Hybrid CNN-LSTM Framework for Predicting Subsurface Energy Production

Jamshed Kaikaus, Mentor: Kate Cauthen (8722), Project Lead: Hongkyu Yoon (8864)

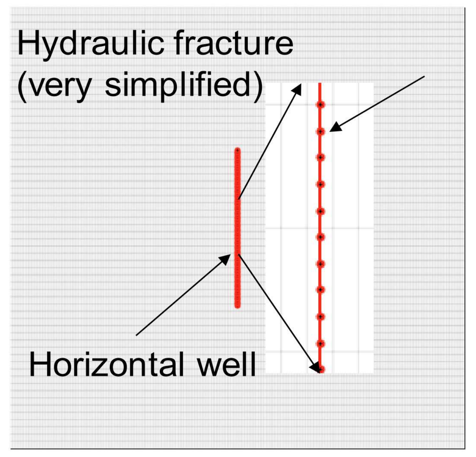
Introduction

Accurate prediction of subsurface energy production necessitates the modeling of complex, multi-scale processes. However, current modeling methods of these processes are often computationally expensive and inflexible.

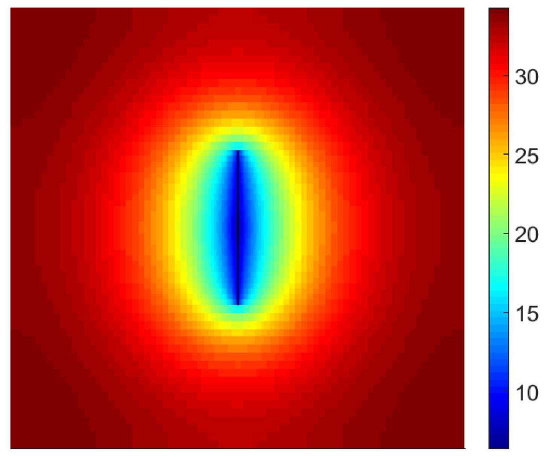
Proposal – This work aims to augment current development of a predictive data-driven platform (PDDP) for subsurface energy systems by researching the effectiveness of a CNN-LSTM Hybrid model architecture towards tackling this spatio-temporal problem.

Dataset

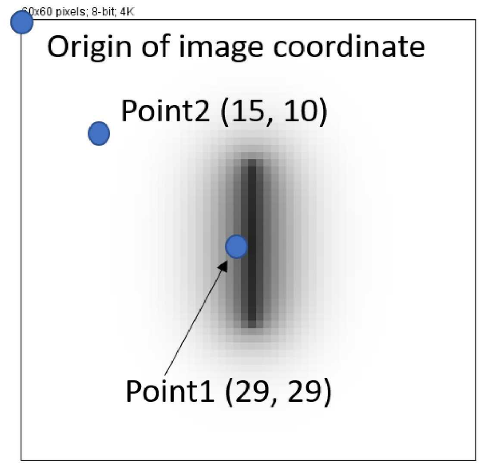
- Data used to model subsurface energy production was obtained using the MRST-Shale simulator.
- The goal of this work is to predict cumulative energy production (*temporal* data) and pressure distribution over time (*spatio-temporal* data)



Horizontal well with hydraulic fractures



Pressure distribution



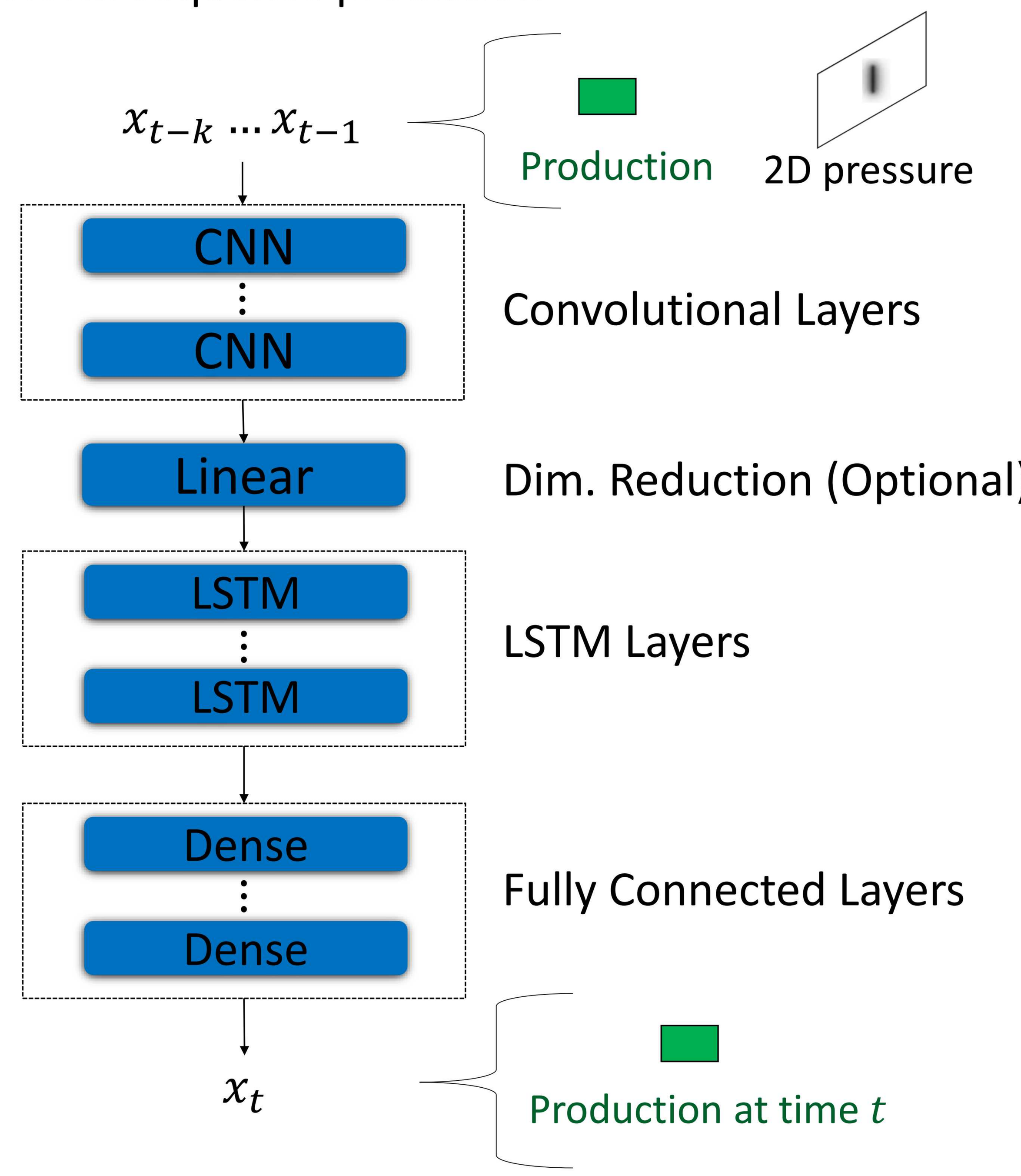
Normalized Pressure dist.

- The dataset contains five features:
  - *Permeability* – measure of interconnectivity
  - *Porosity* – measure of the void spaces
  - *Hydraulic conductivity* – ease with which media can move through fractures
  - *Bottom hole pressure* – pressure acting on the walls of the fracture
  - *Fracture aperture* – perpendicular width of the open fracture
- The dataset contains samples at 30 different time points across 28 total cases
  - Training and validation set – 25 cases
  - Test set – 3 cases

Methodology: CNN-LSTM

The CNN-LSTM hybrid model is an architecture specifically designed to tackle problems where the input data has both spatial and temporal structure

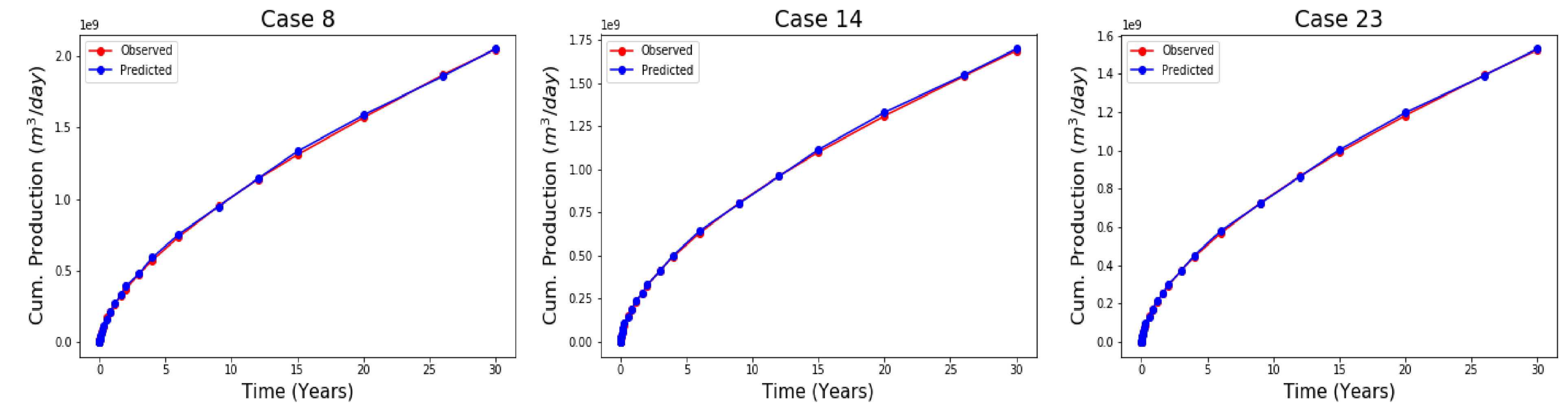
- The CNN Layers perform feature extraction on the multi-dimensional input data
- The LSTM layers interpret these features across time in order to perform sequence prediction



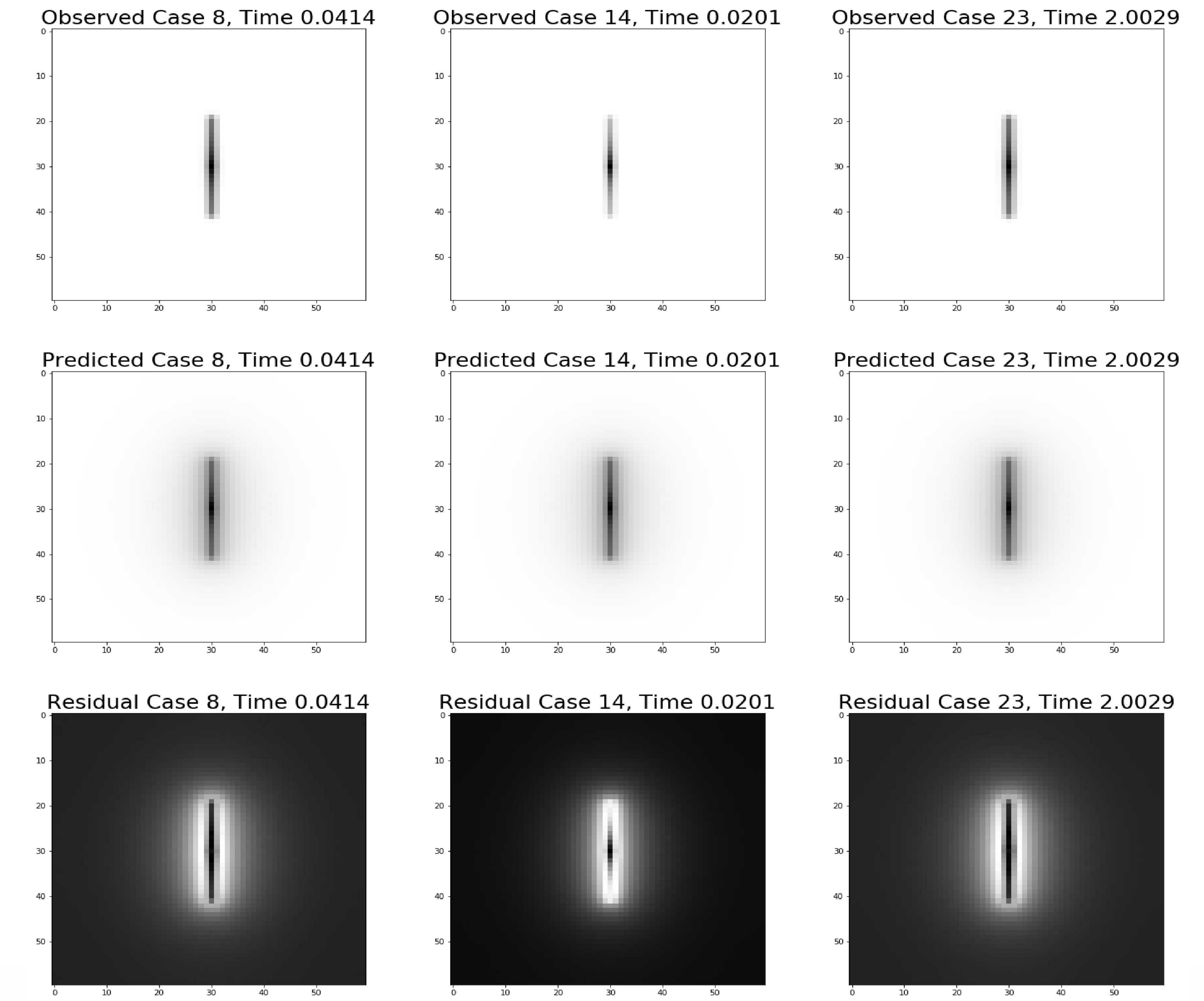
Model Parameters		Models	
		Cum. Energy Production	Pressure Distribution
CNN	# of layers	3	4
	# of hidden units	128, 64, 128	256, 128, 256, 128
	Kernel Size	3x3, 1x1, 3x3	3x3, 1x1, 3x3, 1x1
LSTM	# of layers	2	2
	# of hidden units	128, 64	128, 64
Dim. Reduction Layer		Y	N
Optimizer		Adam	NAdam

Results

Cumulative Energy Production



Pressure Distribution



Discussion

Cumulative Energy Production – Model obtained accurate performance, despite 'lightweight' implementation

Pressure Distribution – Model obtained promising initial results, despite lack of optimization

Future work

- Investigate effect of different model considerations on accuracy, e.g. multi-scale additions for short- and long-term contexts.
- Investigate additional promising model architectures, e.g. Bidirectional LSTMs, etc.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology &Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.
SAND No. ____________

[1] T. N. Sainath, O. Vinyals, A. Senior, and H. Sak, "Convolutional, long short-term memory, fully connected deep neural networks," in *Proc. ICASSP*, 2015, pp. 4580-4584
[2] Q. Zhang, J. C. Lam, V. O. Li, and Y. Han, "Deep-Air: A Hybrid CNN-LSTM Framework for Fine-Grained Air Pollution Forecast," in *arXiv: Signal Processing*, 2020
[3] T. Bogaerts, A. D. Masegosa, J. S. Angarita-Zapata, E. Onieva, and P. Hellinckx, "A graph CNN-LSTM neural network for short and long-term traffic forecasting based on trajectory data", in *Transportation Research Part C: Emerging Technologies, Volume 112*, 202, pp. 62-77