Task 2: Beating the Data Limit: Detecting Very Small CO_2 Signature in Low-Data Regime (FWP-FE-112-19-FY19)

Youzuo Lin

Earth and Environmental Sciences Division

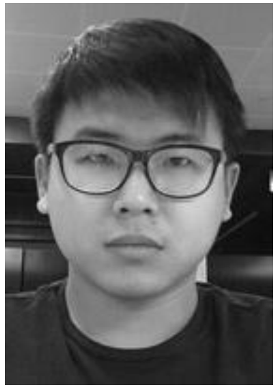
Los Alamos National Laboratory

U.S. Department of Energy

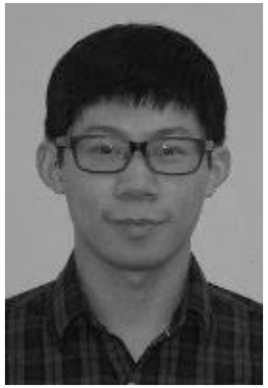
National Energy Technology Laboratory
Carbon Management Project Review Meeting
August 15 - 19, 2022

CONTENTS

Meet Our Team



Hanchen Wang
Micro-seismic Monitoring



Shihang Feng
Seismic Inversion & ML

Bailian Chen
CCUS & ML

Neill Symons
Seismology

Rajesh Pawar
Carbon Sequestration

George Guthrie
Carbon Sequestration

Students

- Will Reichard-Flynn, Post-Master Student, Seismology, Los Alamos National Laboratory
- Yinan Feng, Post-Master Student, Machine Learning, Los Alamos National Laboratory
- Yanhua Liu, Ph.D. Student, Center for Wave Phenomena, Colorado School of Mines
- Daniel Manu, Ph.D. Student, Dept of Electrical & Computer Engineering, University of New Mexico
- Xitong Zhang, Ph.D. Student, Dept of Computational, Mathematics, Science and Engineering, Michigan State University
- Peng Jin, Ph.D. Student, College of Information Sciences and Technology, Penn State University

Problem Statement and Technical Objective

<u>Objective</u>: Capture very small CO_2 or brine leakage over large area

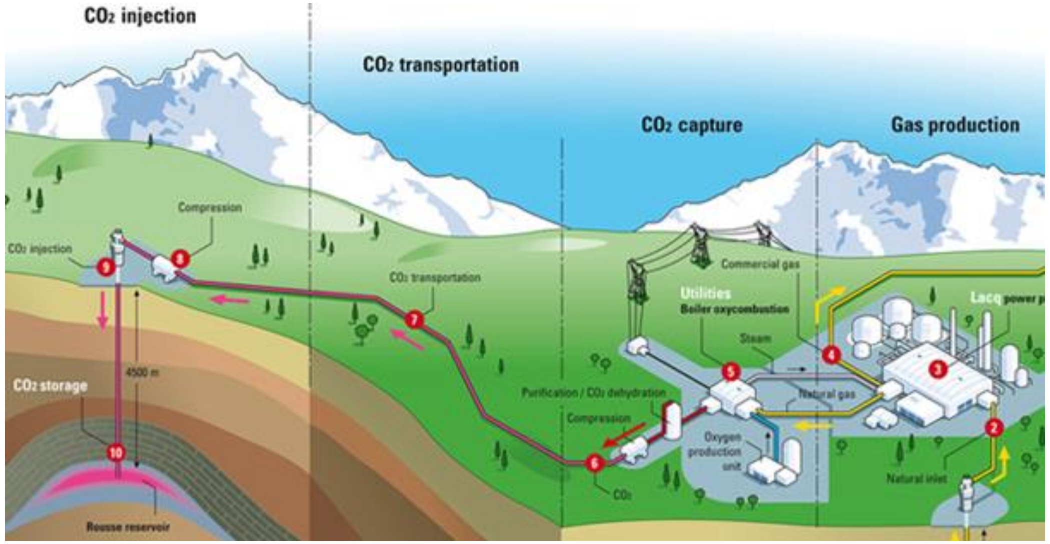
Detect leaks of 100 g/s within 100 d over an area of 100 km^2 for $100k/yr amortized over 10 years.

<u>Challenge 1</u>: Unsatisfactory Detectability

Current geophysical monitoring methods do not yield sufficient detectability to capture very small leakage (due to limitations in data coverage, low spatial resolution, acquisition noise and artifacts, etc.)

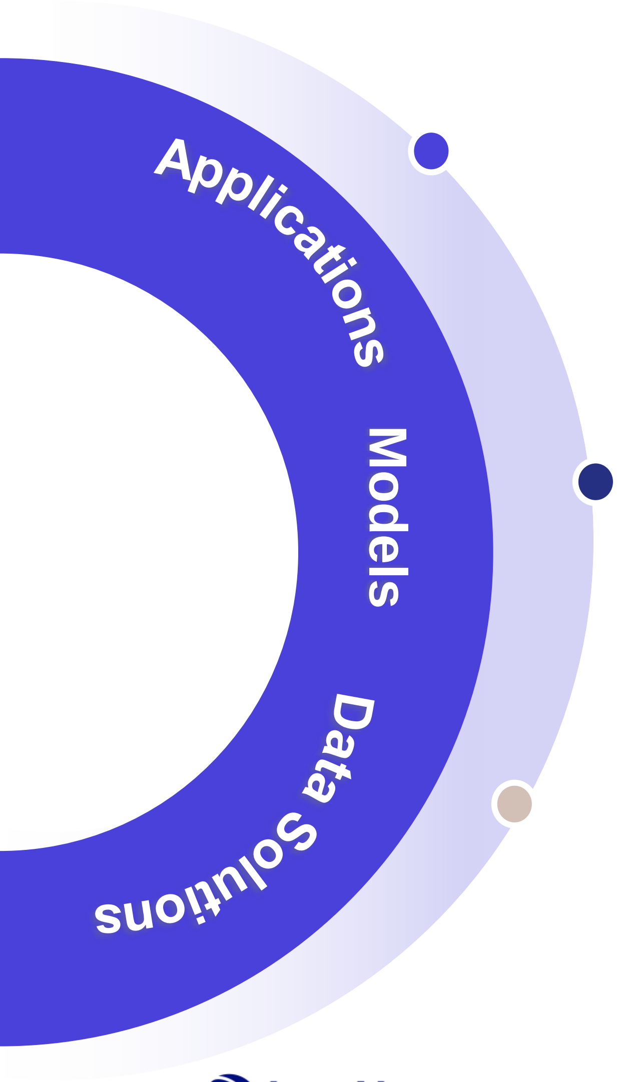
<u>Challenge 2:</u> Expensive Geophysical Monitoring

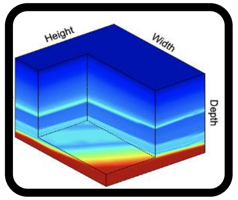
The high financial/computational cost and subjective human factors hinders the applicability of the existing monitoring methods.



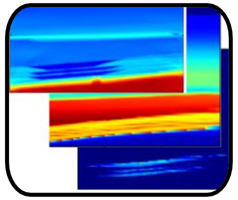
Small Leakage Detection

GeoVision: Seismic Imaging & Inversion Suite – an Overview

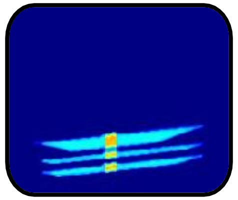




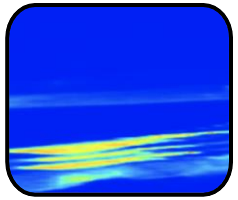
2D/3D Imaging
[Zeng et al. (2021)]



Multiphysics
[Feng et al. (2022)]

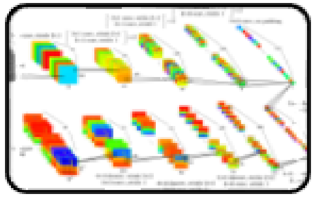


Time-Lapse
[Liu et al. (2021)]

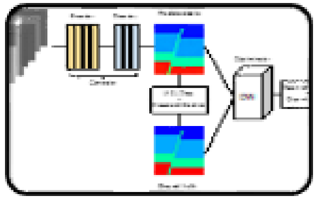


Uncertainty
[Liu et al. (2022)]

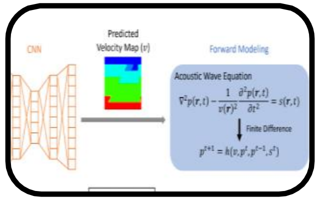
Induced Seismic
[Zhang et al. (2022)]



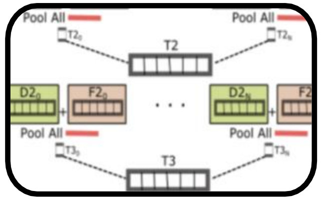
InversionNet
Wu & Lin (2019)]



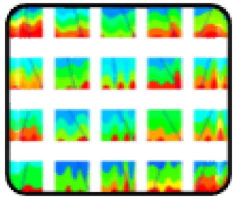
Physics Guided CNN
[Zhang & Lin (2020)]



Unsupervised CNN
[Jin et al. (2021)]

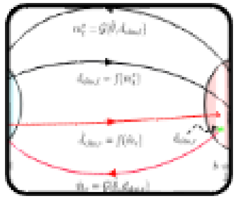


Graph Network
[Zhang et al. (2022)]

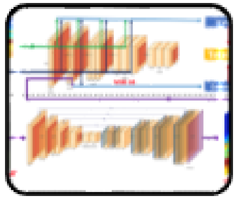


Physics Simulations
[Lin et al. (2017, 2018)]

Transfer Learning
[Zhang & Lin (2019)]



Physics + Data Augmentation
[Gomez et al. (2020)]



Style Learning
[Feng & Lin (2021)]

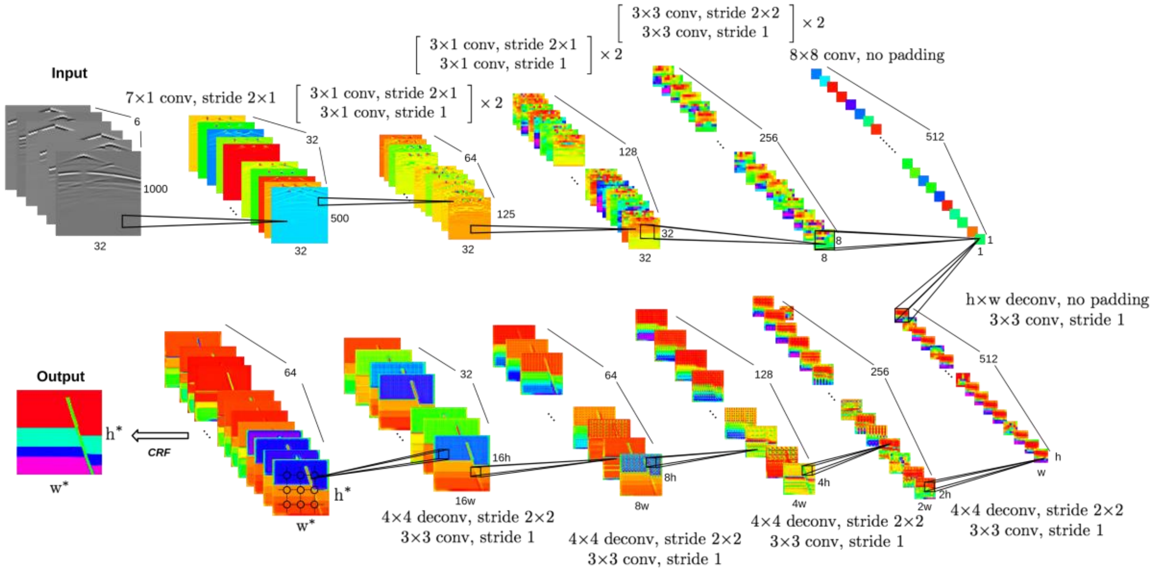
GeoVision Driven by Physics and Machine Learning

What is GeoVision?

- Collection of **site-agnostic** geophysical imaging techniques
- Based on **physics-guided machine learning**

What GeoVision Does?

- **Purely Data-driven Neural Networks** *[Wu and Lin, (2019)]*
  - *Real-time 2D/3D CO_2 Plume Imaging (Saturation)*
  - *Leakage Detection*
  - Uncertainty & Risk Estimate (Data and Model Error)



InversionNet

Yue Wu and Youzuo Lin, "InversionNet: An Efficient and Accurate Data-driven Full Waveform Inversion," IEEE Transactions on Computational Imaging, 6(1):419-433, 2019.

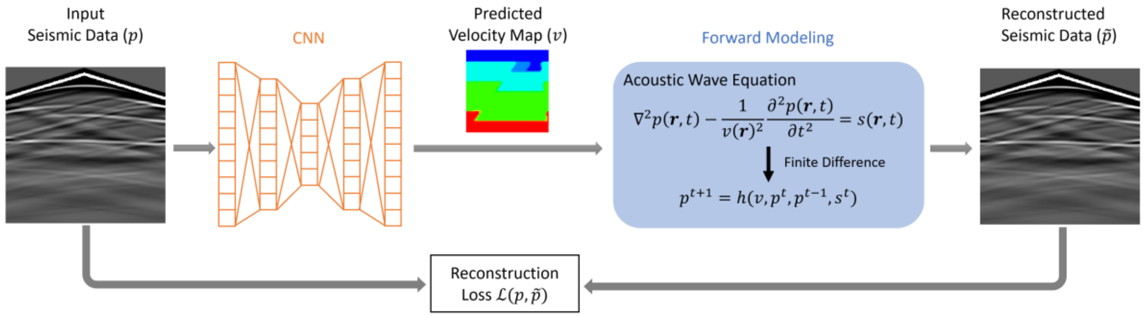
GeoVision Driven by Physics and Machine Learning

What is GeoVision?

- Collection of **site-agnostic** geophysical imaging techniques
- Based on **physics-guided machine learning**

What GeoVision Does?

- **Purely Data-driven Neural Networks** *[Wu and Lin, (2019)]*
  - *Real-time 2D/3D CO_2 Plume Imaging (Saturation)*
  - *Leakage Detection*
  - Uncertainty & Risk Estimate (Data and Model Error)
- **Physics-guided Unsupervised Networks** [Jin et al. (2021)]
  - Enable Imaging without any Label Information



Physics-guided Unsupervised Networks

Peng Jin, Xitong Zhang, Yinpeng Chen, Sharon Xiaolei Huang, Zicheng Liu, and Youzuo Lin, "Unsupervised Learning of Full-Waveform Inversion: Connecting CNN and Partial Differential Equation in a Loop", ICLR, 2022.

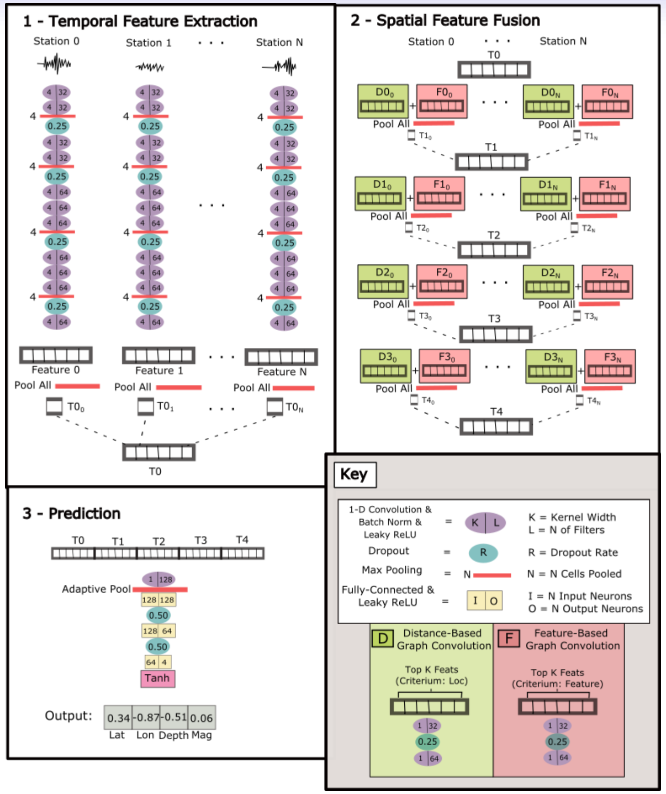
GeoVision Driven by Physics and Machine Learning

What is GeoVision?

- Collection of **site-agnostic** geophysical imaging techniques
- Based on **physics-guided machine learning**

What GeoVision Does?

- **Purely Data-driven Neural Networks** *[Wu and Lin, (2019)]*
  - *Real-time 2D/3D CO_2 Plume Imaging (Saturation)*
  - *Leakage Detection*
  - Uncertainty & Risk Estimate (Data and Model Error)
- **Physics-guided Unsupervised Networks** [Jin et al. (2021)]
  - Enable Imaging without any Label Information
- **Graph Convolution Neural Networks** *[Zhang et al. (2022)]*
  - Induced seismic monitoring and characterization



Graph Neural Network [Zhang et al., 2022]

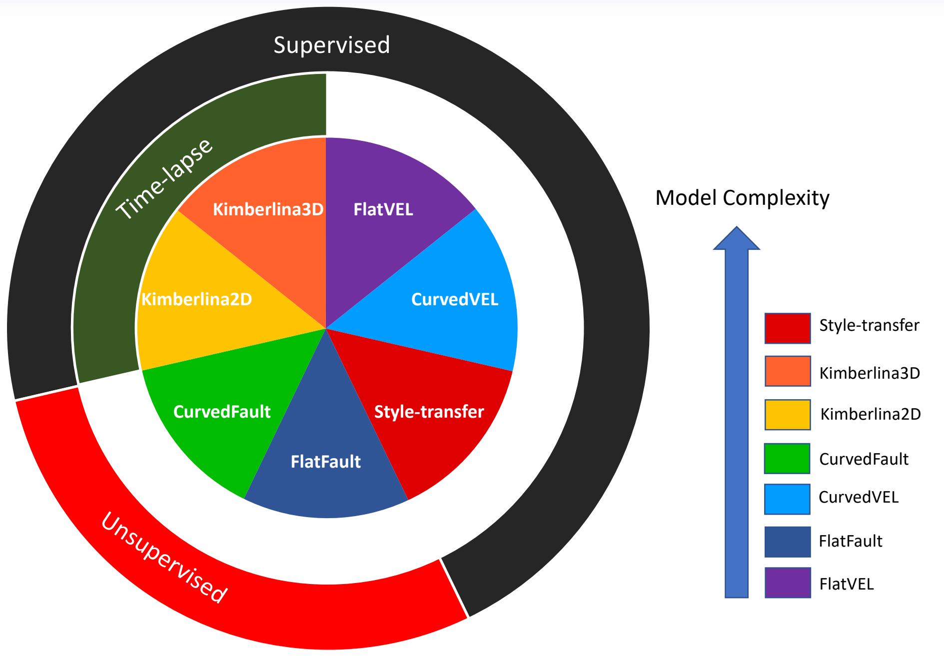
Xitong Zhang, Will Reichard-Flynn, Miao Zhang, Matthew Hirn, and Youzuo Lin, "Spatio-Temporal Graph Convolutional Networks for Earthquake Source Characterization", JGR-Solid Earth, 2022 (Under Review).

GeoVision Enhanced by Large-Scale High Quality Training Data

OpenFWI:
A large-scale open dataset for subsurface geophysics

- **Multi-scale and multi-dimension**
  - Over 500K of samples (labeled and unlabeled)
  - 2D (11 datasets) and 3D (1 dataset)
- **Multi-purpose Applications**
  - Carbon sequestration, fossil fuel energy, and general purposes
- **Multi-complexity Data**
  - Simple layered structures, hypothetic synthetic structures, and physically realistic structures

Data Overview



OpenFWI (https://openfwi-lanl.github.io/)

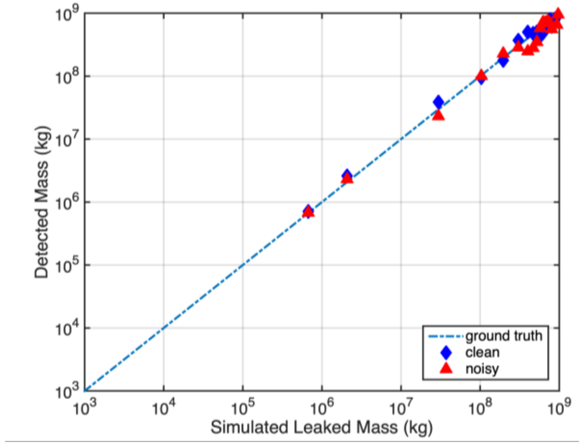
Chengyuan Deng, Shihang Feng, Hanchen Wang, Xitong Zhang, Peng Jin, Yinan Feng, Qili Zeng, Yingpeng Chen, and Youzuo Lin, "OpenFWI: Large-scale Multi-structural Benchmark Datasets for Full Waveform Inversion", arXiv, 2022 (under review NeurIPS).

Project Scope: Task & Milestone

Major Milestone

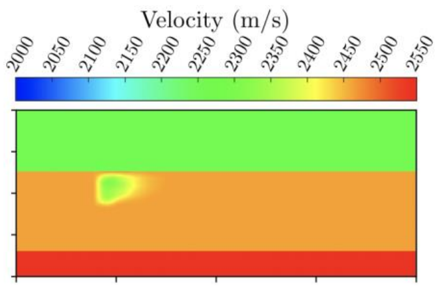
- Project Kick-off
  - Preliminary Study & Data Preparation
- R&D Tasks
  - Task 1 – Synthetic Data Test (Kimberlina 1.2)
  - Task 2 – Controlled Experiment Test 1 (Cranfield)
  - Task 3 – Controlled Experiment Test 2 (Sleipner)
  - Task 4 – Controlled Experiment Test 2 (San Juan)

Previous Task 1: Leakage Detection using Kimberlina 1.2 Data



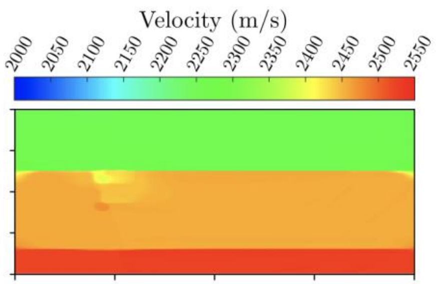
Leakage Mass Detection
[Zheng et al., 2019]

a



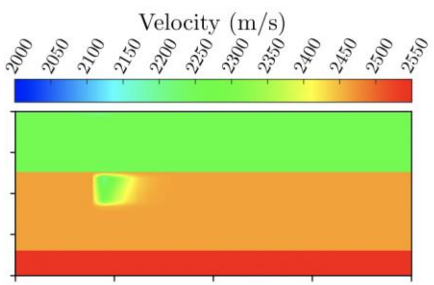
Ground Truth
[Gomez et al., 2020]

b



Traditional
[Gomez et al., 2020]

c



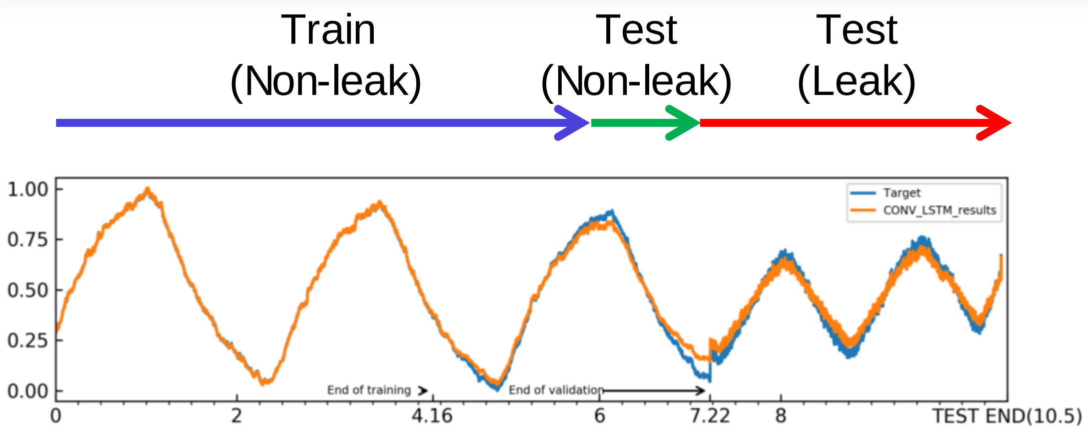
GeoVision
[Gomez et al., 2020]

d

- GeoVision learns critical information from ***massive amount of data*** to predict leakage mass and plume.
- Collaboration with Zan Wang and Bob Dilmore via NRAP.

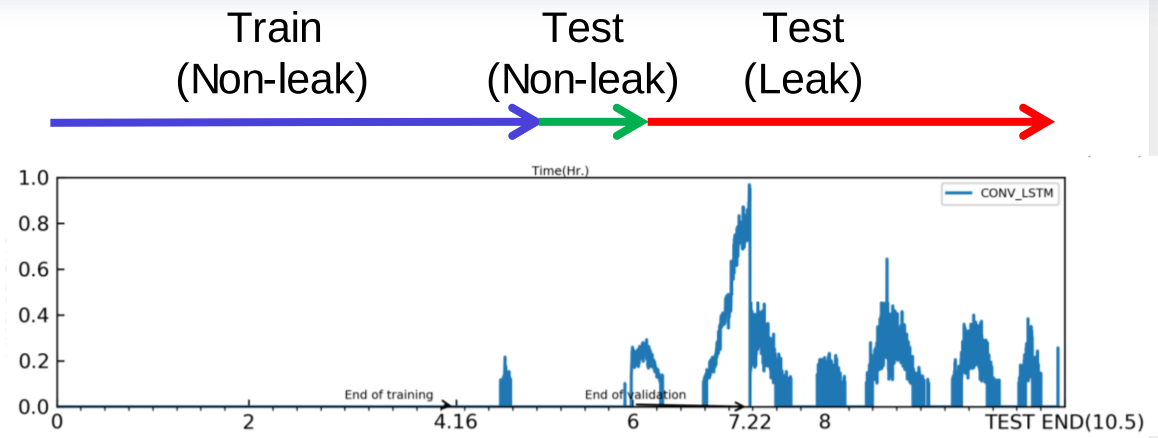
Zheng Zhou, Youzuo Lin, Zhongping Zhang, Yue Wu, Zan Wang, Robert Dilmore, and George Guthrie, "A Data-Driven CO2 Leakage Detection Using Seismic Data and Spatial-Temporal Densely Connected Convolutional Neural Networks," International Journal of Greenhouse Gas Control, Vol 90, 2019.

Previous Task 2: Leakage Detection using Cranfield Data



Signature Prediction
[Sinha et al., 2020]

a



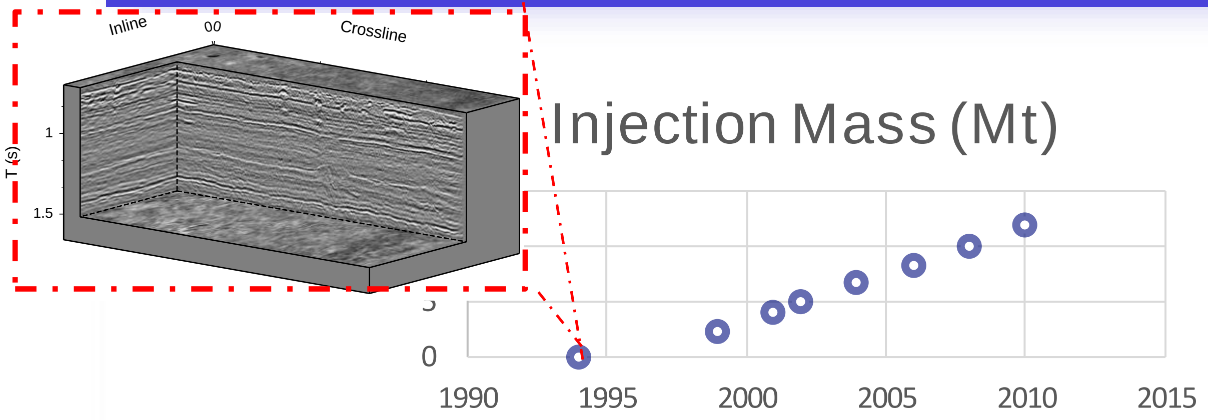
Leakage Detection
[Sinha et al., 2020]

b

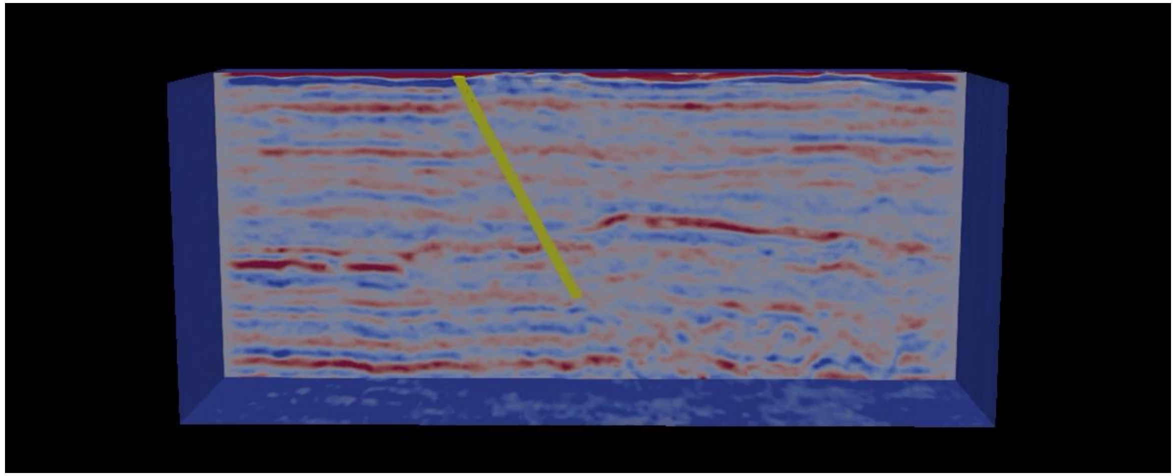
- ML model, trained on non-leak temporal pressure data, can predict leakage.
- Through the collaboration with Alex Sun and BEG.

Saurabh Sinha, Rafael Pires de Lima, Youzuo Lin, Alexander Y. Sun, Neill Symons, Rajesh Pawar, and George Guthrie, “Normal or Abnormal? Machine Learning for the Leakage Detection in Carbon Sequestration Projects Using Pressure Field Data,” International Journal of Greenhouse Gas Control, Vol. 103, 2020.

Previous Task 3: In-Situ Monitoring using Sleipner Data



Repeated Seismic Acquisition (10 years)



In-situ Monitoring
[Feng et al., 2021]

a

b

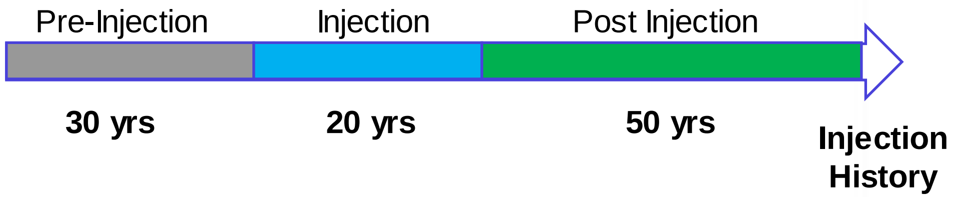
When sufficiently trained, ML can ***fill in the gap*** of static data to inform the dynamics of the plume.

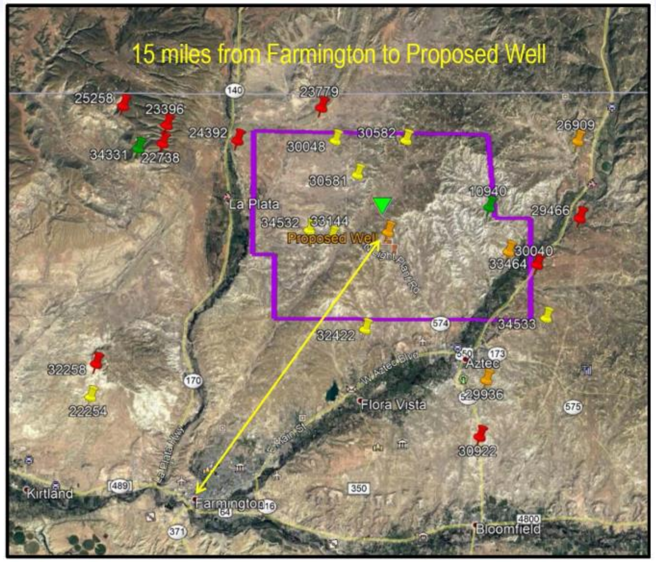
Shihang Feng, Xitong Zhang, Brendt Wohlberg, Neill Symons, and Youzuo Lin "Connect the Dots: In Situ 4D Seismic Monitoring of CO2 Storage with Spatio-temporal CNNs," IEEE Transactions on Geoscience and Remote Sensing, vol 60, 1-- 16 2021.

Task 4: Time-lapsed Imaging using San Juan Basin Data

San Juan Basin Dataset (CarbonSAFE)

- **Data Availability**
  - Baseline velocity model built from well logs
  - Time-lapse velocity models built from reservoir simulation (100 yrs)
  - Seismic Data Simulation (5 sources and 70 receivers)

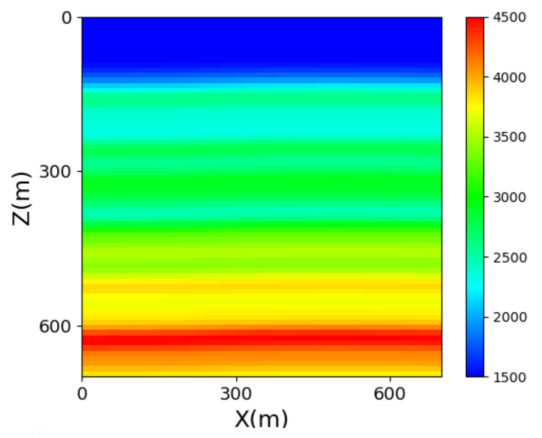




San Juan Basin

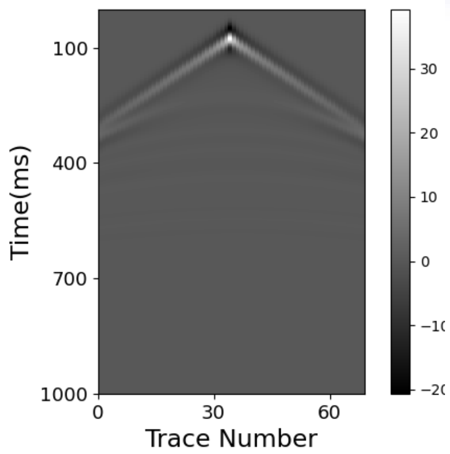
- Collaboration with William Ampomah (NMT) via CarbonSAFE.

Task 4: Time-lapsed Imaging using San Juan Basin Data



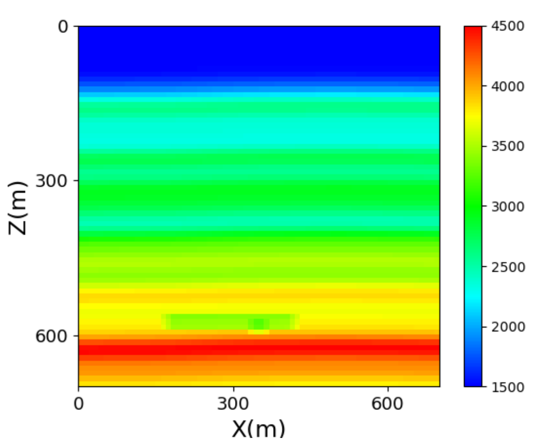
Baseline Velocity

a



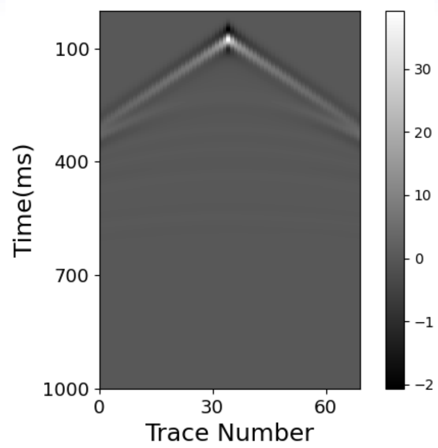
Baseline Seismic

b



Time-lapse Velocity

c



Time-lapse Seismic

d

Test ML's performance with ***very few data***

- ***Sub-Task 1: Seismic Data & Well Logs (Given b & d & well logs, infer a & c)***
- ***Sub-Task 2: Seismic Data (Given b & d, infer a & c)***

Motivation

"*What I **cannot create,***
I do not understand"

Richard P. Feynman

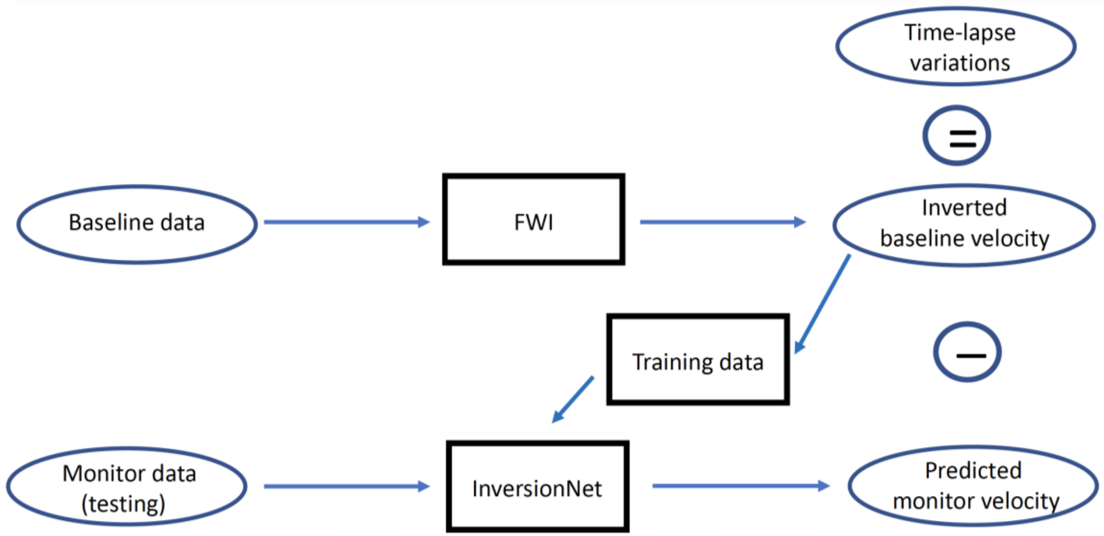
Task 4.1: Create Velocity from Reservoir Simulation & Well Logs

What *Physics* do we know

- Wave equation
- Well logs values
  - Location of the reservoir
  - Range of velocity perturbation

A hybrid imaging model

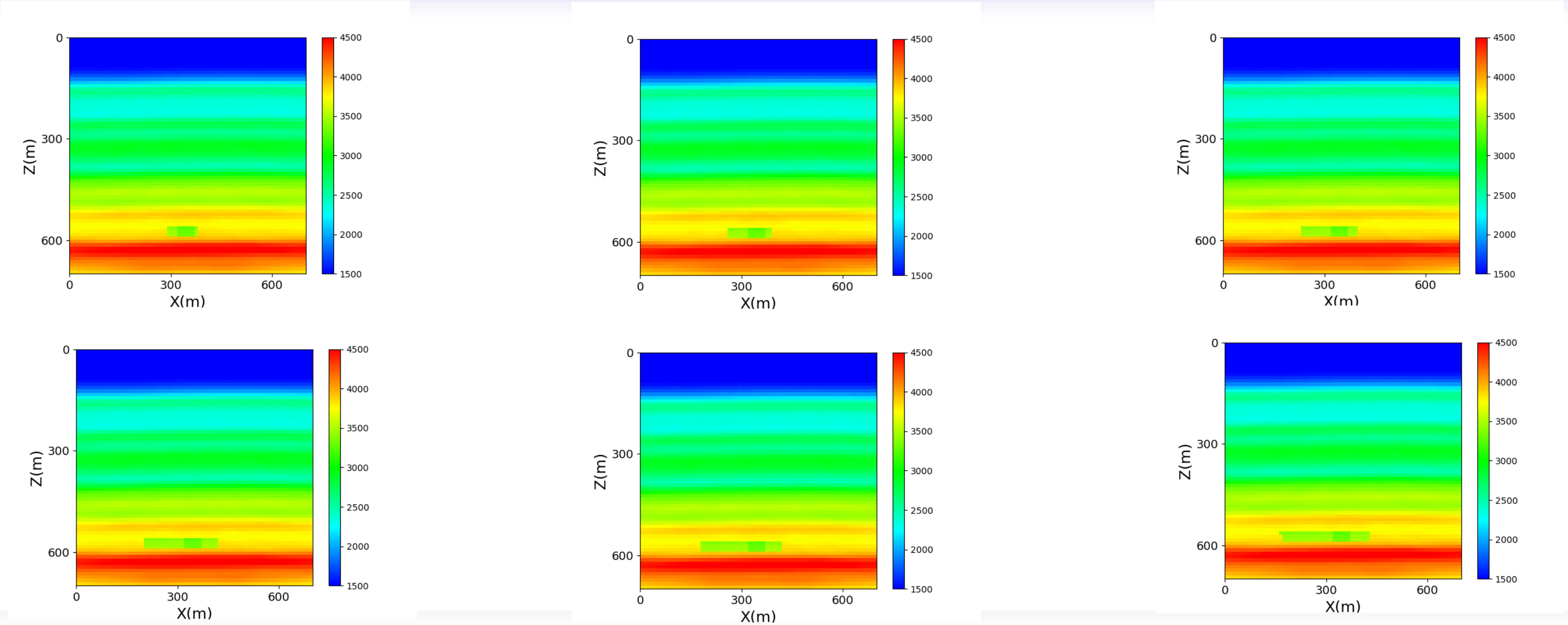
- Baseline Inversion: traditional FWI
- Time-lapse Inversion: Machine Learning



Proposed Idea [Liu et al., 2022]

Yanhua Liu, Shihang Feng, Ilya Tsvankin, David Alumbaugh, and Youzuo Lin, "Joint Physics-Based and Data-Driven Time-Lapse Seismic Inversion: Mitigating Data Scarcity," under review in Geophysics, 2022.

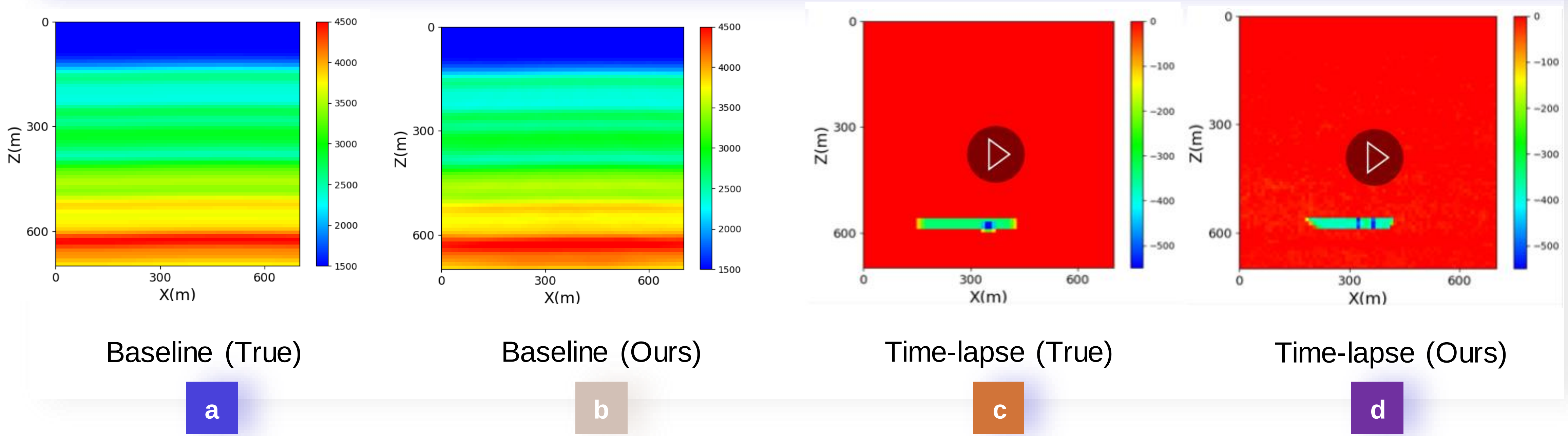
Task 4.1: Create Velocity from Reservoir Simulation & Well Logs



- High quality synthetic velocity models are generated.

Yanhua Liu, Shihang Feng, Ilya Tsvankin, David Alumbaugh, and Youzuo Lin, “Joint Physics-Based and Data-Driven Time-Lapse Seismic Inversion: Mitigating Data Scarcity,” under review in Geophysics, 2022.

Task 4.1: Create Velocity from Reservoir Simulation & Well Logs

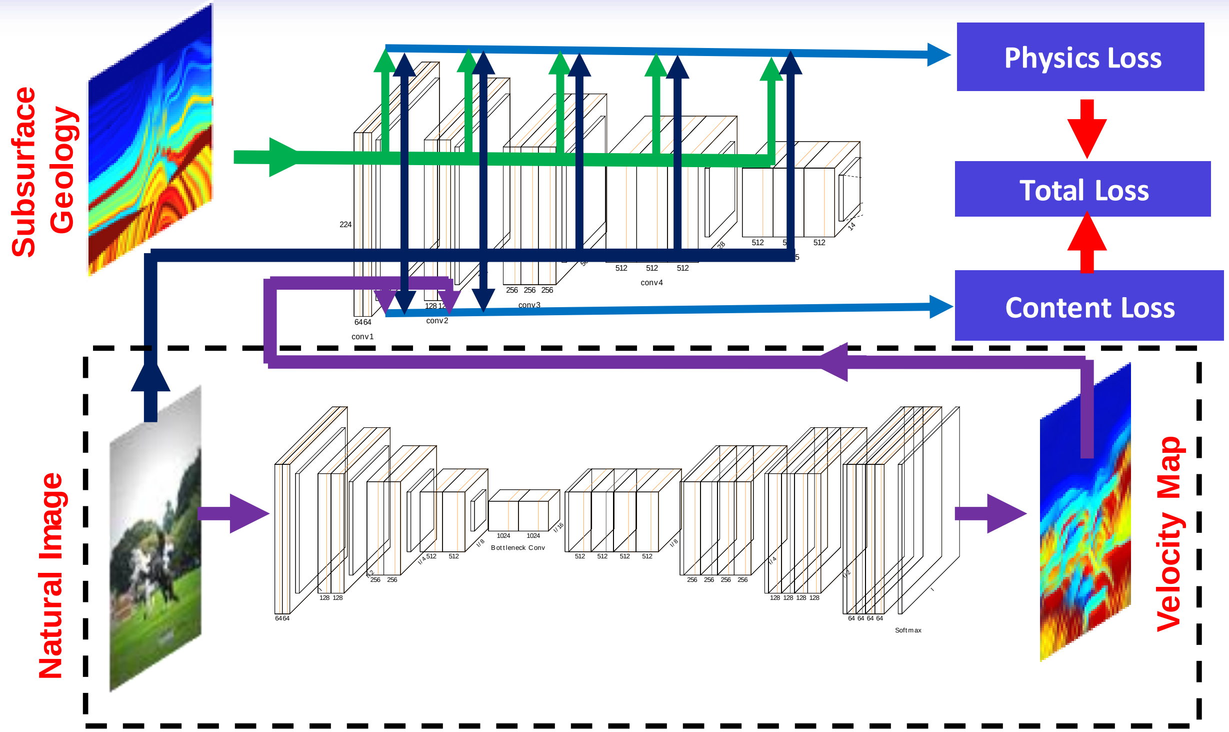


- High quality synthetic velocity models are generated.
- CO_2 plume dynamics are very well captured.

Task4.2: Create Velocity from Style-Transform Learning

VelocityGen: Transform from Natural Images into Velocity Models

Synthesize large volume of velocity maps from natural images that would yield realistic representations of subsurface kinematics and geology.



VelocityGen [Feng et al., 2021]

Shihang Feng, Youzuo Lin, and Brendt Wohlberg, "Multiscale Data-driven Seismic Full-waveform Inversion with Field Data Study", IEEE Transactions on Geoscience and Remote Sensing ,vol 60, p1 – 14, 2021.

Task4.2: Create Velocity from Style-Transform Learning

Input Dataset

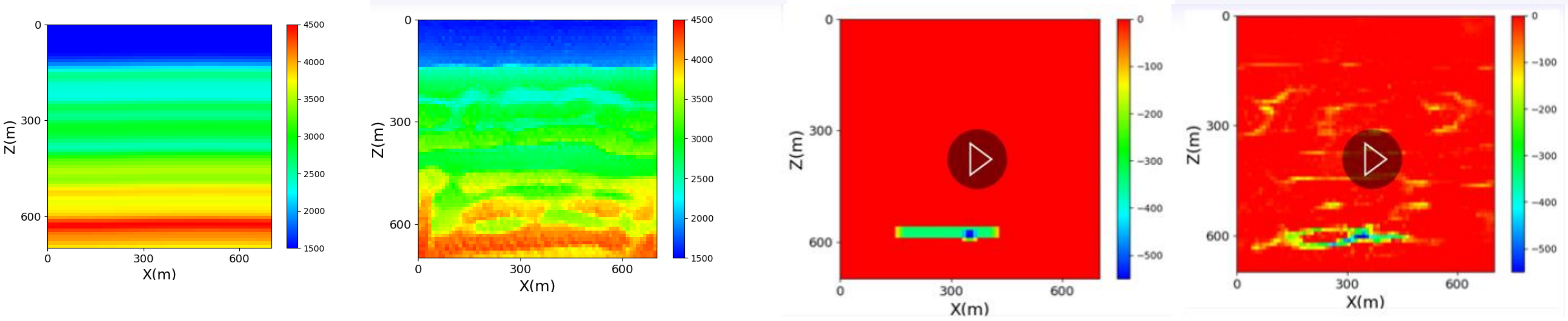
- **Content Images**
  - **COCO dataset** contains 67,000 natural images
- **Geologic Style**
  - **Marmousi velocity model** is used as geologic style image

Our method yields physically realistic velocity models

Velocity Models Generated from Natural Images

Shihang Feng, Youzuo Lin, and Brendt Wohlberg, "Multiscale Data-driven Seismic Full-waveform Inversion with Field Data Study", IEEE Transactions on Geoscience and Remote Sensing ,vol 60, p1 – 14, 2021.

Task 4.2: Create Velocity from Style-Transform Learning



Baseline (True) **a**

Baseline (Ours) **b**

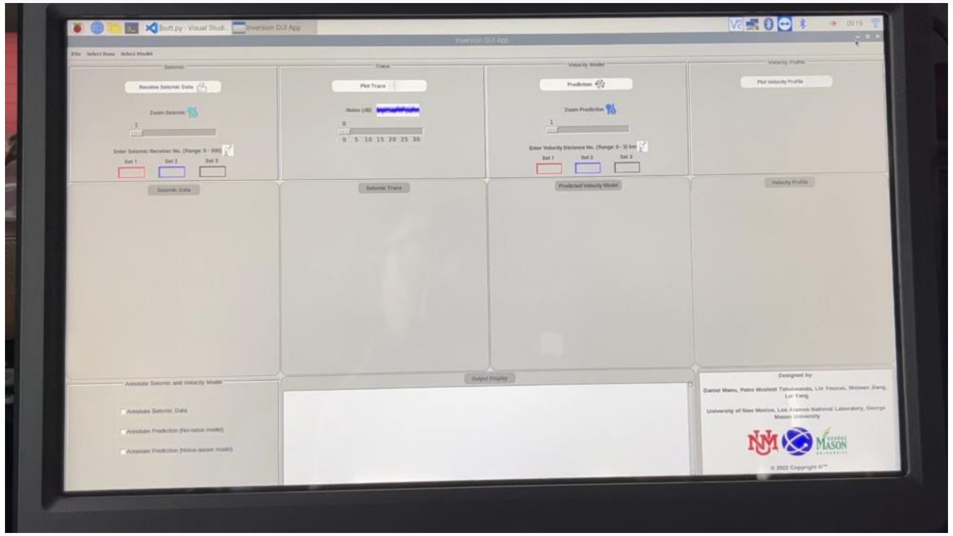
Time-lapse (True) **c**

Time-lapse (Ours) **d**

- Natural images are good sources of images for subsurface structures!
- Additional R&D effort would be needed to further improve the imaging quality.

Imaging on the *Edge*: a Step Toward Real-Time Decision Making

a



b

- Run GeoVision on an edge device (Raspberry Pi) to provide real-time imaging.
- Won 2nd Place in SIGDA University Demonstration at the 2022 Design Automation Conference.
- Collaboration with U New Mexico and George Mason.

Summary

01 Machine Learning

Machine learning, particularly, deep neural network models provide great potential in improving seismic imaging performance (i.e., enhancing the spatial resolution and accelerating imaging speed).

02 Physics Knowledge

The incorporation of physics knowledge will significantly improve model robustness and generalization.

03 Data Scarcity

One of the major challenges employing deep learning-based imaging models for CO_2 monitoring is the lack of data.

04 GeoVision

- We demonstrate the great potential of GeoVision using four case studies (Kimberlina, Cranfield, Sleipner, and SJB).
- **Look for more applications!**

Acknowledgement

- The computation was performed using both Darwin cluster and HPC facilities of LANL's Institutional Computing Program.

ylin@lanl.gov

YouzuoLin1

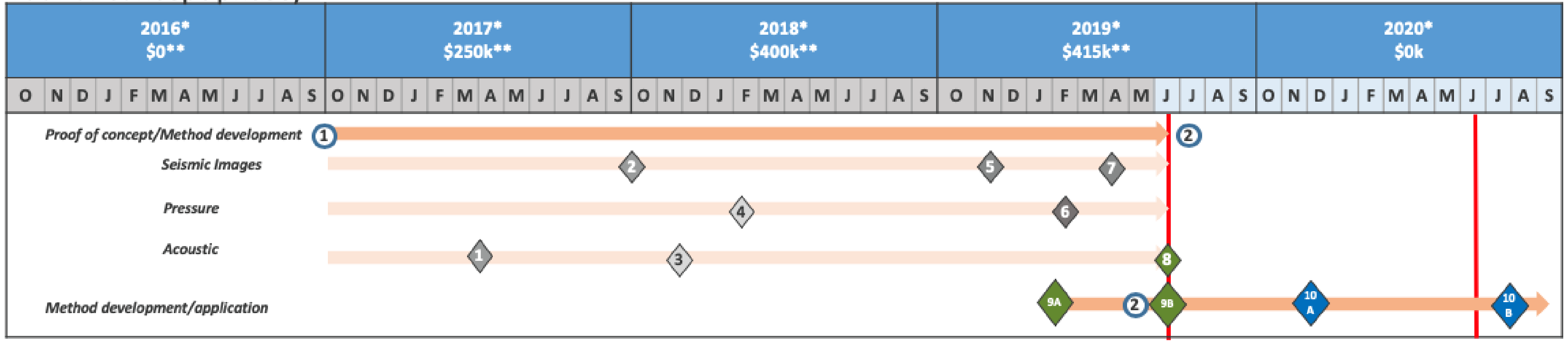
Gantt Chart

Task 2: Project Timeline Overview

NATIONAL ENERGY TECHNOLOGY LABORATORY

Monitoring for small leaks over large areas post-injection using conventional datasets

(Proof of concept phase)



Milestones

1. Collected field data set for testing extraction of small acoustics signals associated with fluid movement (leakage)
2. Seismic Images: Development/testing of machine-learning (ML) method for extracting large-leak signal from synthetic data
3. Acoustic: Development/testing of ML method for extracting large-leak signal from acoustic
4. Pressure: Development/testing of ML method for extracting large-leak signal from pressure data
5. Seismic Images: Testing of ML potential to extract large-leak signal from noisy synthetic seismic images
6. Pressure: Testing of ML method for extracting small-leak signal from noisy pressure data
7. Seismic Images: Testing of ML potential to extract small-leak signal from synthetic seismic images
8. Acoustic: Testing of ML method for extracting small-leak signal from noisy acoustic data

9a. Develop requirements for Method Development/Application
9b. Develop multi-data ML integration method
10. Initial testing on real data sites (a, b, c…)

Chart Key

TRL Score; Go / No-Go Timeframe; Project Completion; Milestone; Complete; FY19; FY20

Go / No-Go

1. Determine which method(s) should be used to detect a leak of 100 g/s over an area of 100 km^2 for 10-year amortized cost of \$100k/year
2. *Initiate development of multi-data ML integration platform and test/demonstrate on field data?* Decision based on proof-of-concept that analysis of conventional data (seismic ± pressure ± acoustic) could meet performance/cost goals