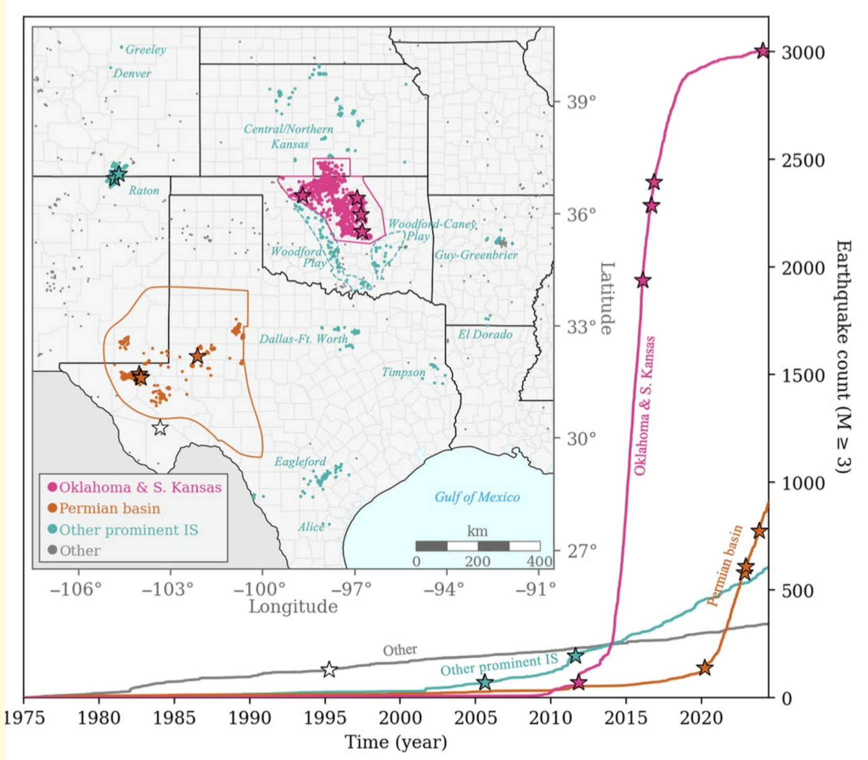
Managing Induced Earthquake Potential with Deep Learning

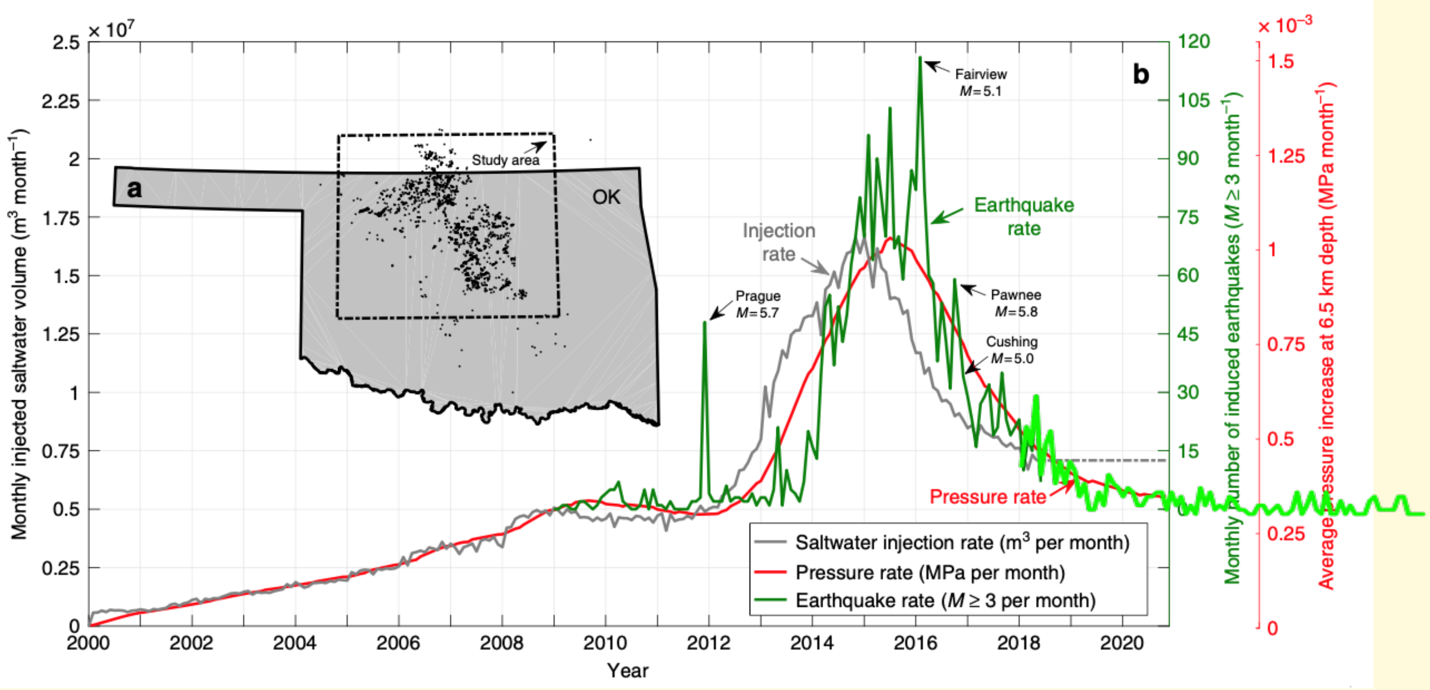
William L. Ellsworth[1], Gregory Howe[2], Brandon Liu[2], Bridge Elmon[2], Matthew Gebara[2], Owen Murphy[2], Gregory C. Beroza[1]

1 Stanford University, USA; **2** Stottler Henke Associates, USA



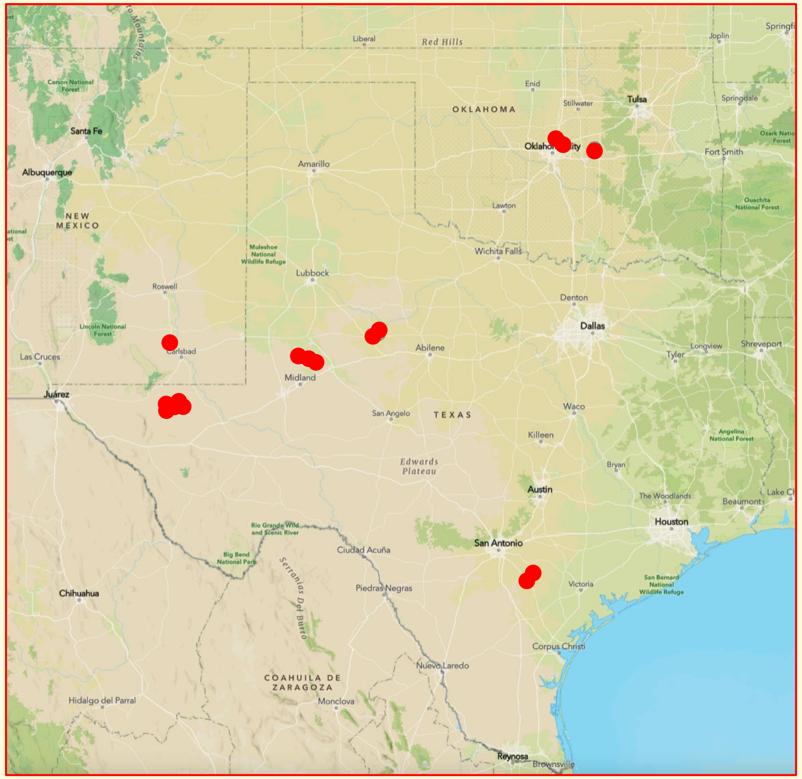
Skoumal, Barbour, Rubinstein and Glasgow (TSR, 2024)

Oklahoma Seismicity Forecast

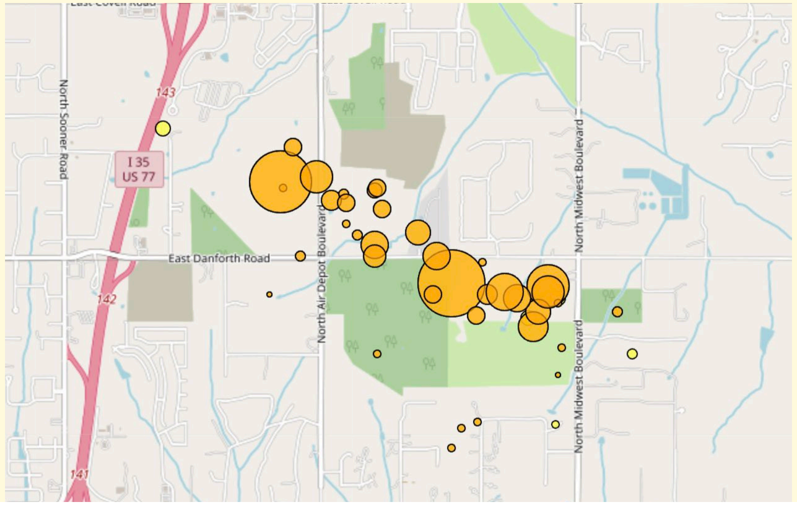


Langenbruch, Weingarten and Zoback (2018)

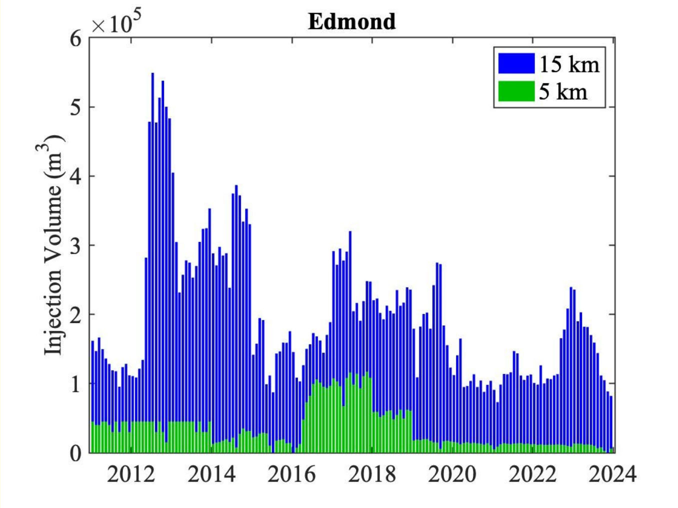
In the last 14 months there have been 23 earthquakes M 4.0 – 5.1 in Oklahoma, Texas and New Mexico



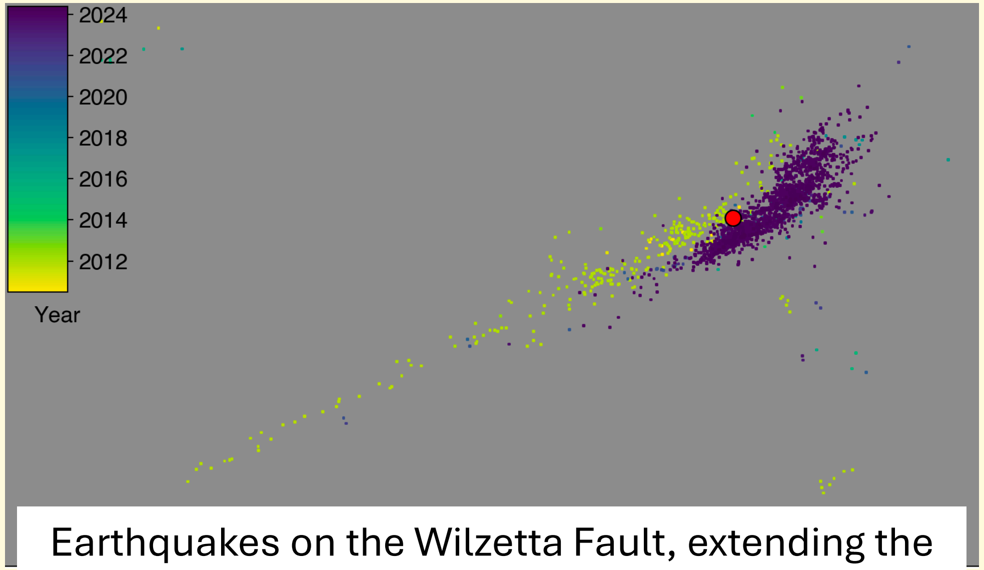
Edmond, Oklahoma January 13, 2024
M_W 4.1, 4.3



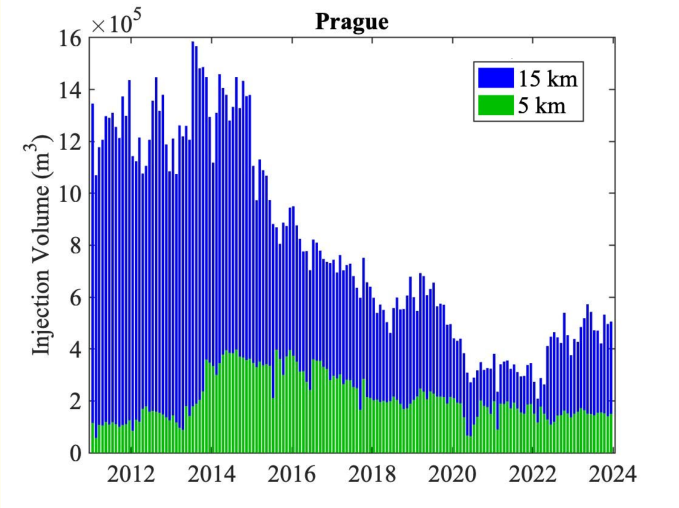
Earthquakes on the same faults that ruptured in 2015 (M 4.3), 2016 (M 4.2) and 2017 (M 4.2)



Prague, Oklahoma February 2, 2024
M_W 5.1, 4.3

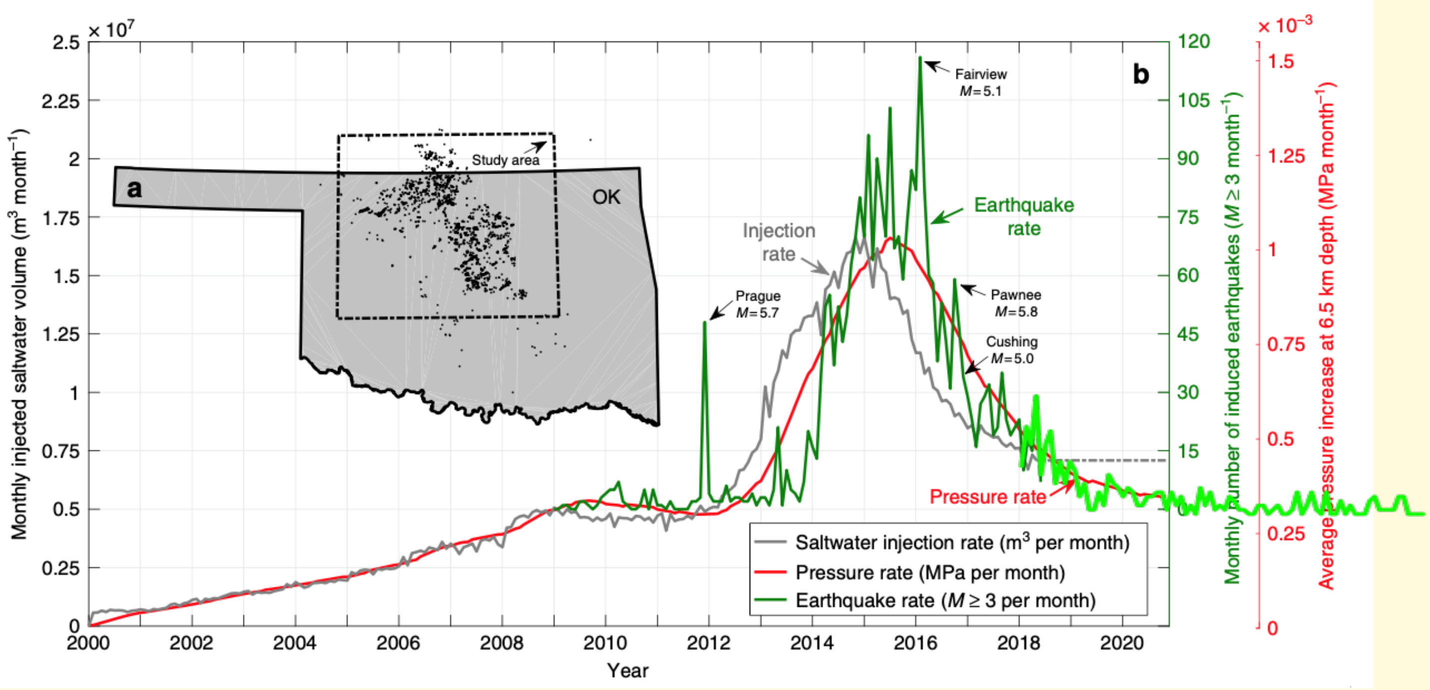


Earthquakes on the Wilzetta Fault, extending the 2011 M 5.7 sequence to the northeast

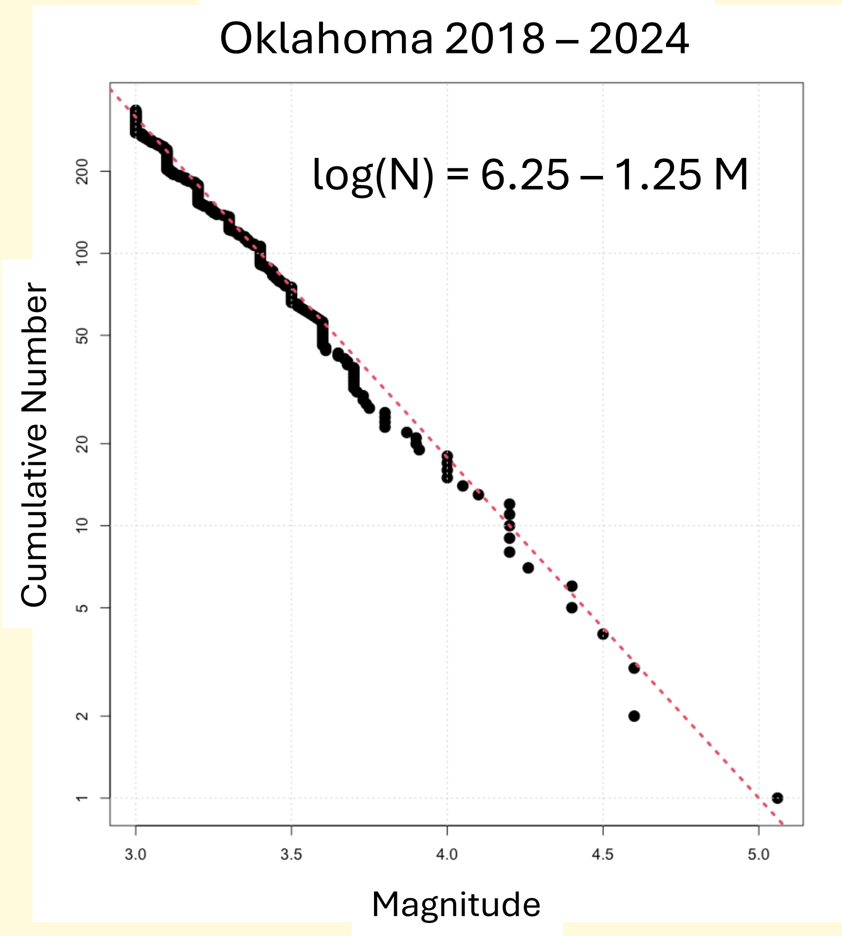


Work done with:
Jake Walter, Paul Oguari, Ben Allen (OGS)
Jeong-Ung Woo (LANL)
Margaret Glasgow, Rob Skoumal (USGS)

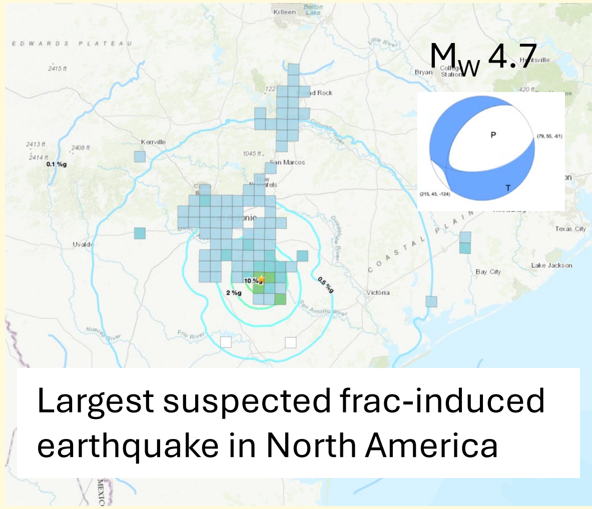
Oklahoma Seismicity Forecast



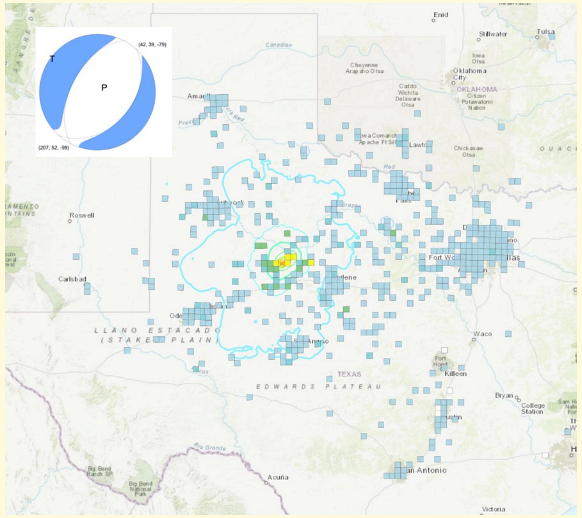
Langenbruch, Weingarten and Zoback (2018)



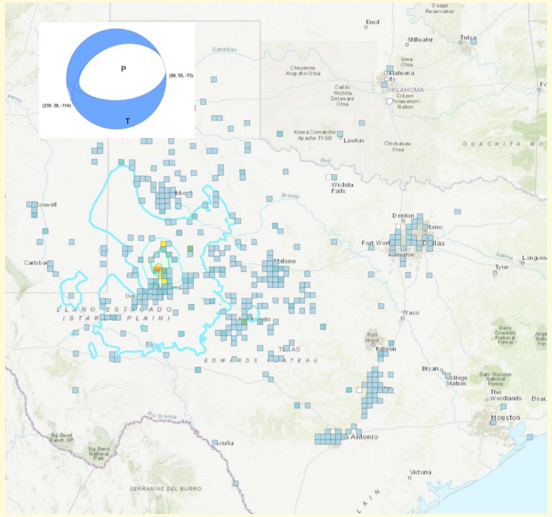
Falls City, Texas
M_W 4.3, 4.7 February 17, 2024
M_W 4.6 January 30, 2025



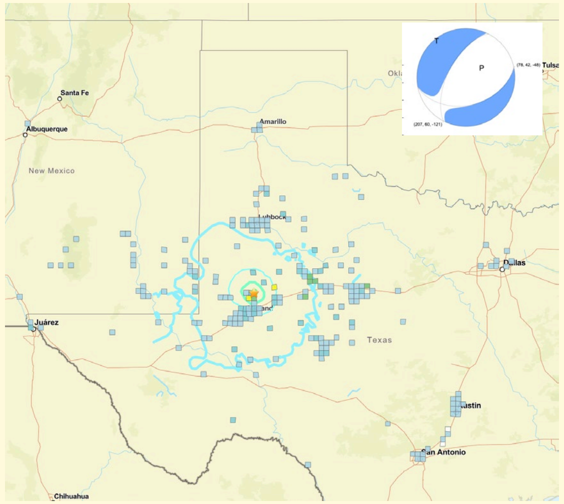
Hermleigh, Texas
July 23 - 27, 2024
M_W 4.6, 4.8, 4.1

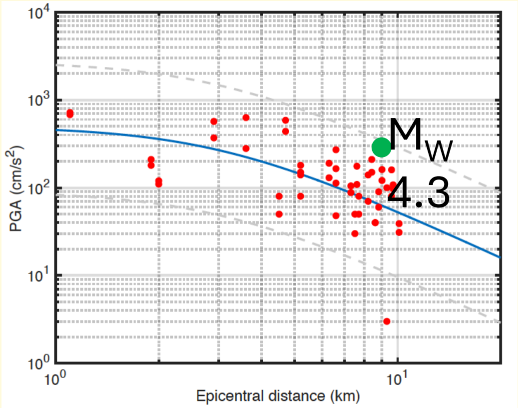


North Tarzan, Texas
September 17, 2024
M_W 4.8

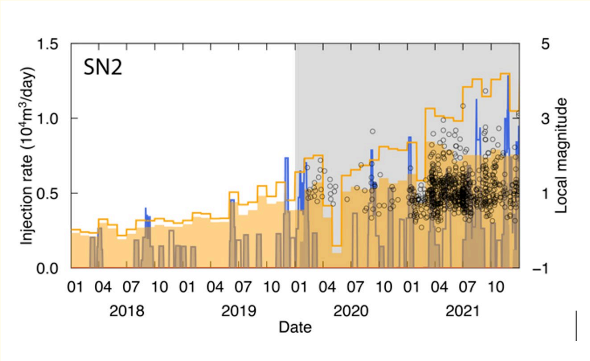


Ackerly, Texas
February 28, 2025
M_W 4.7

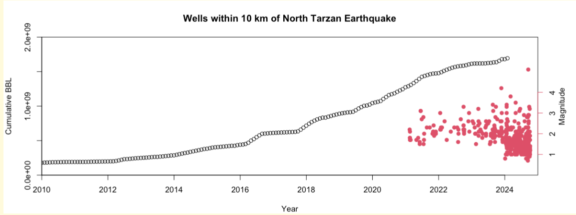




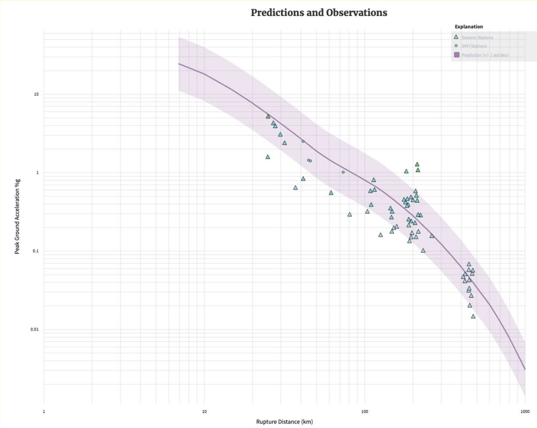
Peak Ground Acceleration M_W 4 to 4.5
Huang et al. (2017)



Woo and Ellsworth (2023)



Cumulative Wastewater Disposal and Seismicity



Predicted and Observed Peak Ground Acceleration

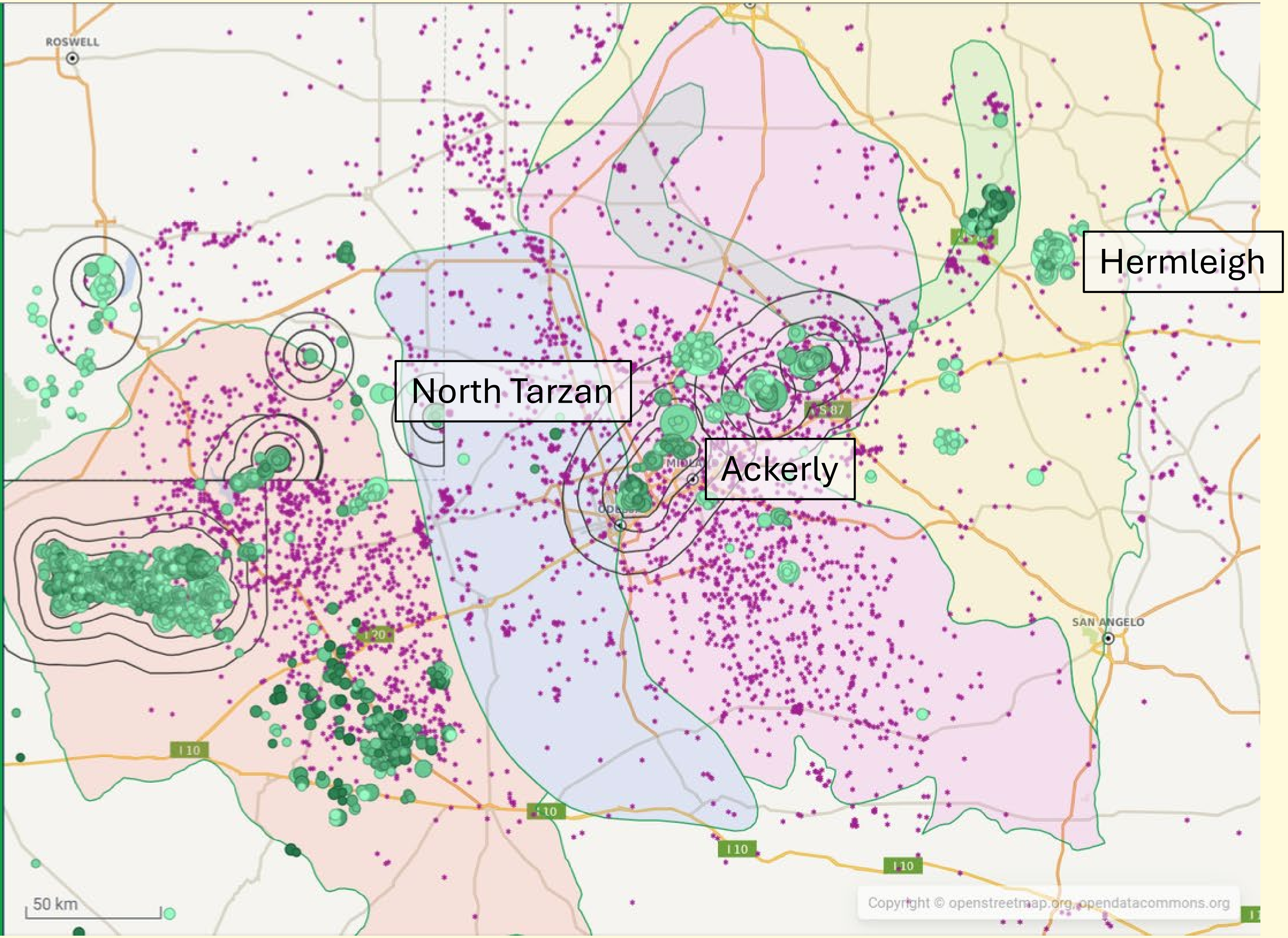
Permian Basin

Delaware, Midland, CBP

Seismicity >M2.5 2018-2024

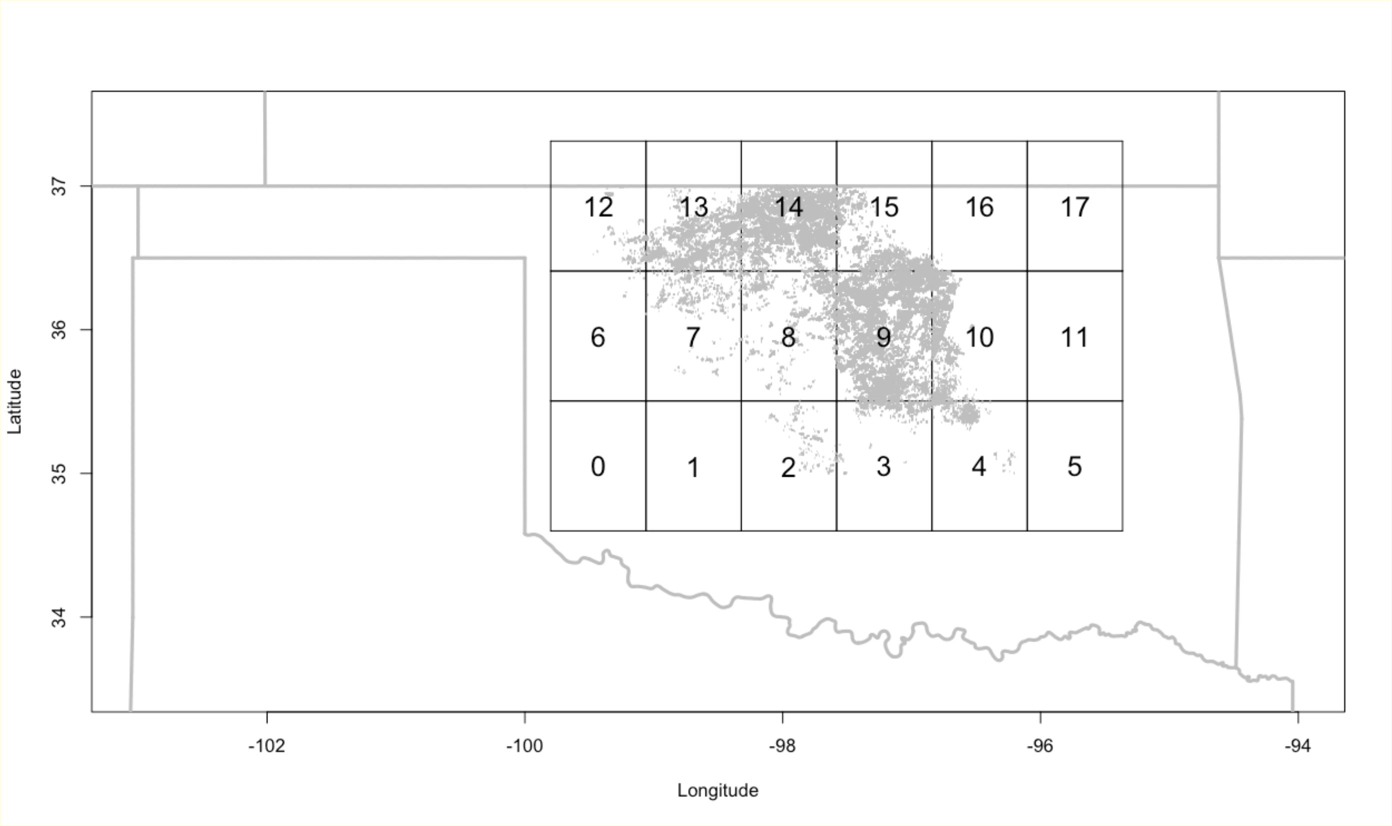
Seismic Response Areas (SRAs)

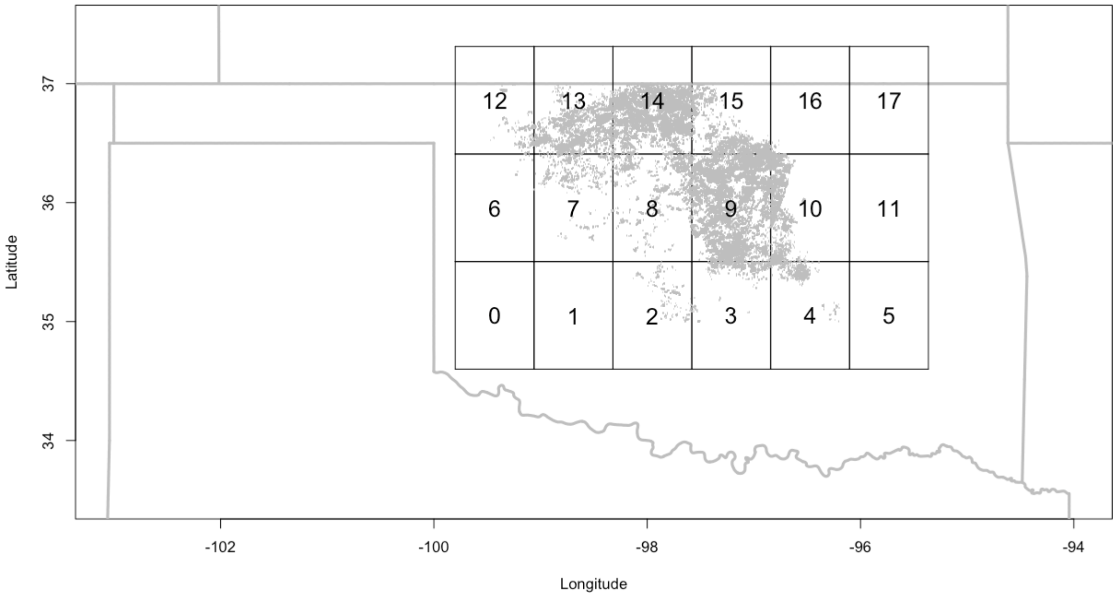
Saltwater Disposal (Deep & Shallow)

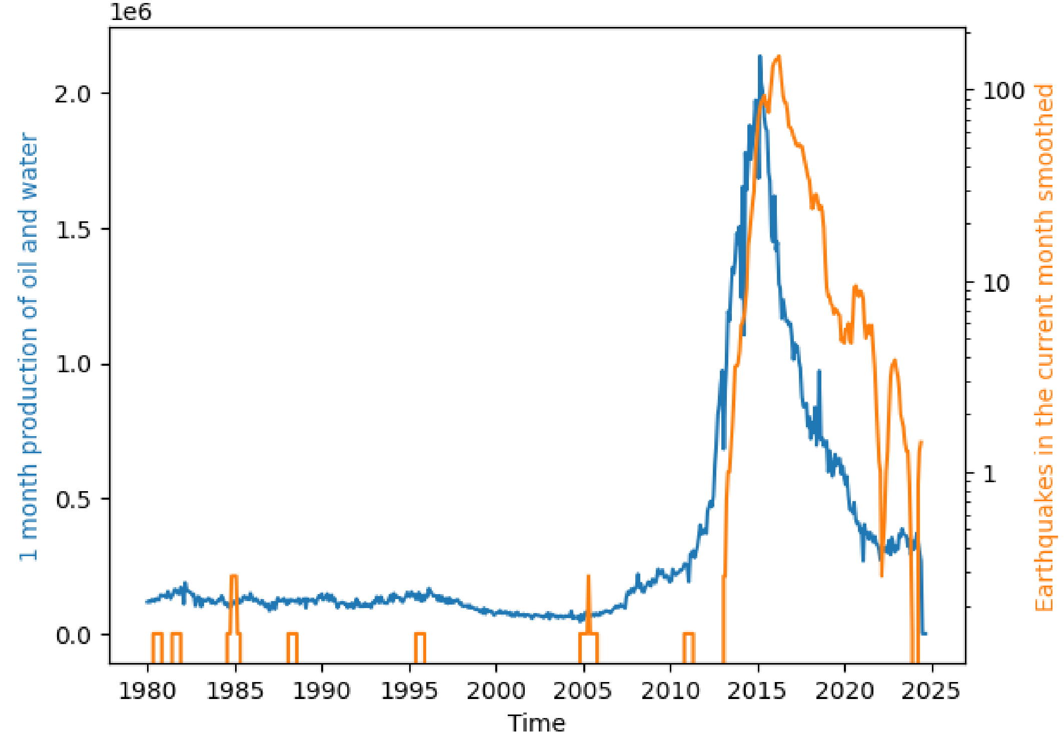


A Deep Learning Model for Induced Seismicity Forecasting

- Develop a *time-series forecasting model* that incorporates injection and production information
- Demonstrate feasibility by outperforming baseline statistic models (such as rolling means)
- Keep it simple!





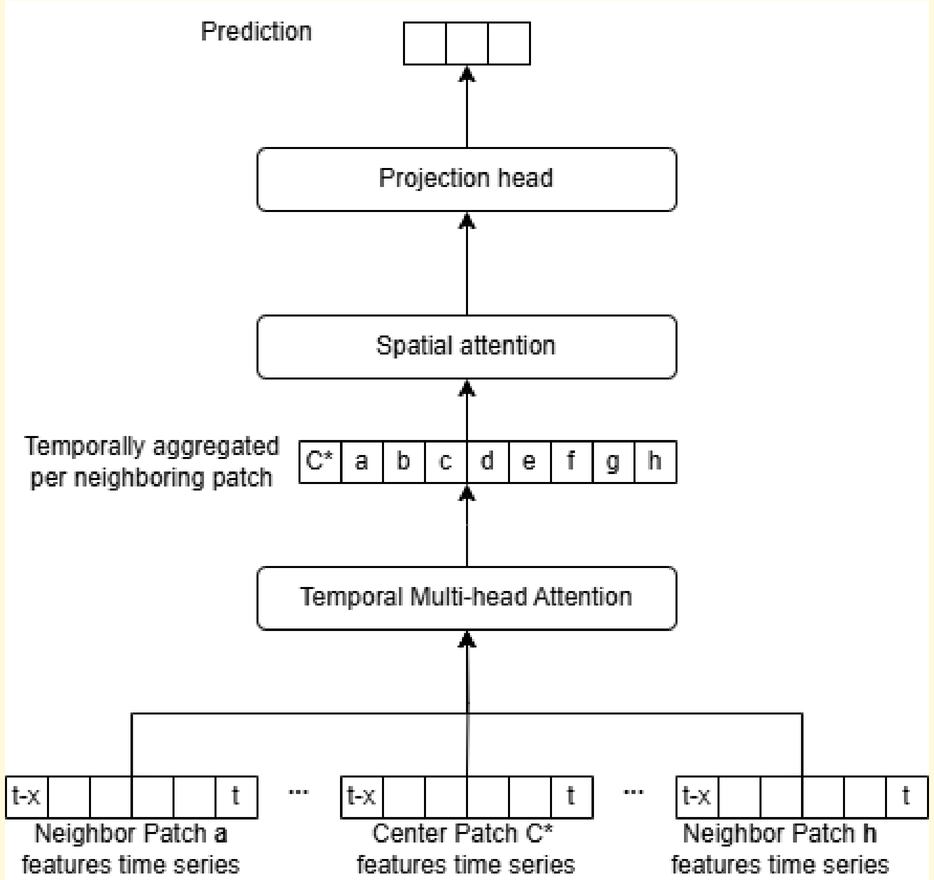


Problem Formulation / Data

- Divide region into subgrids
- Extract temporal time-series **features** from monthly operational data taken from ([data source])
  - Production of oil/water, deep disposal, shallow diosposal, fracking
- Extract **labels** from earthquake catalogs (TexNet in Texas, and OGS in Oklahoma)
  - Monthly earthquake count
  - Monthly maximum magnitude

Model Architecture

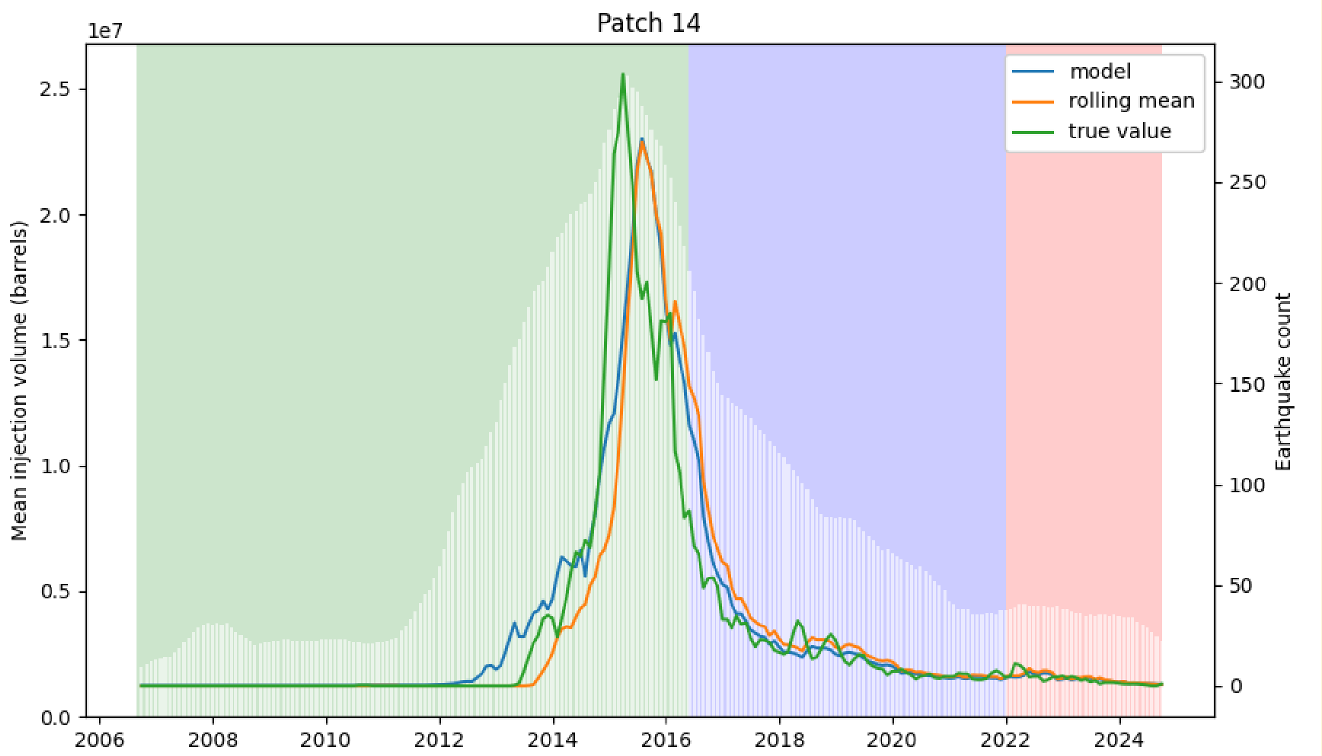
- We use a graph neural network architecture that uses attention to aggregate information
  - Temporal attention attends to the temporal context of each patch
  - Spatial attention aggregates temporal context of neighboring patches
- *Crucially*, use the baseline model as a feature
  - Learning against the residual error

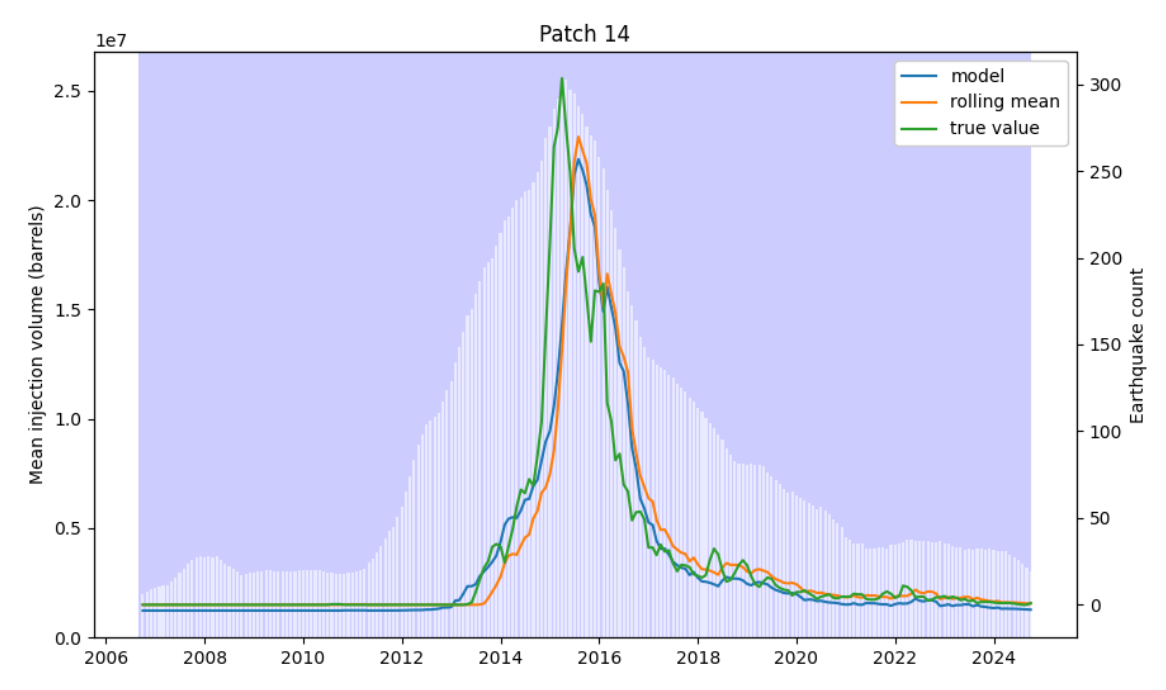


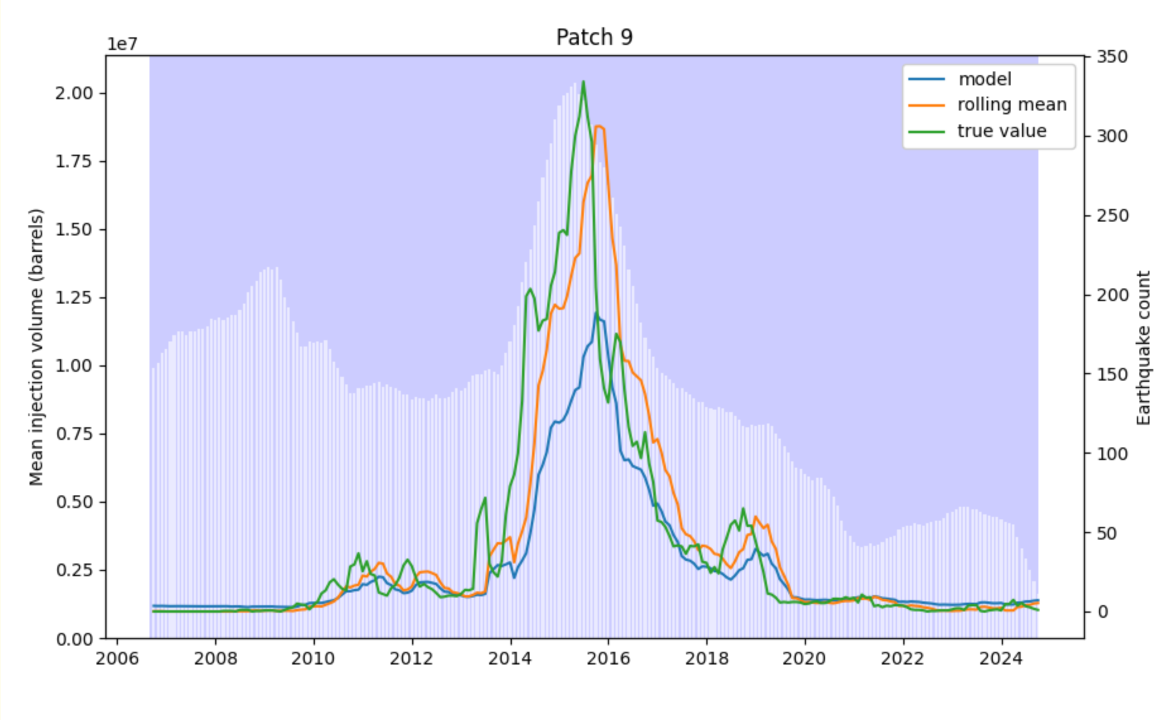
Results (temporal)

- On a 55/30/15 train/test/validation split, **39.04%** improvement over a 6-month rolling means estimate
- On Patch 14, we see a **59.42%** improvement over a 6-month rolling means estimate
- Strong temporal generalization

Patch	RM window (months)	RM train MSE	RM val MSE	RM test MSE	Model train MSE	Model val MSE	Model test MSE	Model relative val improvement
All	**6**	**215.321**	**84.828**	**4.340**	**170.756**	**51.716**	**4.055**	**39.034%**
All	3	170.172	67.147	5.463	128.795	47.750	4.671	28.887%
9	6	1649.308	409.899	5.898	1409.183	171.347	4.995	58.198%
9	3	1309.352	274.569	5.700	1046.254	148.888	4.631	45.773%
13	6	478.811	388.140	5.673	375.706	244.065	5.134	37.119%
13	3	429.906	357.859	4.104	345.936	246.150	3.516	31.216%
14	**6**	**1493.753**	**279.029**	**5.477**	**1009.786**	**113.228**	**4.800**	**59.420%**
14	3	1039.098	108.156	6.178	706.127	55.266	5.067	48.902%







Results (spatial)

- Validating on random patches was too noisy, many patches saw little seismic activity
- Training on 9 and 13, validating on 14: **+26.917%**
- Training on 13 and 14, validating on 9: **–54.368%**
- Spatial generalization remains challenging

Holdout patch	6-month RM validation score	Model validation score	Relative model performance	Average Euclidean patch distance to other patches in the training set
9	1016.713	1569.485	**-54.368%**	1.825 patches
13	378.864	451.153	-19.080%	1.618 patches
14	892.372	652.172	**+26.917%**	1.207 patches

Discussion and Outlook

- Gains relative to rolling mean are modest with the current architecture. Patches may be too large, or attention mechanism too limited in space.

- Transfer learning is challenging. Can we generalize from Oklahoma to Texas despite geological differences?

- We anticipate that better methods for data normalization will improve performance.