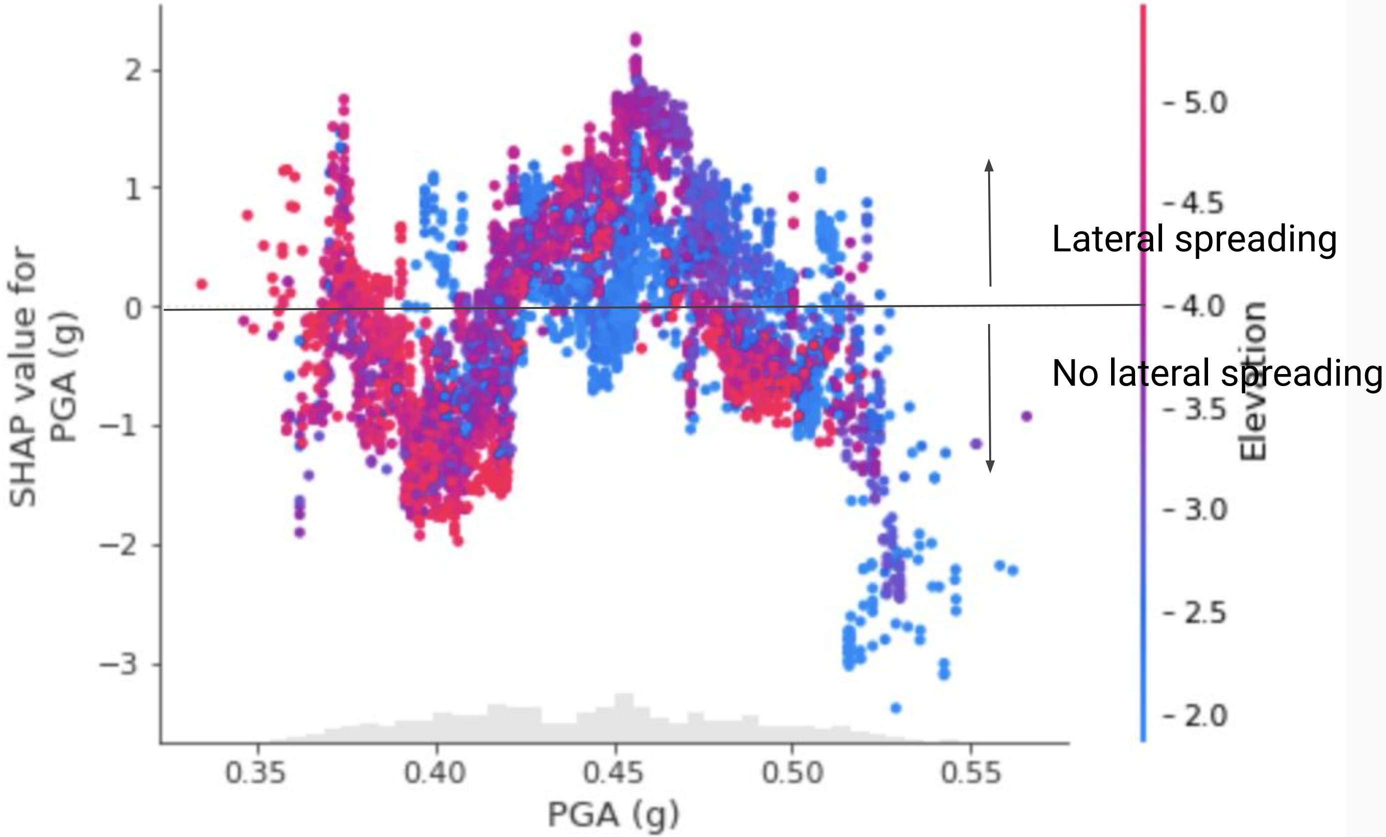




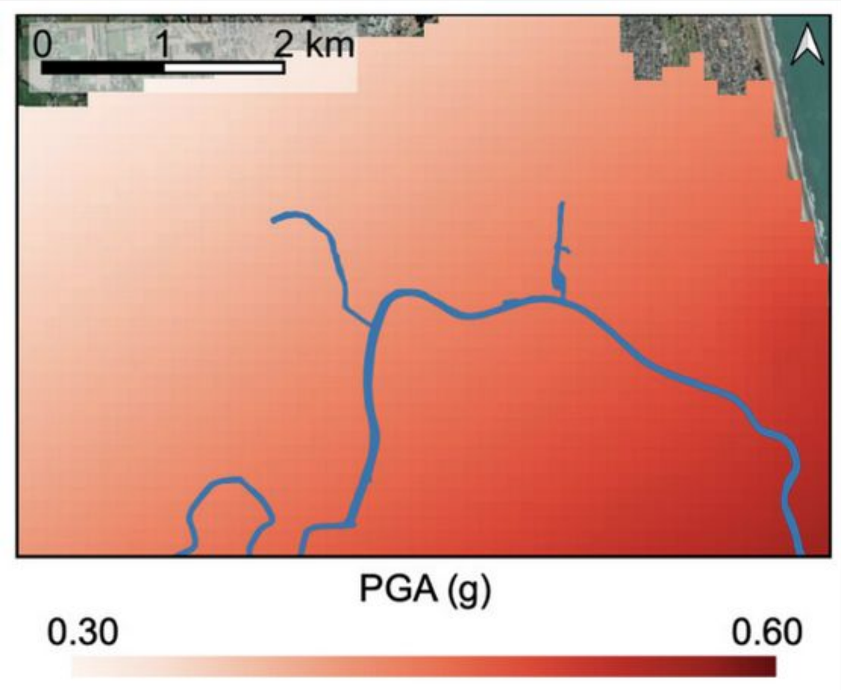
Incorrectly predicting no lateral spreading
Prob (Lat spread) = 8%

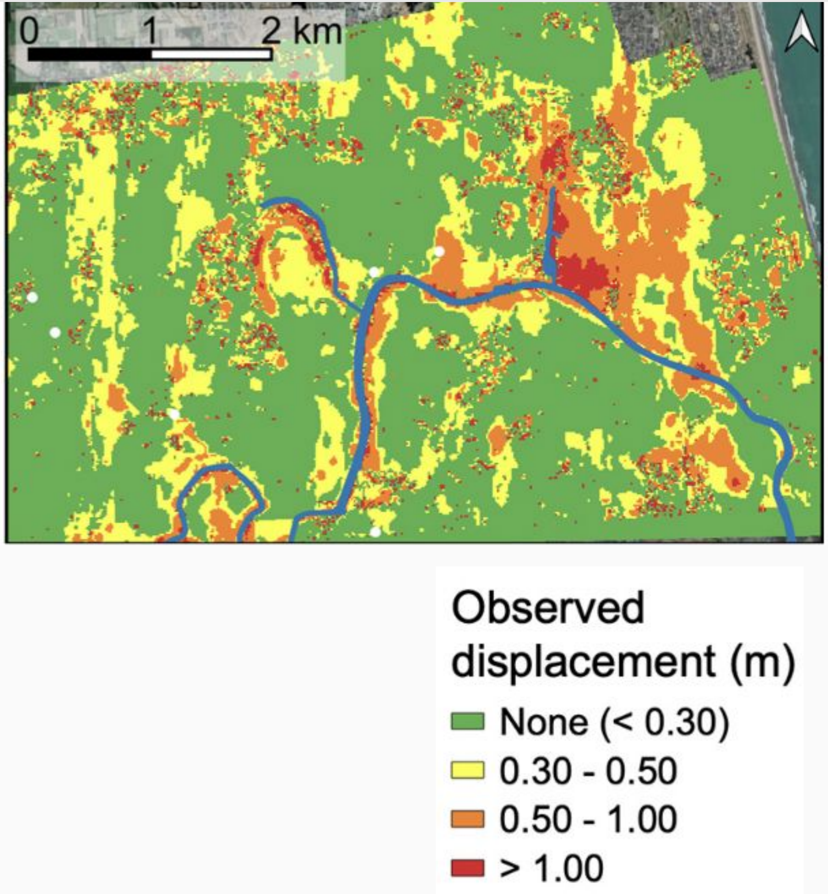
Incorrectly predicting lateral spreading
Prob (Lat spread) = 53%

Why PGA has a bad influence?

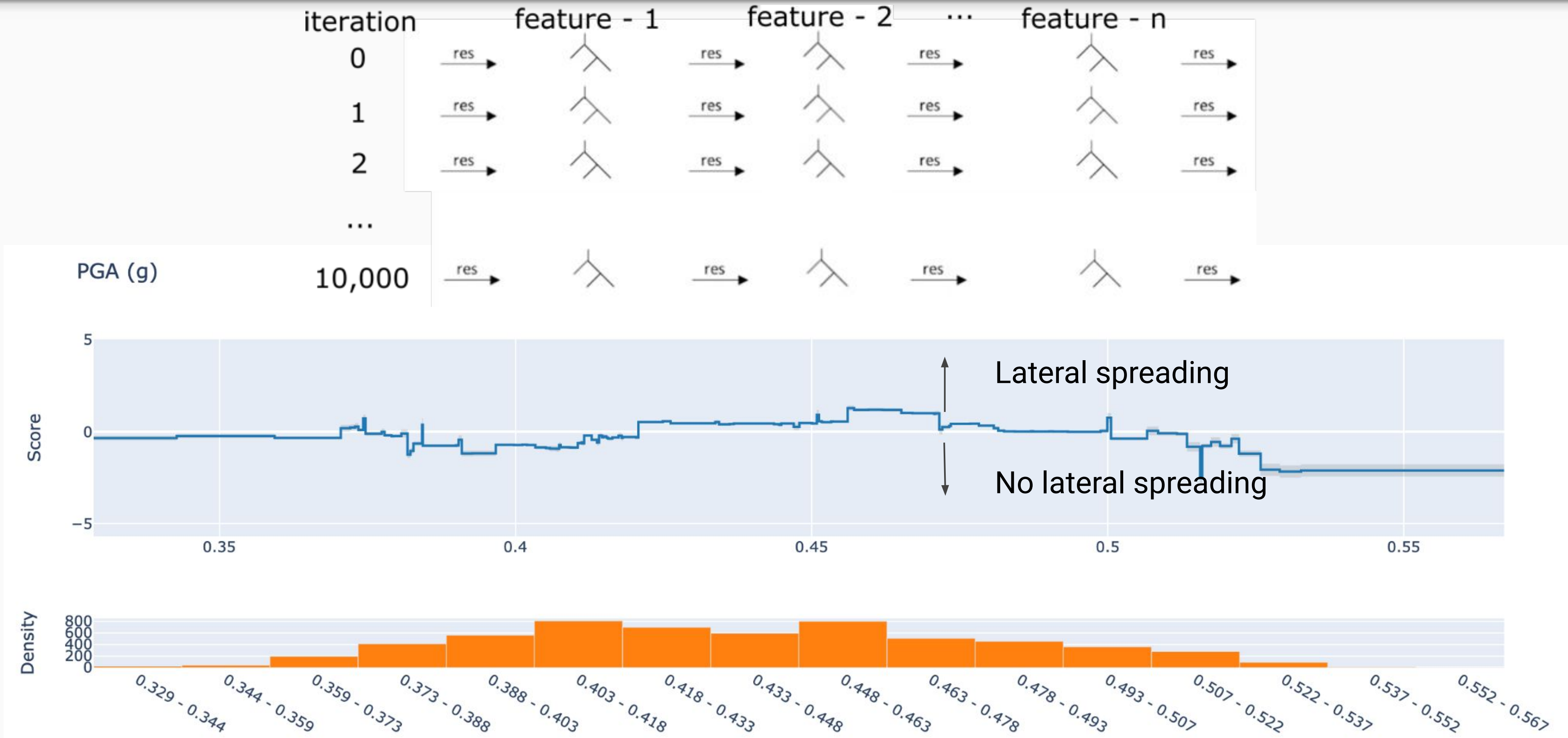


Bad learning of PGA relation

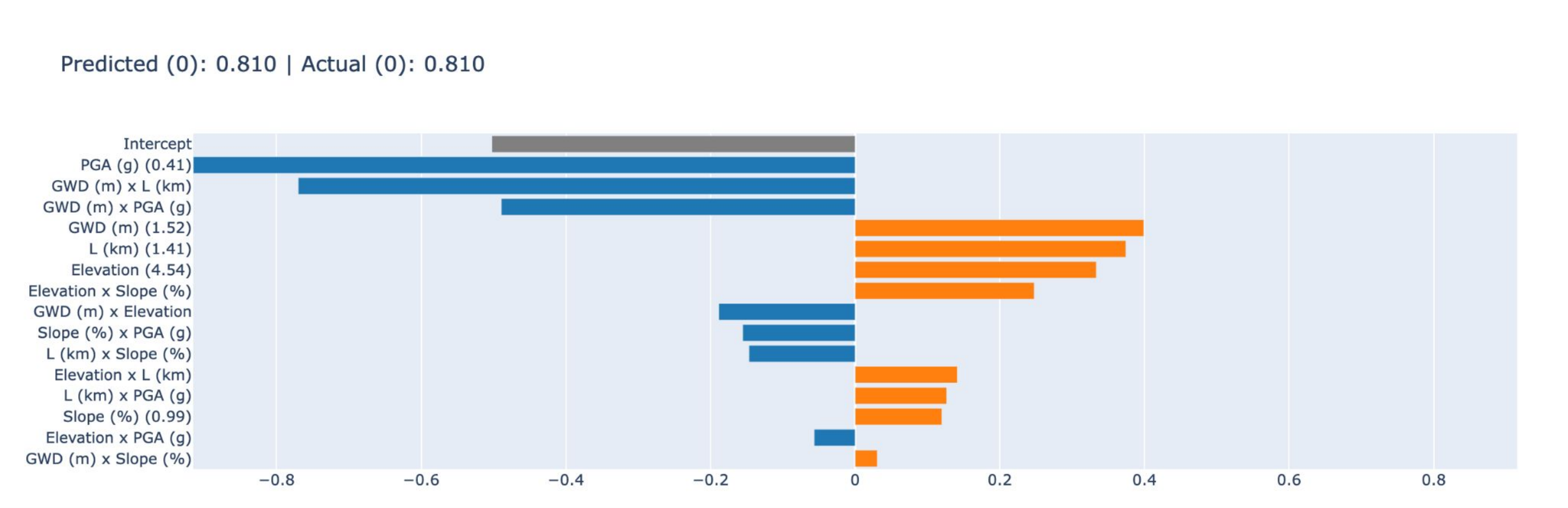




GlassBox models - Physics Informed (PIX) Trees



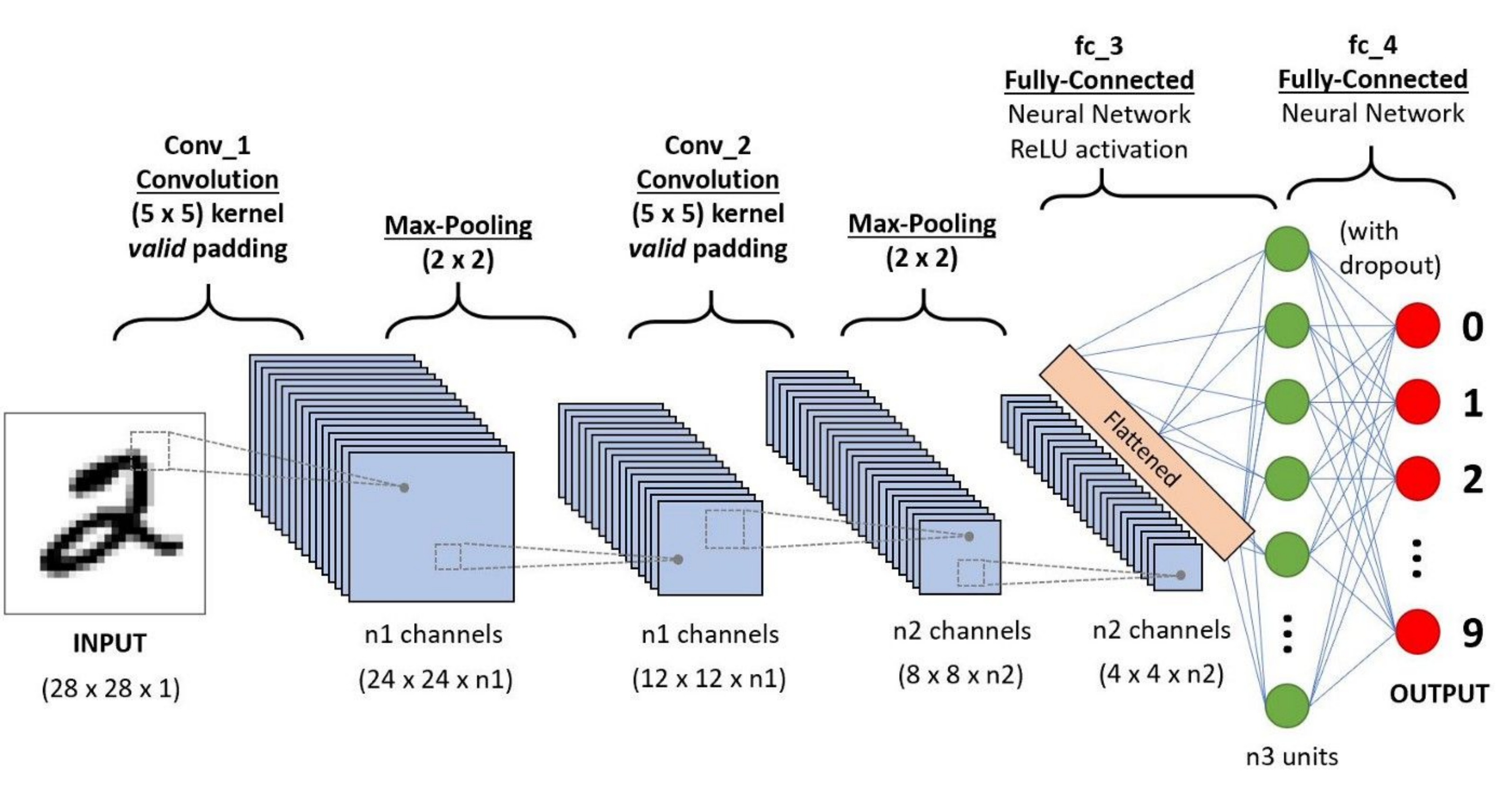
GlassBox models - PIX Trees



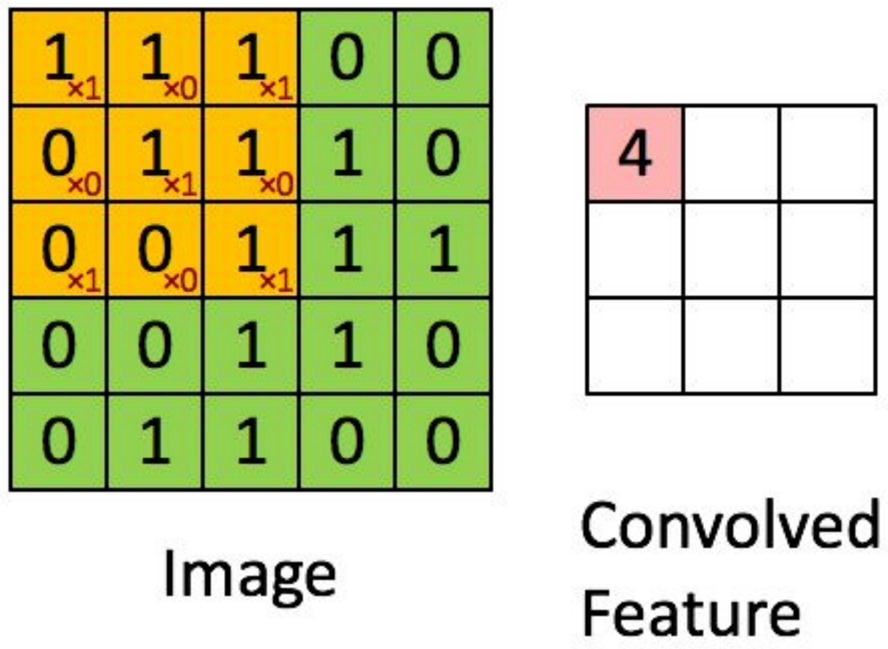
Convolutional Neural Network (CNN)

Full-Waveform Inversion

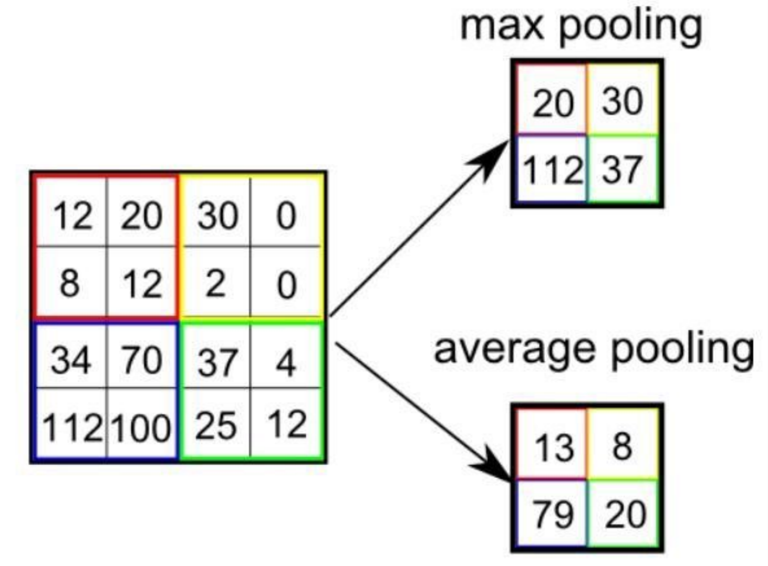
Convolutional Neural Network



CNN: Kernel



Convolutional Layer



Pooling Layer

Building damage classifier - an explanation

Undamaged

Partially-damaged

Completely-destroyed

Explaining a building image classifier

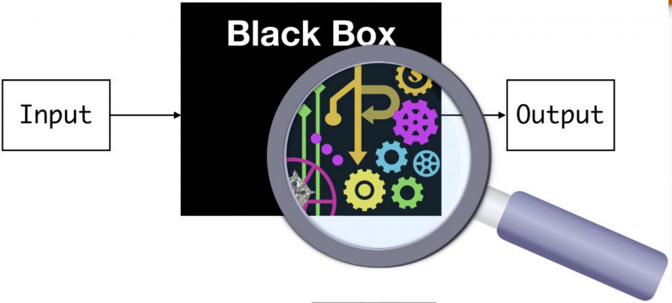
Sky

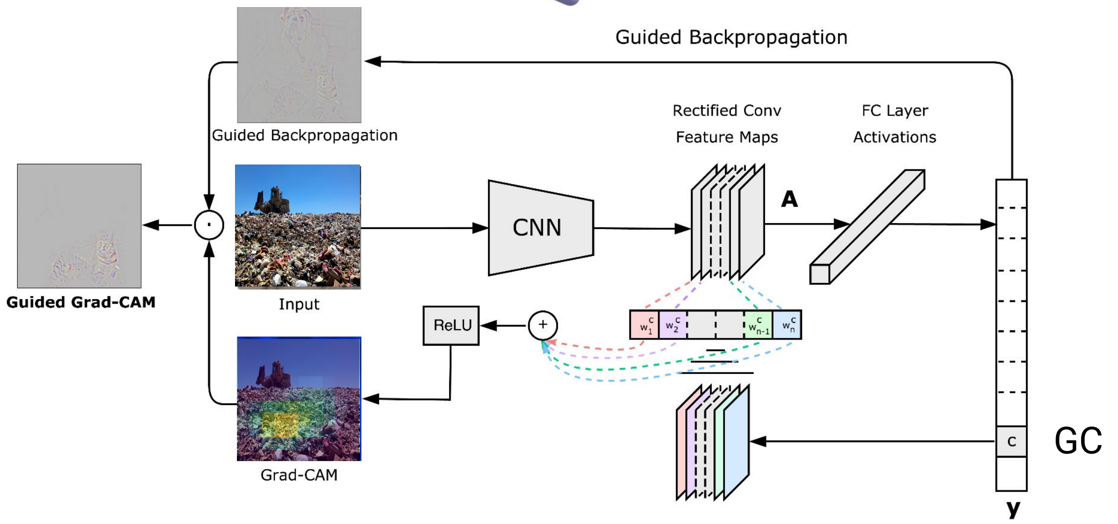
rubble

Predicted class: General Collapse

Transfer learning prediction: Ants

GradCAM Explanation





http://gradcam.cloudcv.org/

Post-hoc explanations - LIME Explainer

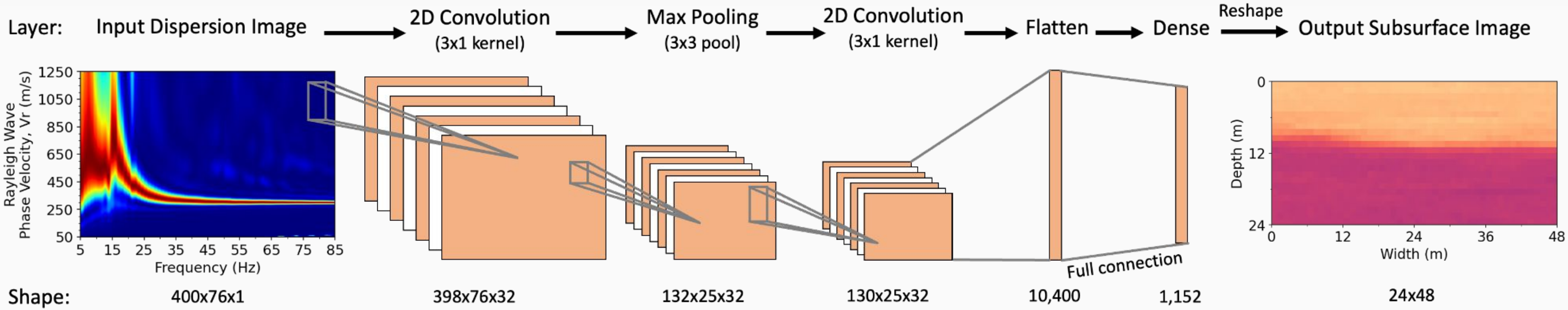
Original Image

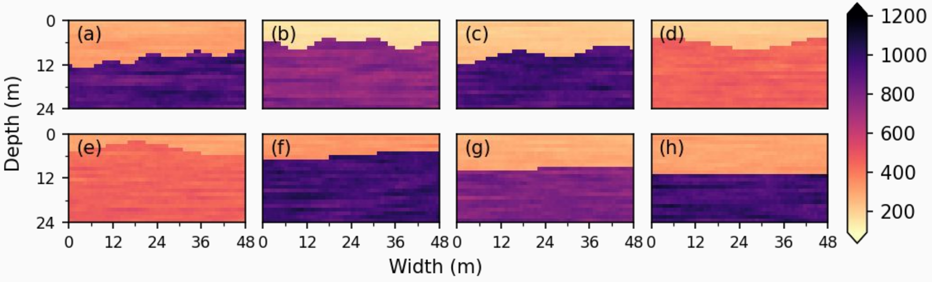
Tree Frog (P = 0.54)

Interpretable Components

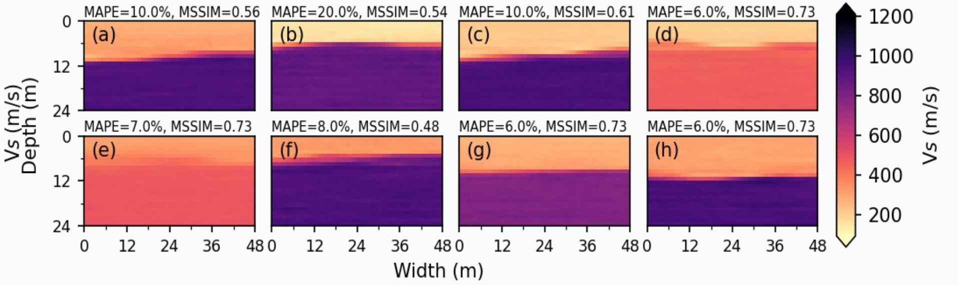
Perturbed Instances	P(tree frog)
	0.85
	0.00001
	0.52

Convolutional Neural Network for inverse problem



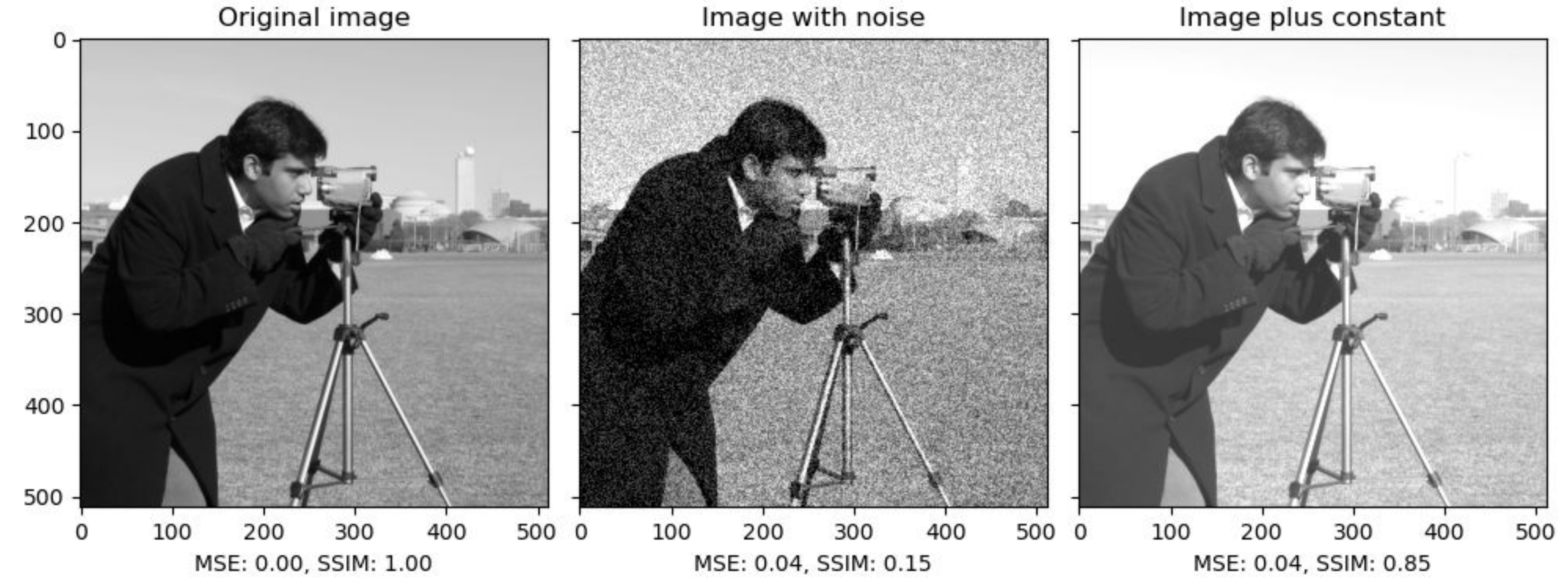


True Model



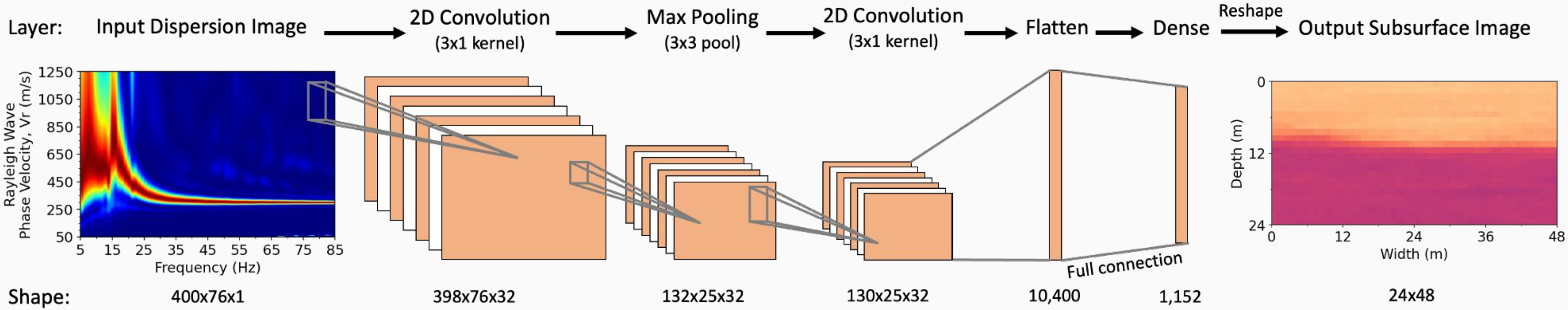
Predictions

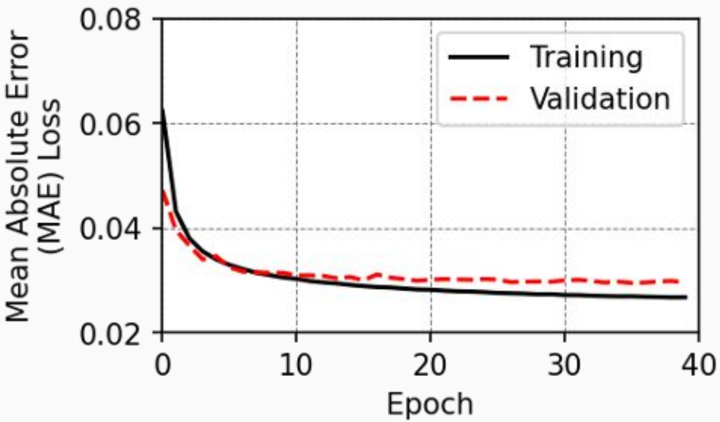
Structural Similarity Index



https://scikit-image.org/docs/stable/auto_examples/transform/plot_ssim.html

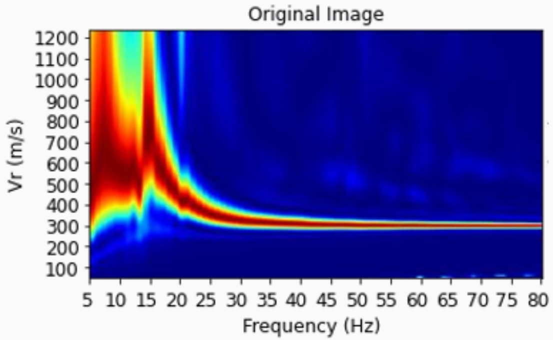
How to design a CNN





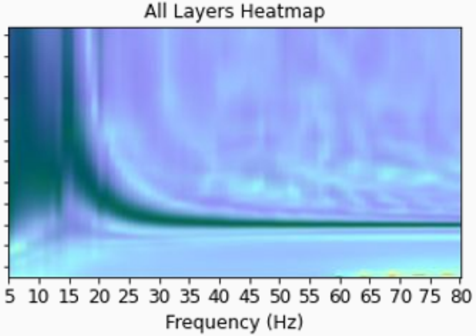
MAPE	**Kernel Size: 3x1**	**Kernel Size: 3x3**
Deep (5)	11.3 %	**8.5 %**
Shallow (3)	12.1 %	11.2 %

Designing a neural network with explainability



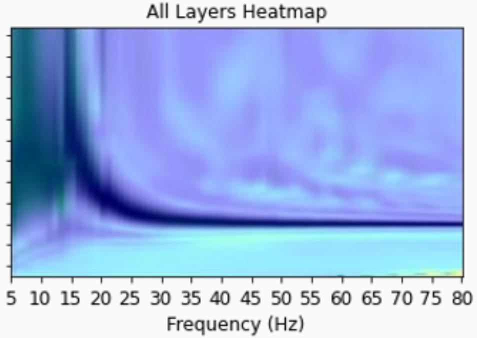
3x1 kernel

deep architecture



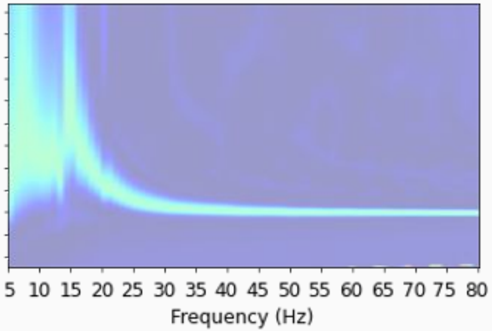
3x3 kernel

deep architecture



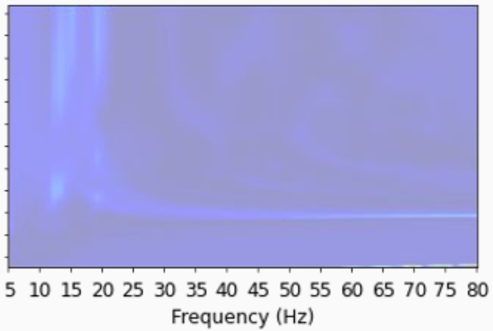
3x1 kernel

shallow architecture

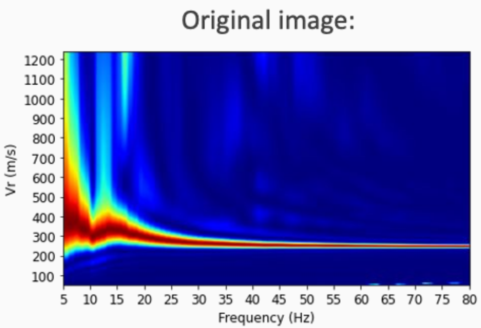


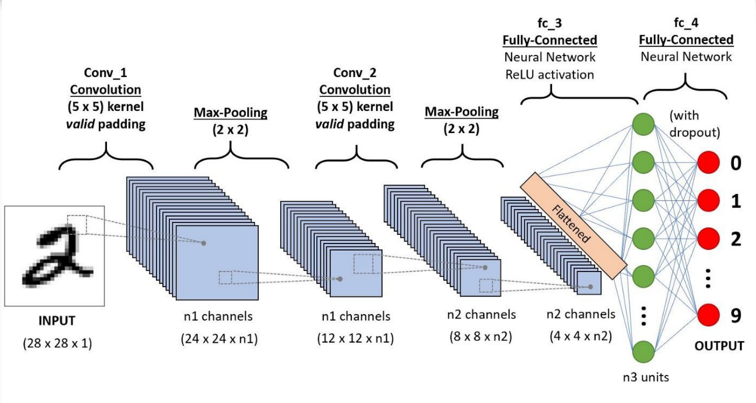
3x3 kernel

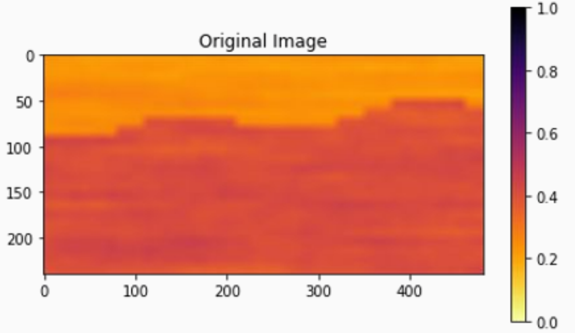
shallow architecture



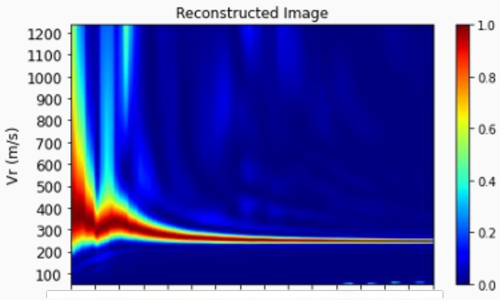
Inverting a neural network - Iteratively

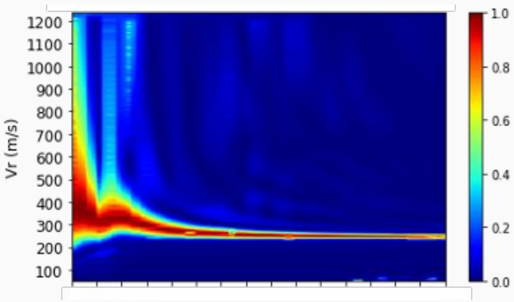


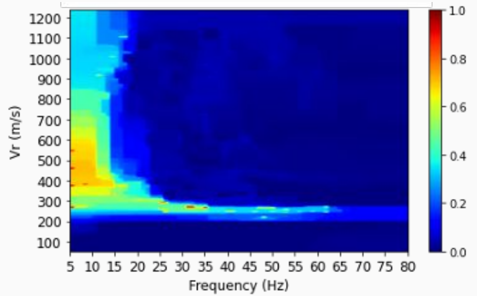




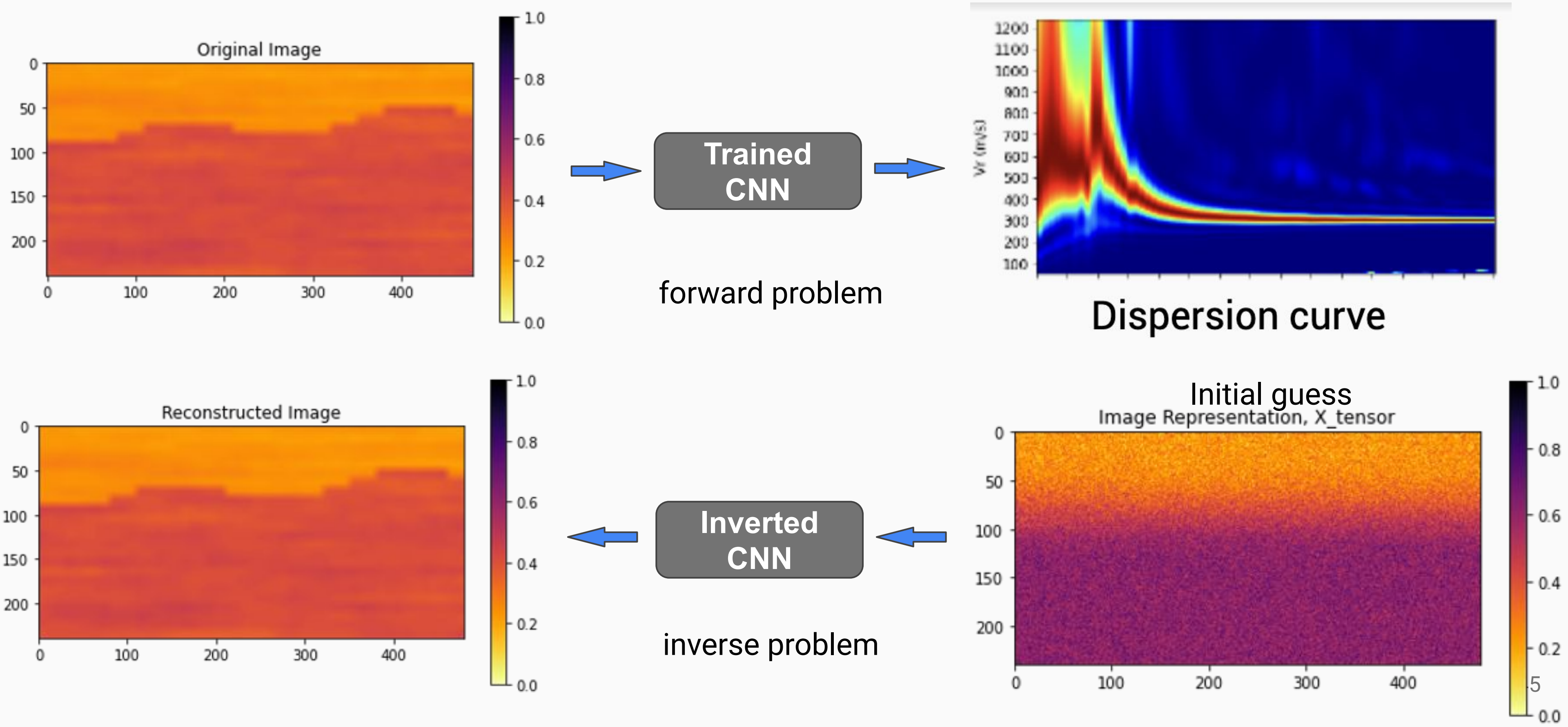
Deeper layers







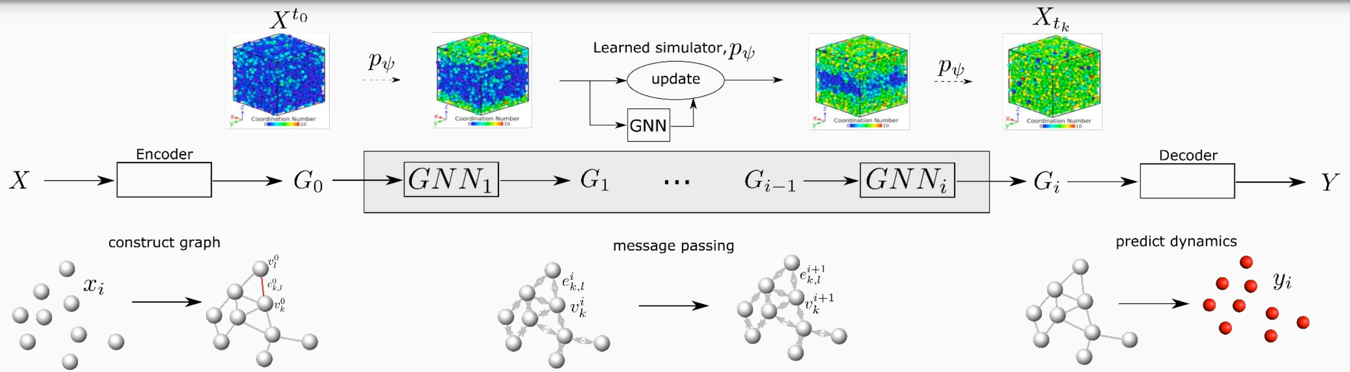
Using inversion to solve the inverse problem

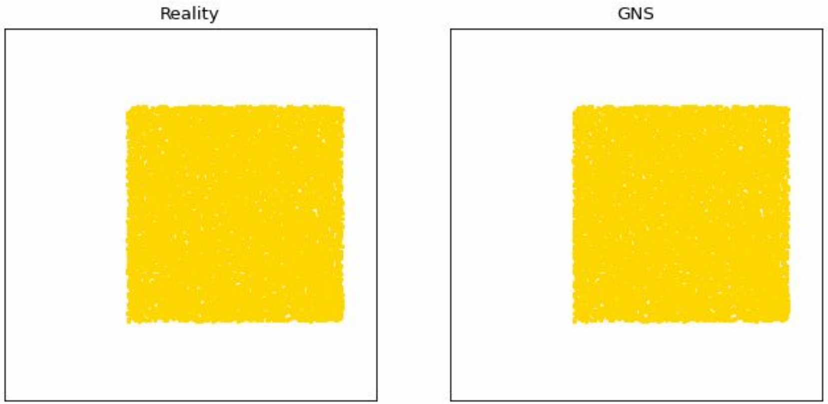


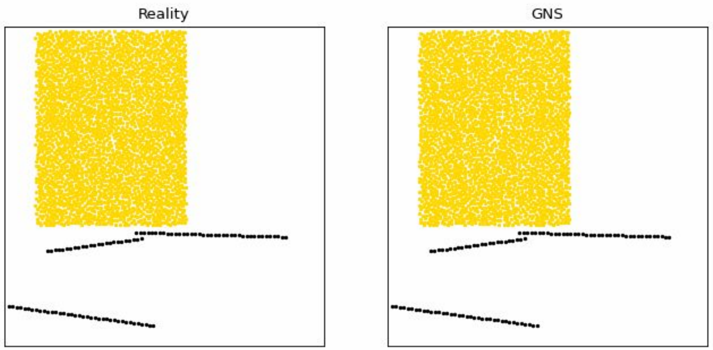
Future of Physics simulations

Knowledge Discovery and GNS

ML prediction of granular flows



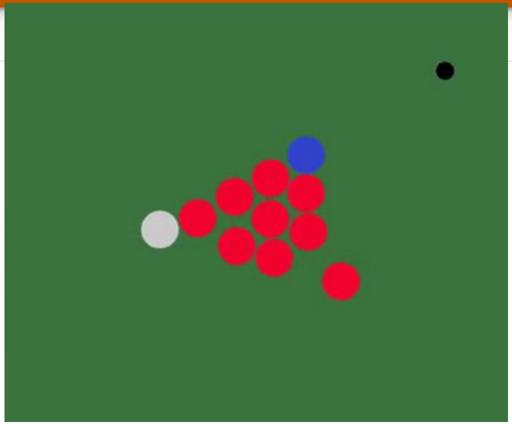




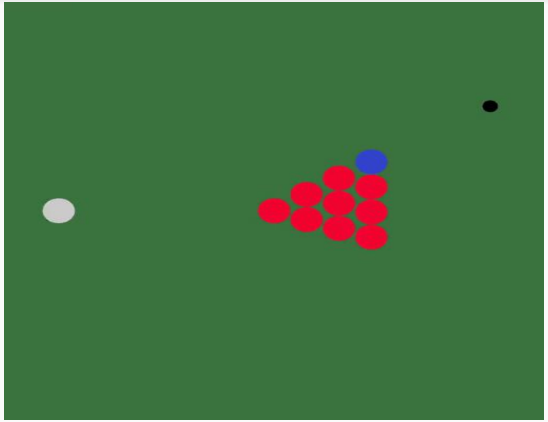
2.5 hours

20 s

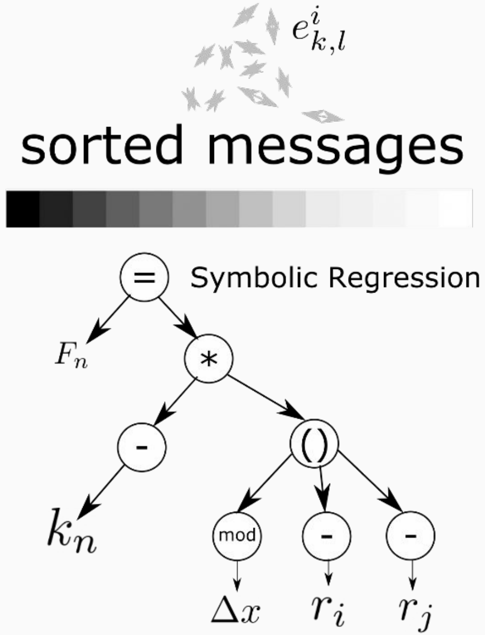
Knowledge discovery with AI

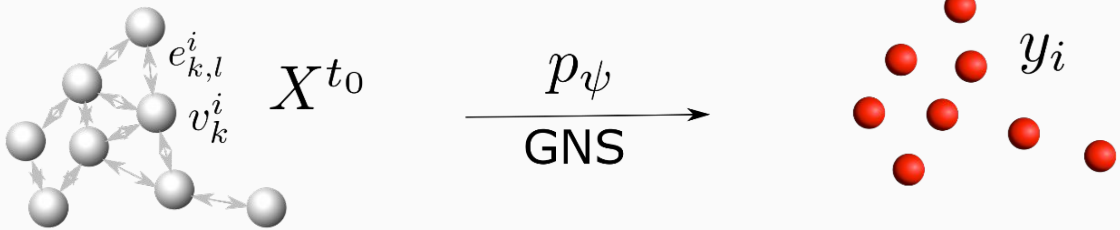


initial state



predicted dynamics





$$F_n = -k_n * (|\Delta x| - r_i - r_j)$$

Deriving analytical eq. linear spring with AI

Eq.	Derived equation	MSE	C_x	D_a
1	-198.72363	64747.52	1	Y
2	$(\Delta x + -198.92792)$	63938.996	3	N
3	$(-203.1408 + exp(\Delta x))$	60907.68	4	N
4	$((\Delta x + -2.3484528) * 92.79602)$	27031.744	5	Y
5	$((\Delta x * (\mathbf{X}_1 + 92.75565)) + -218.16481)$	26830.58	7	N
6	$((\Delta x + (abs(r_1) * -1.1491286)) * 100.23312)$	21227.312	8	Y
7	$((\Delta x + ((abs(r_1) + 1.1013538) * -0.8038518)) * 98.86028)$	18721.219	10	Y
8*	$\boldsymbol{((\Delta x + (\mathbf{abs}((r_2 * -1.0) + r_1) * -1.0)) * 100.0)}$	**3.76E-10**	**12**	**Y**
9	$((\Delta x + (abs((r_2 * inv(-1.0)) + r_1) * -1.0)) * 99.9998)$	3.01E-10	15	N

C_x is the complexity and D_a represents if the expression passes dimensional analysis.
$*$ denotes the chosen solution.

$$F_n = k_n * abs(\Delta x - r_i - r_j)$$


Hook 'em Horns!
https://geoelements.org
This work was partially funded by NSF OAC: #2103937