CREWES

UNIVERSITY OF CALGARY
FACULTY OF SCIENCE
Department of Geoscience

PennState
College of Earth and Mineral Sciences

PennState
Institute for Computational and Data Sciences

Physics-guided deep learning for seismic inversion: hybrid training and uncertainty analysis

Jian Sun[1,2], Kristopher Innanen[1], Chao Huang[3]

[1] CREWES Project, University of Calgary

[2] College of Earth and Mineral Sciences, The Pennsylvania State University

[3] School of Ocean and Earth Science, Tongji University

Dec 4, 2020

CREWES Sponsor Meeting

Outline

- Seismic Inversion: physics-based or data-driven?
- Methodology
- Training a Physics-guided Neural Network
- Numerical Examples
- Error & Uncertainty Analysis
- Conclusions & Acknowledgements

Seismic Inversion Methods

- Physics-based deterministic methods:
  - Tomography
  - Full waveform inversion

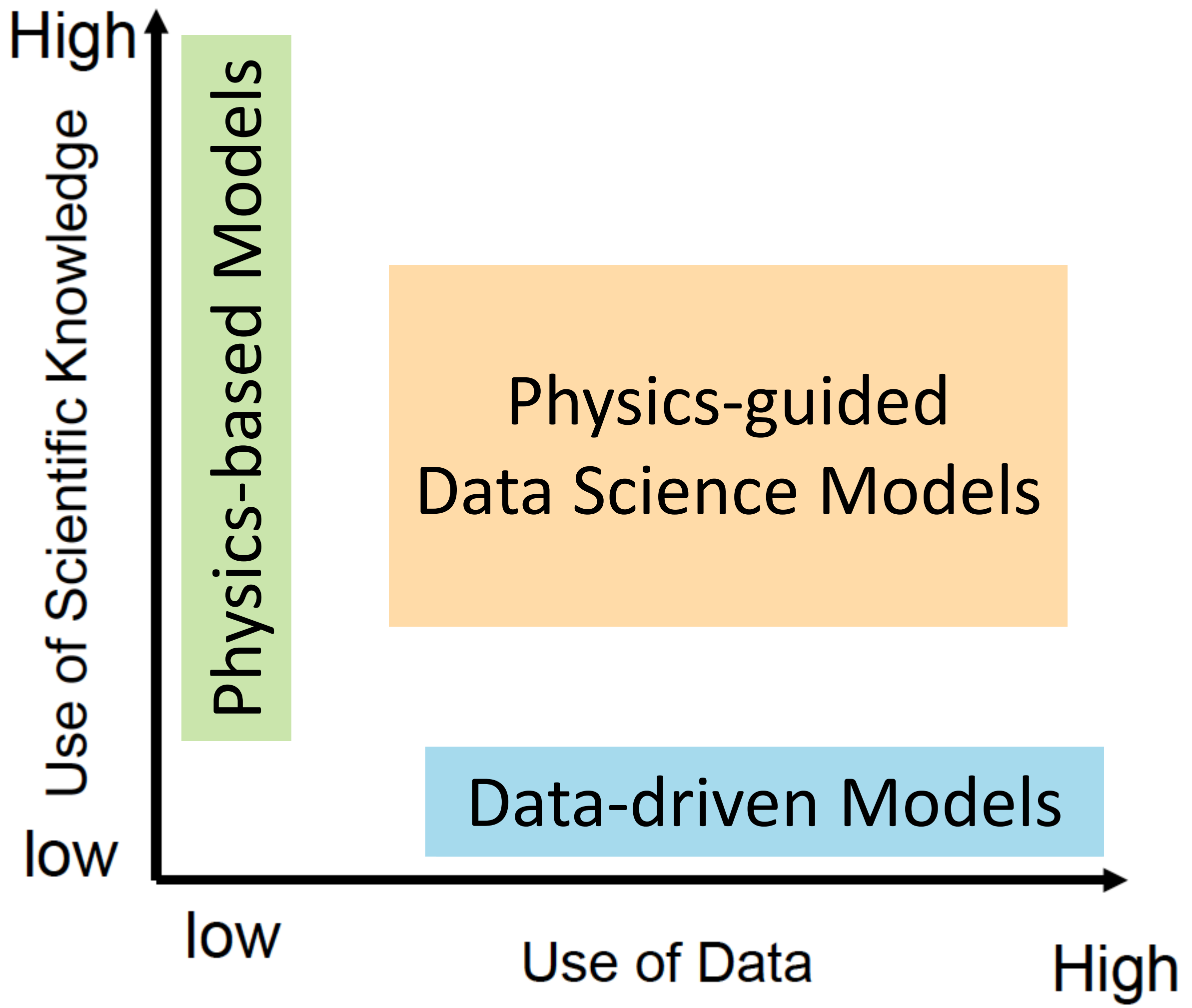
Pros	Cons
✓ High generalization ability	✗ Iterative process
✓ Physics-incorporated	✗ Accurate initial models
✓ Less data requirement	✗ Computational expensive

Seismic Inversion Methods

- Data-driven methods
  - Bayesian models
  - Deep learning

Pros	Cons
✓ Fast and efficient	✗ Big data requirement
✓ End-to-end framework	✗ No physics-involved
✓ Less computational burden	✗ Generalization ability is dependent on training dataset
✓ No expert required	

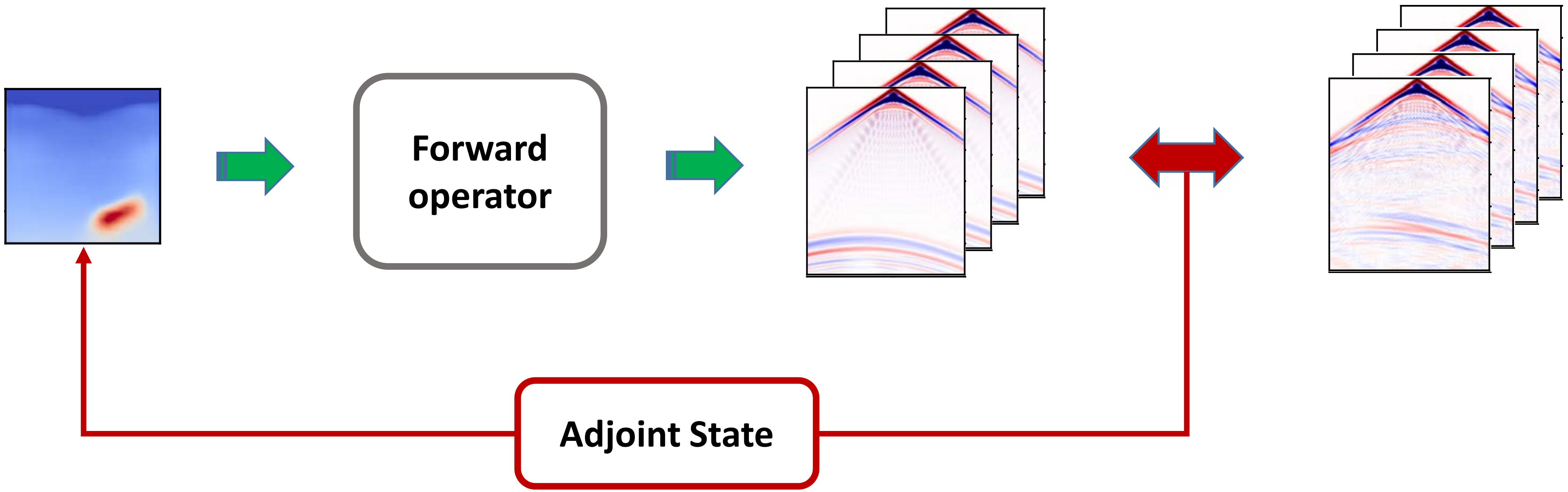
Seismic Inversion Methods



Karpatne et al. 2017

Downton and Hampson, 2018

Methodology: Physics-based



- Objective function: $\mathcal{L}_d^k(\tilde{\mathbf{m}}, \mathbf{d}) = \frac{1}{Tkn_s} \sum_{i=1}^{k} \sum_{s=1}^{n_s} \sum_{t=0}^{T} \left\| \mathbf{d}_i^{s,t} - \mathcal{F}^{s,t}(\tilde{\mathbf{m}}_i) \right\|^2$

Methodology: Physics-based

GEOPHYSICS, VOL. 85, NO. 2 (MARCH-APRIL 2020); P. R87–R99, 19 FIGS., 4 TABLES.
10.1190/GEO2019-0138.1

Check for up

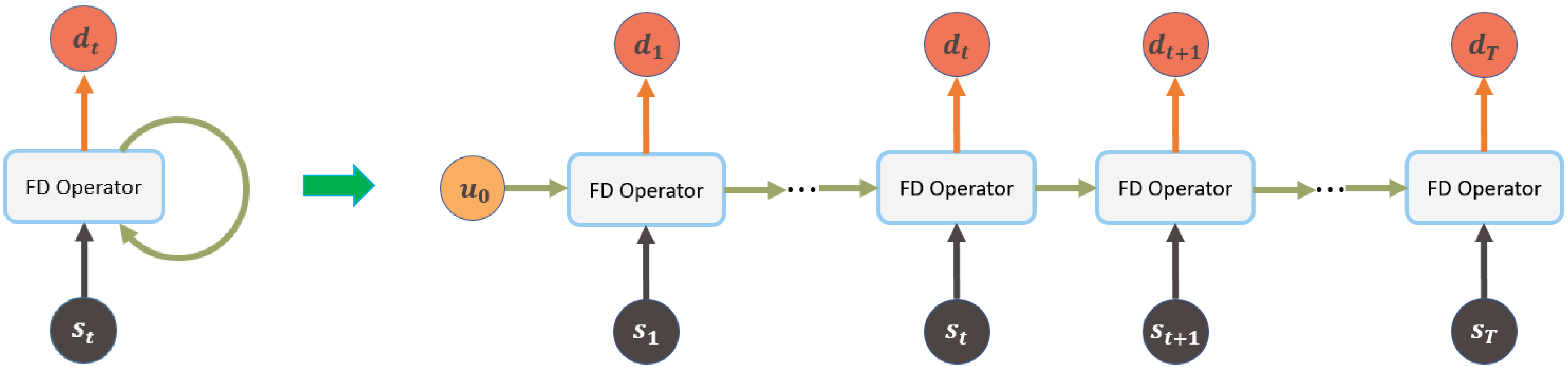
A theory-guided deep-learning formulation and optimization of seismic waveform inversion

Jian Sun[1], Zhan Niu[2], Kristopher A. Innanen[2], Junxiao Li[3], and Daniel O. Trad[2]

ABSTRACT

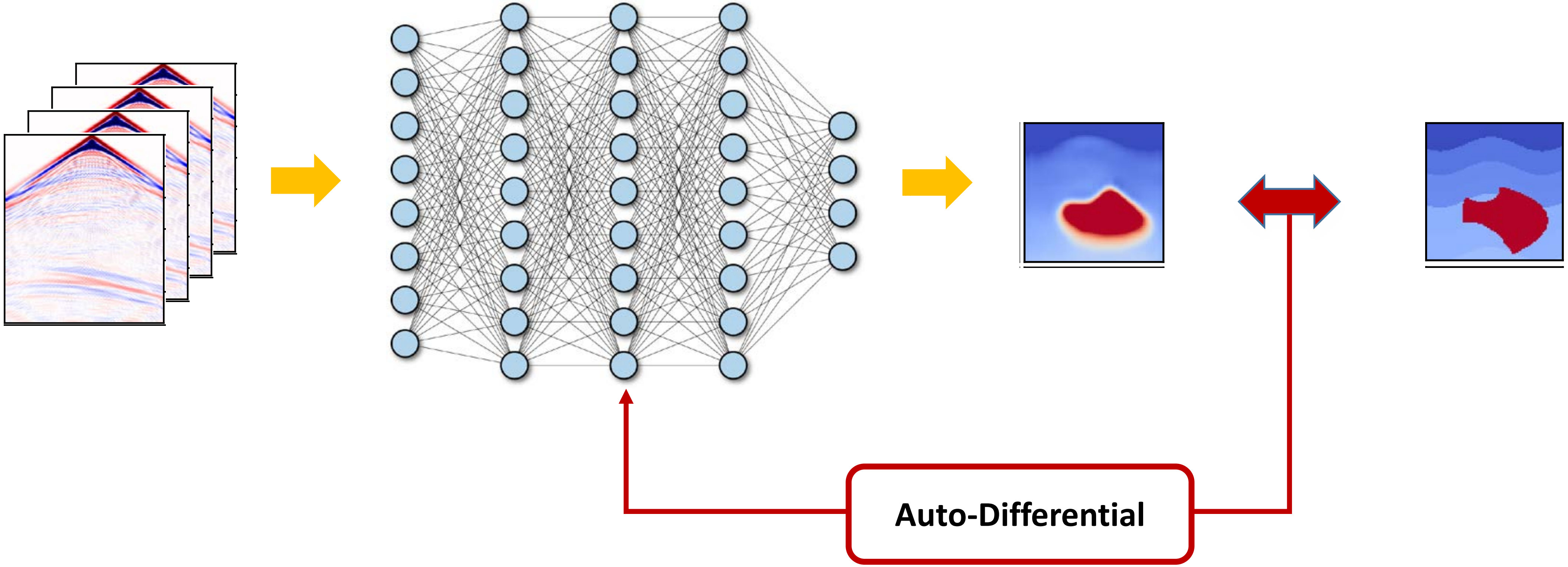
Deep-learning techniques appear to be poised to play very important roles in our processing flows for inversion and interpretation of seismic data. The most successful seismic applications of these complex pattern-identifying networks will, presumably, be those that also leverage the deterministic physical models on which we normally base our seismic interpretations. If this is true, algorithms belonging to theory-guided data science, whose aim is roughly this, will have particular applicability in our field. We have developed a theory-designed recurrent neural network (RNN) that allows single- and multidimensional scalar acoustic seismic forward-modeling problems to be set up in terms of its forward propagation. We find that training such a network and updating its weights using measured seismic data then amounts to a solution of the seismic inverse problem and is equivalent to gradient-based seismic full-waveform inversion (FWI). By refining these RNNs in terms of optimization method and learning rate, comparisons are made between standard deep-learning optimization and nonlinear conjugate gradient and limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) optimized algorithms. Our numerical analysis indicates that adaptive moment (or Adam) optimization with a learning rate set to match the magnitudes of standard FWI updates appears to produce the most stable and well-behaved waveform inversion results, which is reconfirmed by a multidimensional 2D Marmousi experiment. Future waveform RNNs, with additional degrees of freedom, may allow optimal wave propagation rules to be solved for at the same time as medium properties, reducing modeling errors.

Methodology: Physics-based



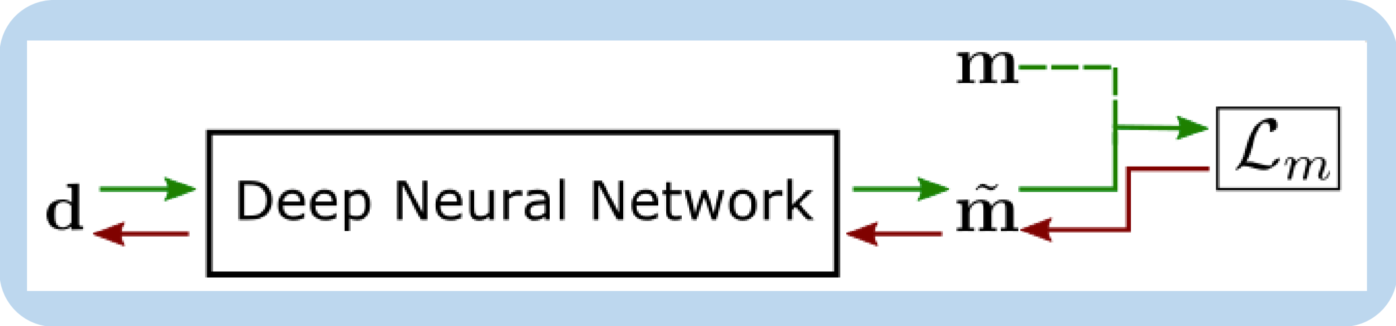
- Objective function: $\mathcal{L}_d^k(\tilde{\mathbf{m}}, \mathbf{d}) = \frac{1}{Tkn_s} \sum_{i=1}^{k} \sum_{s=1}^{n_s} \sum_{t=0}^{T} \left\| \mathbf{d}_i^{s,t} - \mathcal{F}^{s,t}(\tilde{\mathbf{m}}_i) \right\|^2$

Methodology: Data-driven

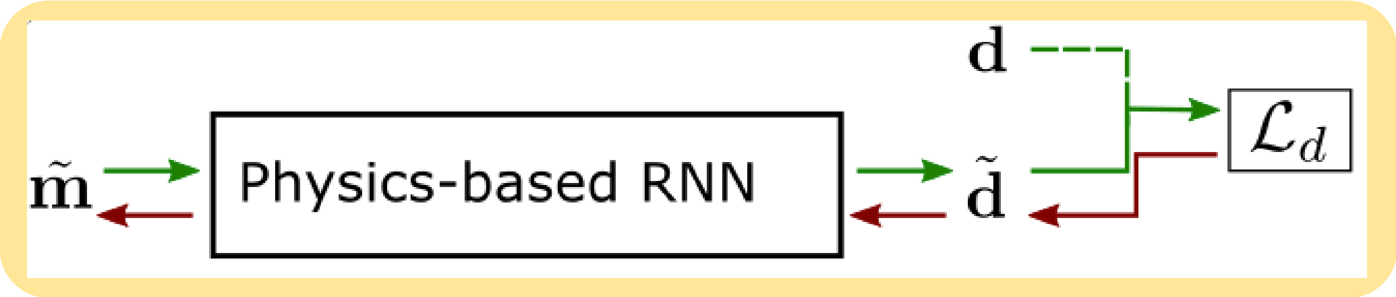


- Objective function: $\mathcal{L}_m^n(\mathbf{m},\mathbf{d}) = \frac{1}{n}\sum_{i=0}^{n}\|\mathbf{m}_i - \tilde{\mathbf{m}}_i\|^2 = \frac{1}{n}\sum_{i=0}^{n}\|\mathbf{m}_i - \Omega(\mathbf{d}_i;\Theta)\|^2$

Methodology: Data-driven & Physics-based

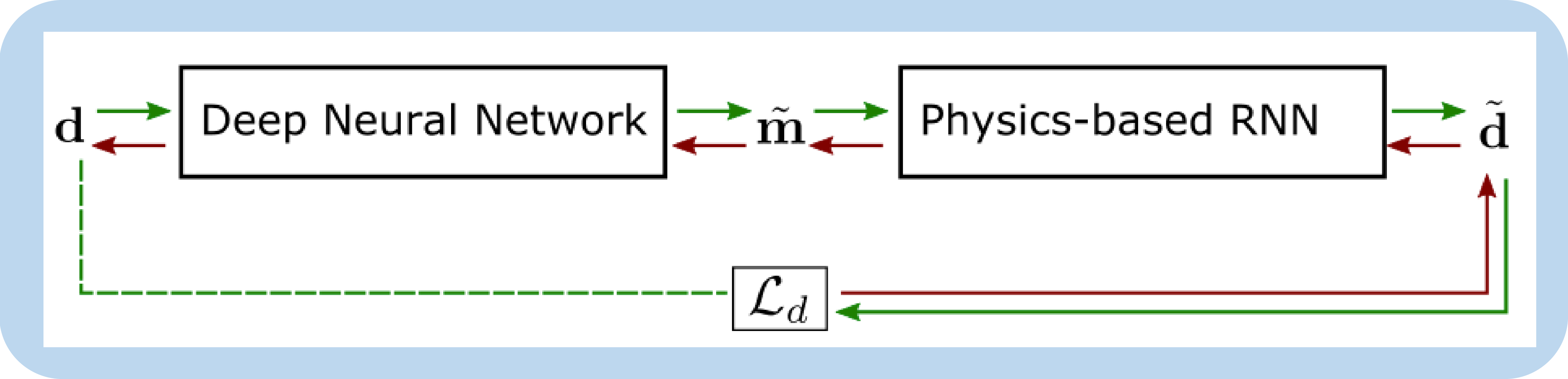


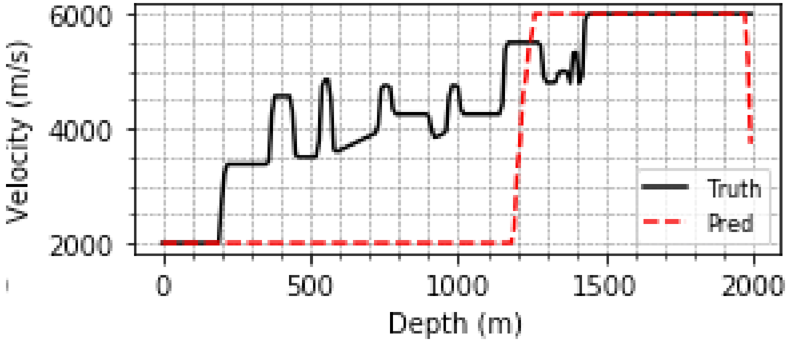
$$\mathcal{L}_m^n(\mathbf{m}, \mathbf{d}) = \frac{1}{n} \sum_{i=0}^{n} \|\mathbf{m}_i - \Omega(\mathbf{d}_i; \Theta)\|^2$$



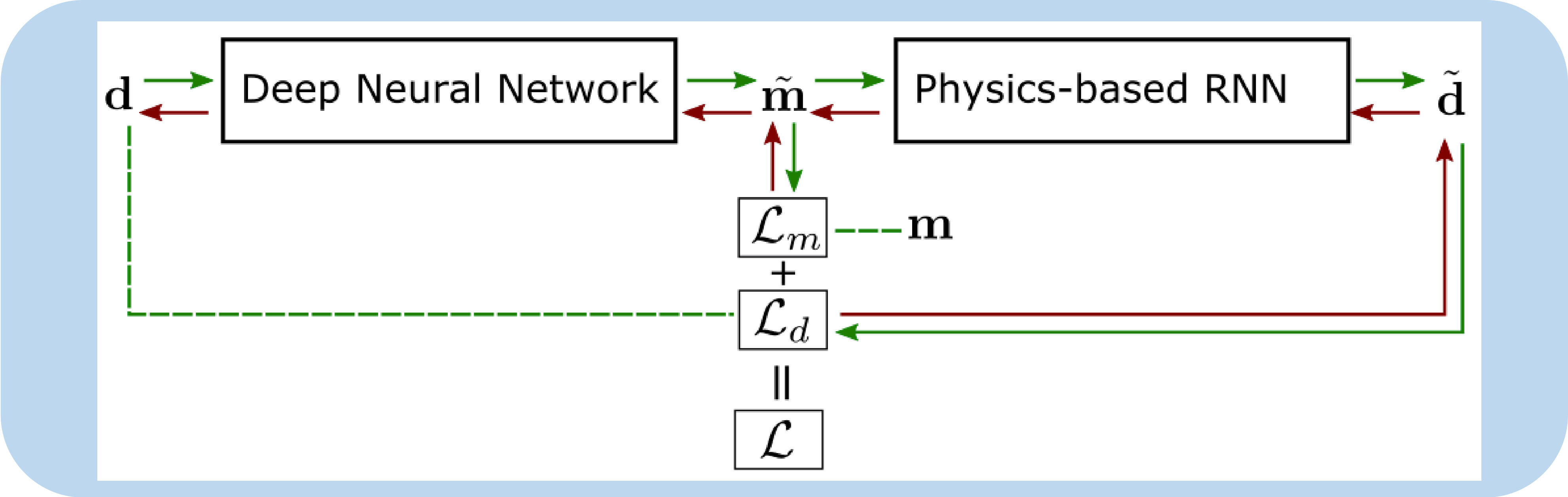
$$\mathcal{L}_d^k(\tilde{\mathbf{m}}, \mathbf{d}) = \frac{1}{Tkn_s} \sum_{i=1}^{k} \sum_{s=1}^{n_s} \sum_{t=0}^{T} \left\|\mathbf{d}_i^{s,t} - \mathcal{F}^{s,t}(\tilde{\mathbf{m}}_i)\right\|^2$$

Methodology: Physics-guided



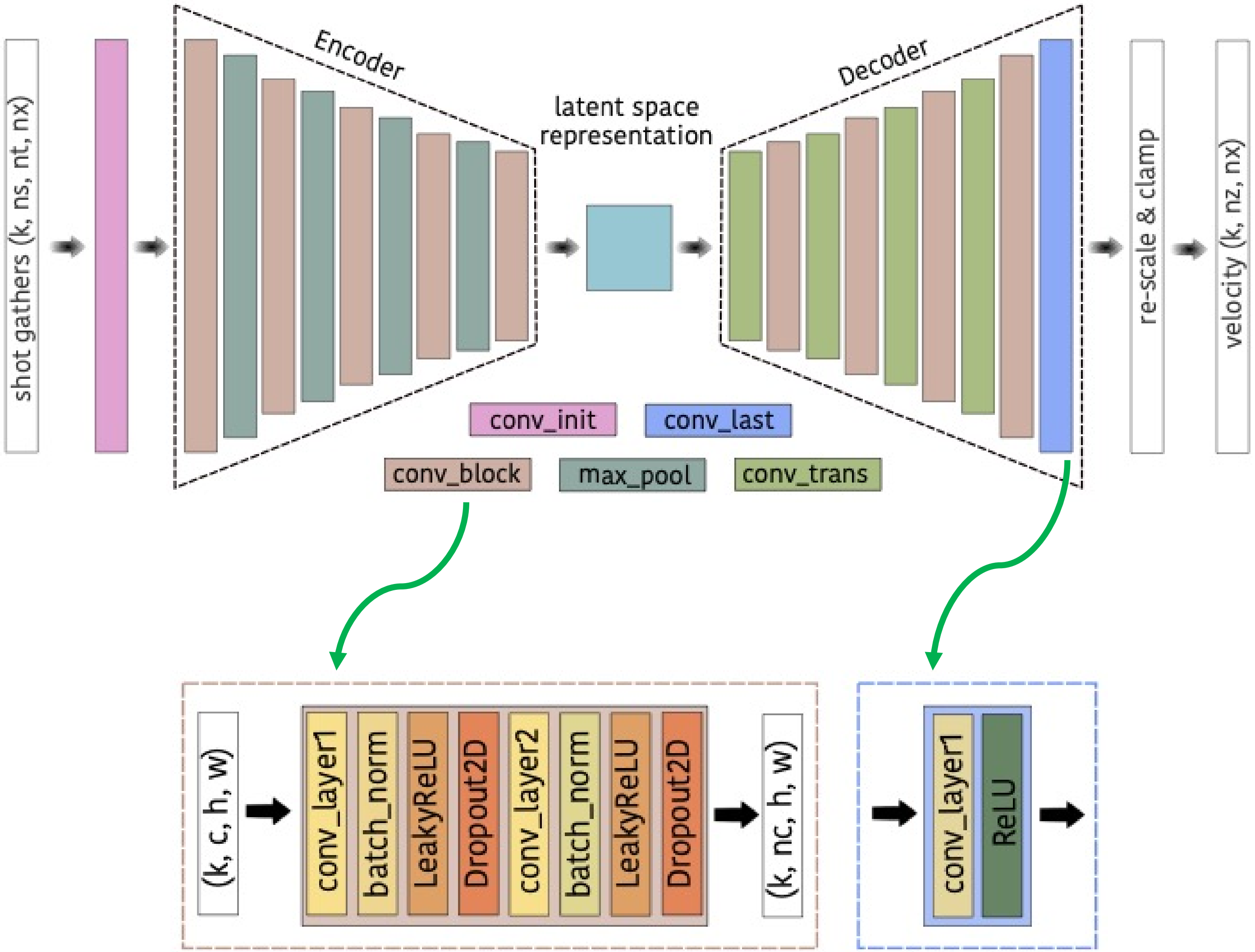


Methodology: Hybrid Physics-guided

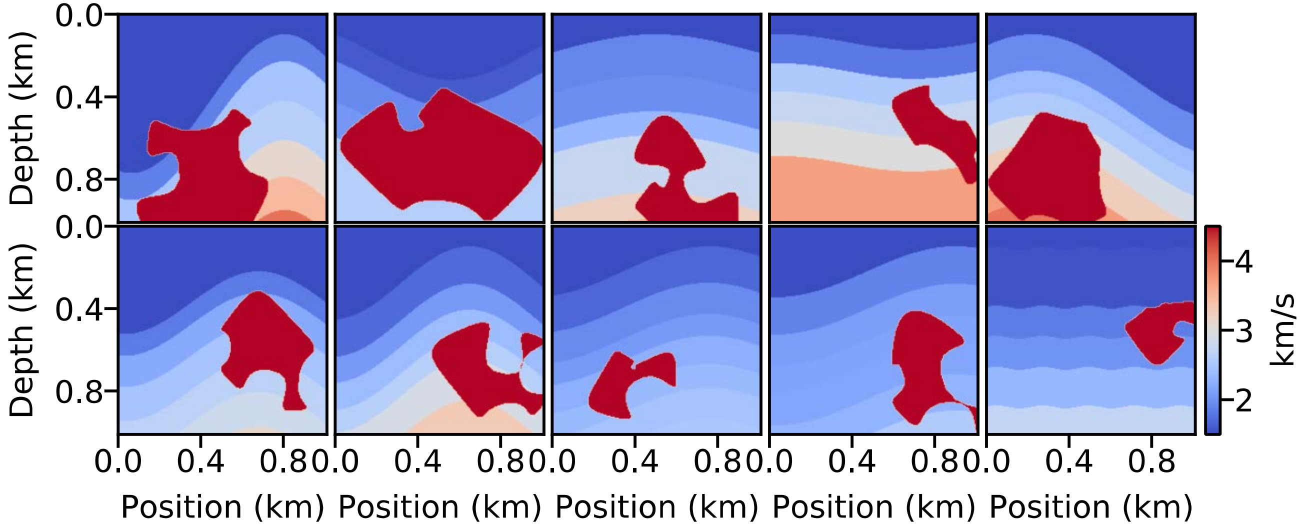


$$\mathcal{L}^k = \lambda_m \tilde{\mathcal{L}}_m^k + \lambda_d \tilde{\mathcal{L}}_d^k$$

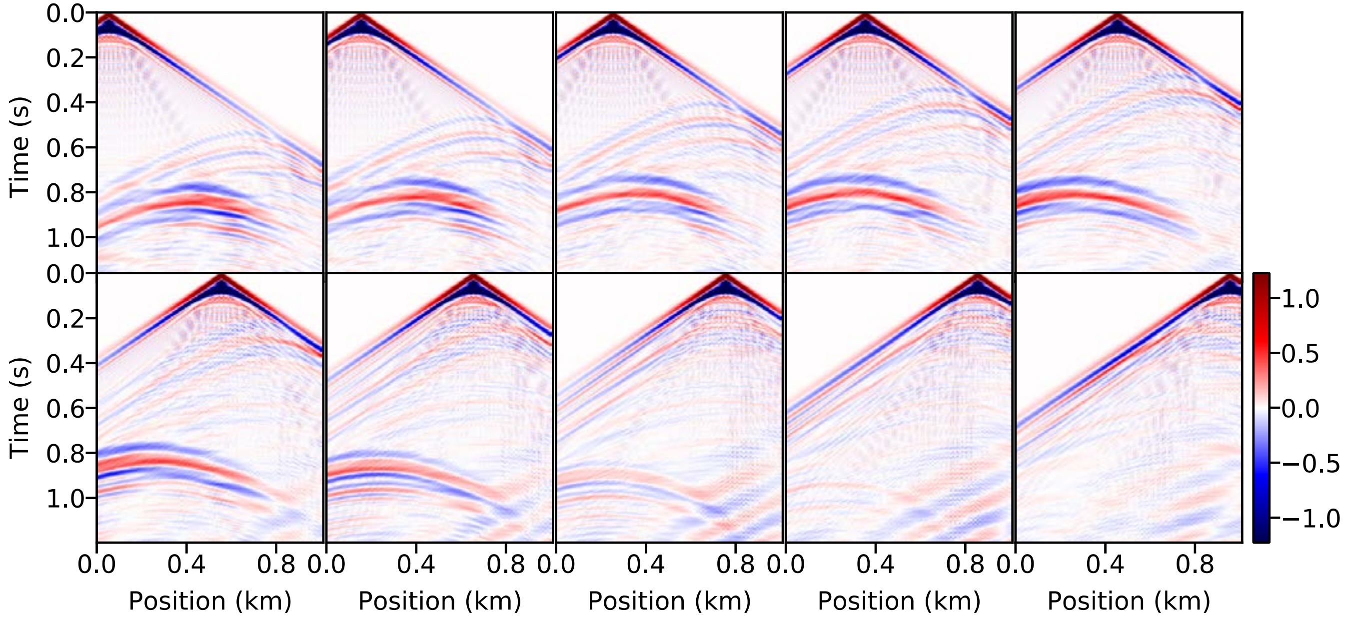
Methodology: Network Architecture



Training a Physics-guided Neural Network (PGNN)



10 Synthetic Salt Models



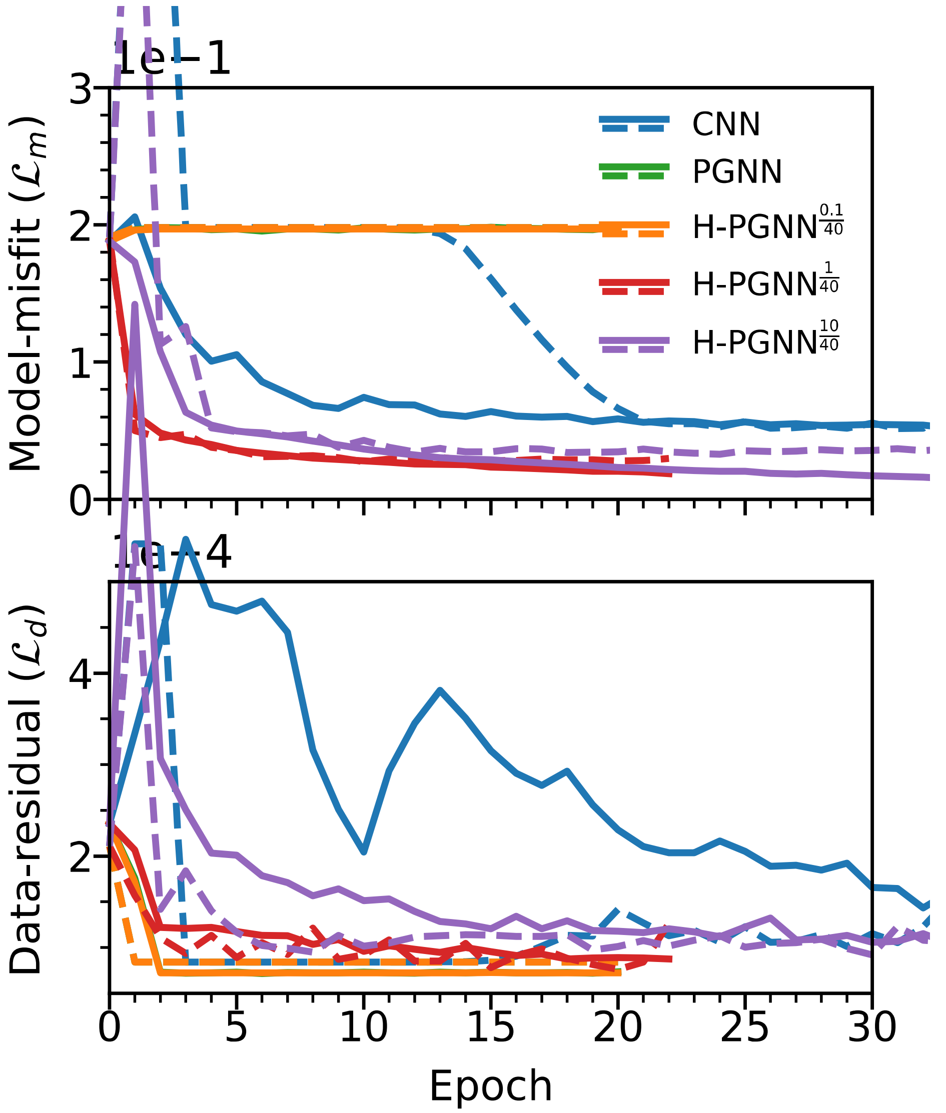
10 Shot gathers

Training a Physics-guided Neural Network (PGNN)

- Trade-off parameters

$$\mathcal{L}^k = \lambda_m \tilde{\mathcal{L}}_m^k + \lambda_d \tilde{\mathcal{L}}_d^k$$

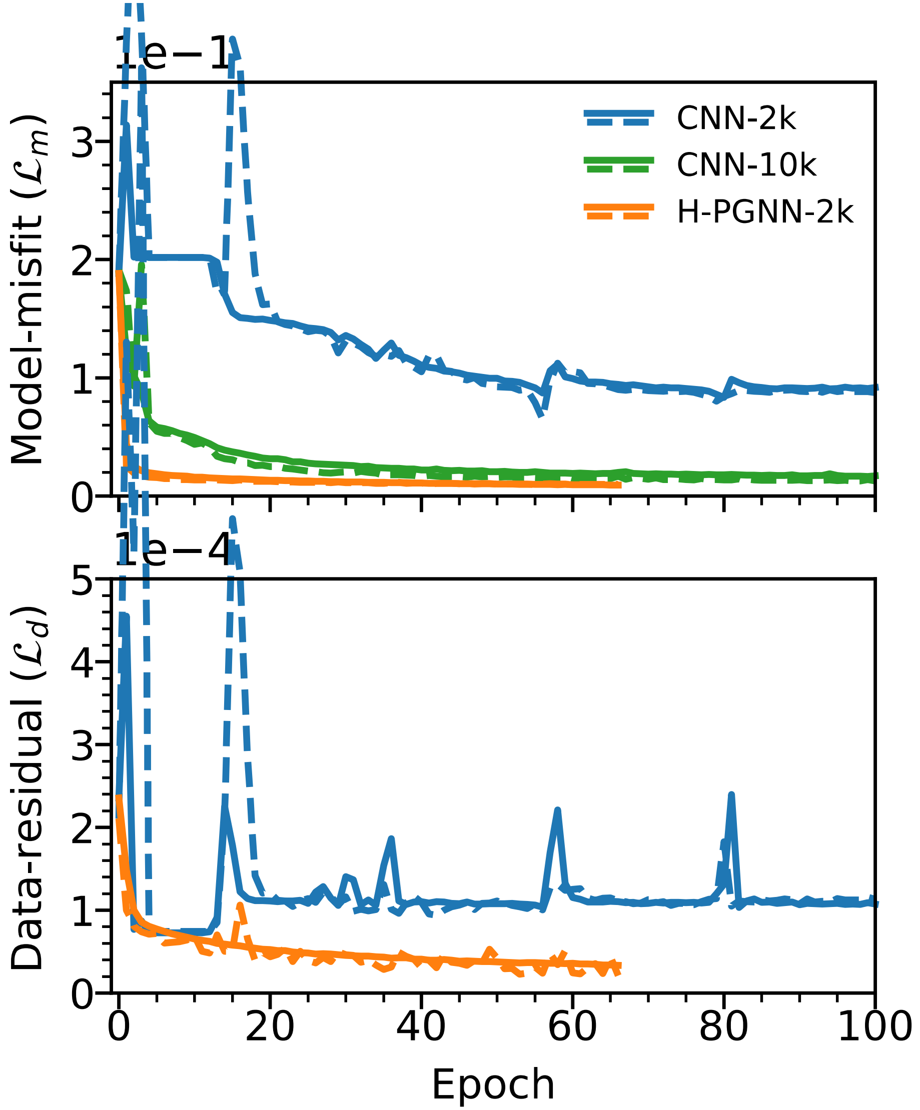
Parameters / Mode	Number of samples in Train (Val)	Batch-size in Train (Val)	Trade-off (λ_m, λ_d)
CNN	320 (120)	16 (64)	(1, 0)
PGNN	320 (120)	16 (32)	(0, 1)
H-PGNN $\frac{0.1}{40}$	320 (120)	16 (32)	(0.1, 40)
H-PGNN $\frac{1}{40}$	320 (120)	16 (32)	(1, 40)
H-PGNN $\frac{10}{40}$	320 (120)	16 (32)	(10, 40)



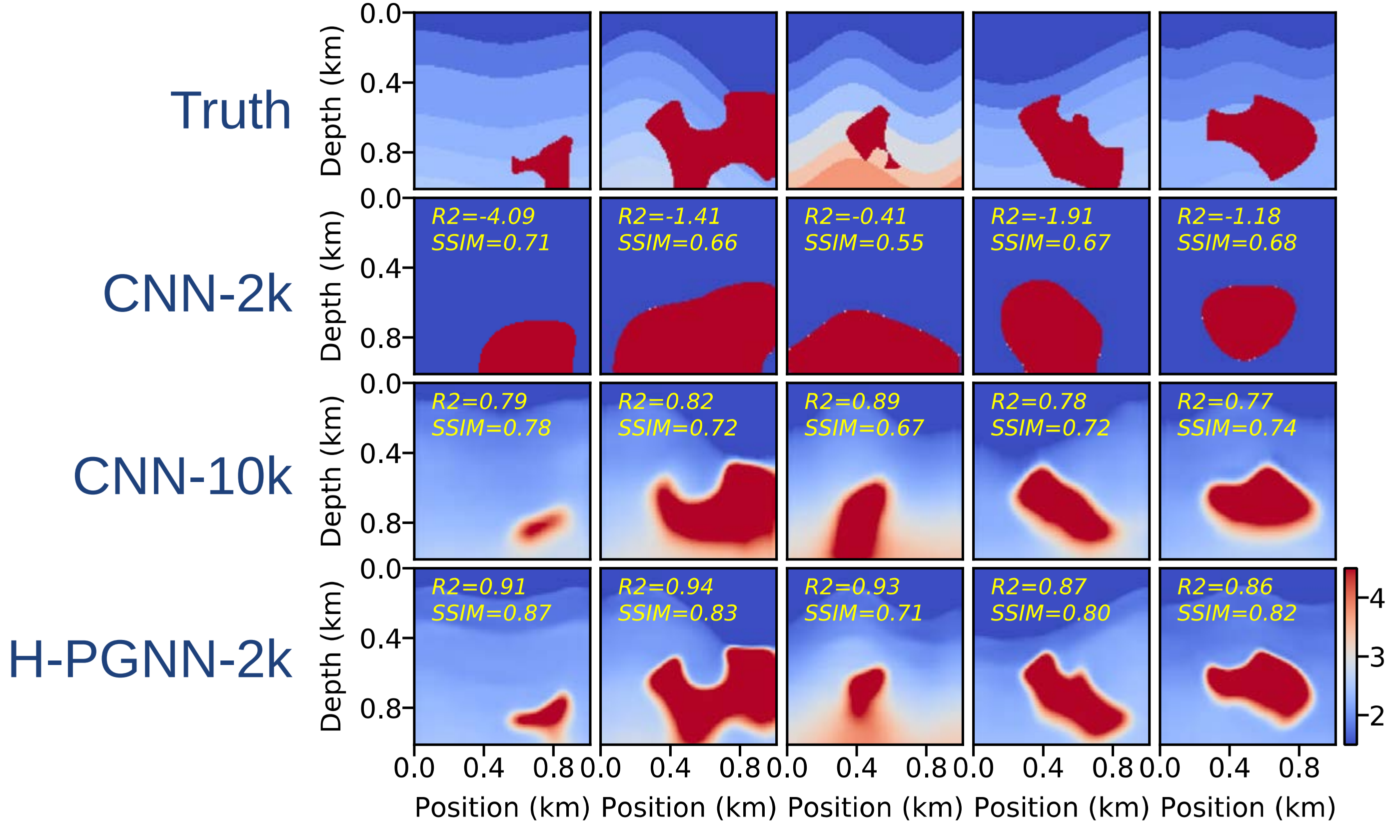
Numerical Examples: Training & Validation

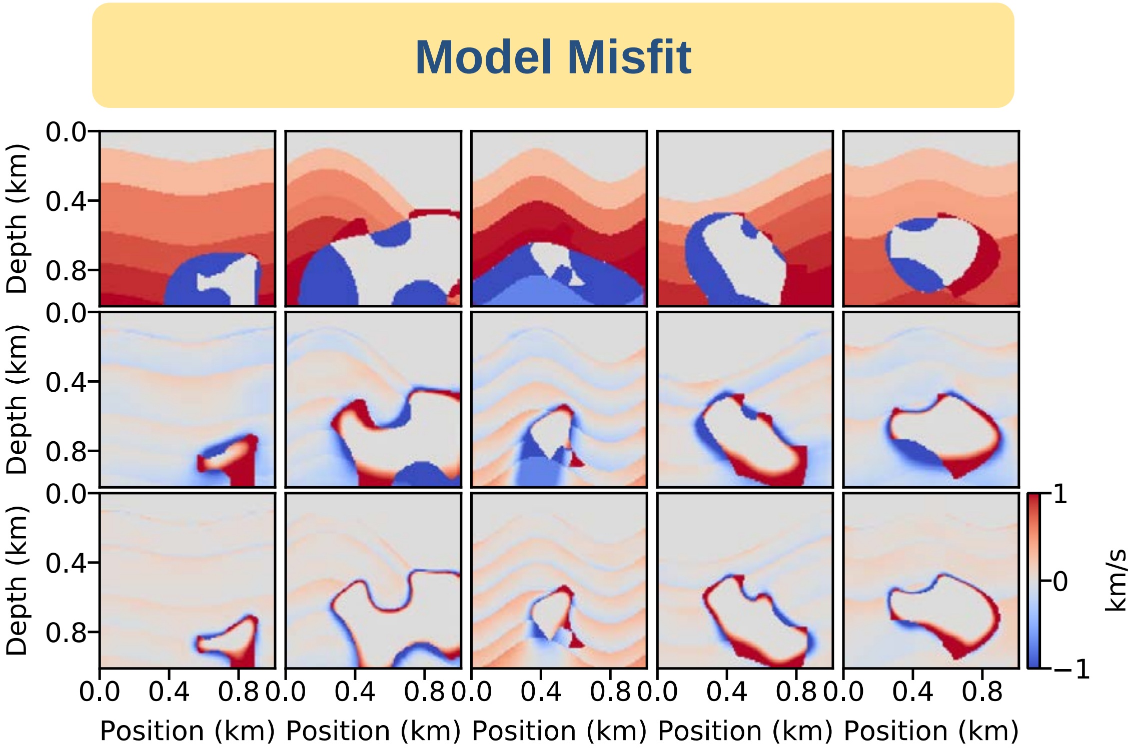
$$\mathcal{L}^k = \lambda_m \tilde{\mathcal{L}}_m^k + \lambda_d \tilde{\mathcal{L}}_d^k$$

Parameters / Mode	Number of samples in Train (Val, Test)	Batch-size in Train (Val, Test)	Trade-off (λ_m, λ_d)	Training time-cost (per epoch 10GPUs)
CNN-10k	10000 (3000, 3000)	32 (64, 64)	(1, 0)	2.58mins
CNN-2k	2000 (600, 3000)	32 (64, 64)	(1, 0)	0.75mins
H-PGNN-2k	2000 (600, 3000)	16 (32, 64)	(1→0.1, 40)	20.3mins

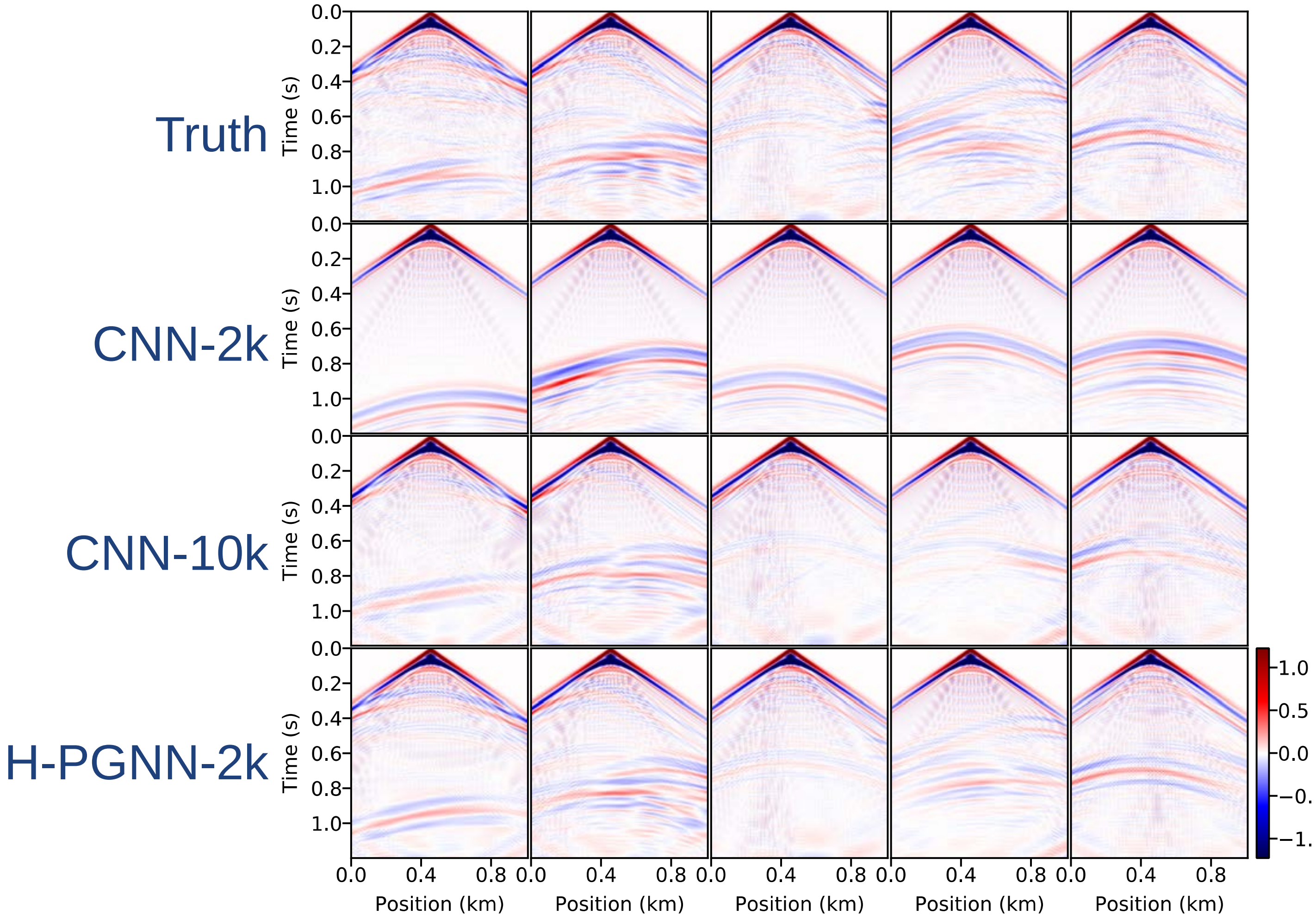


Numerical Examples: Test Results

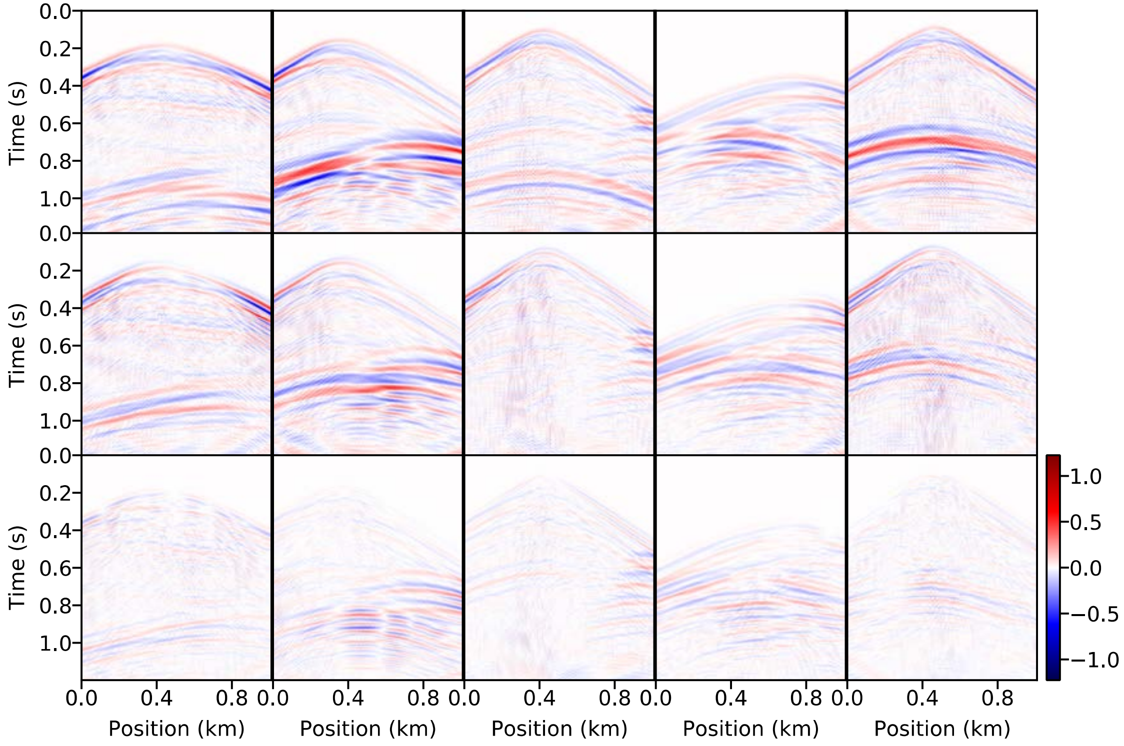




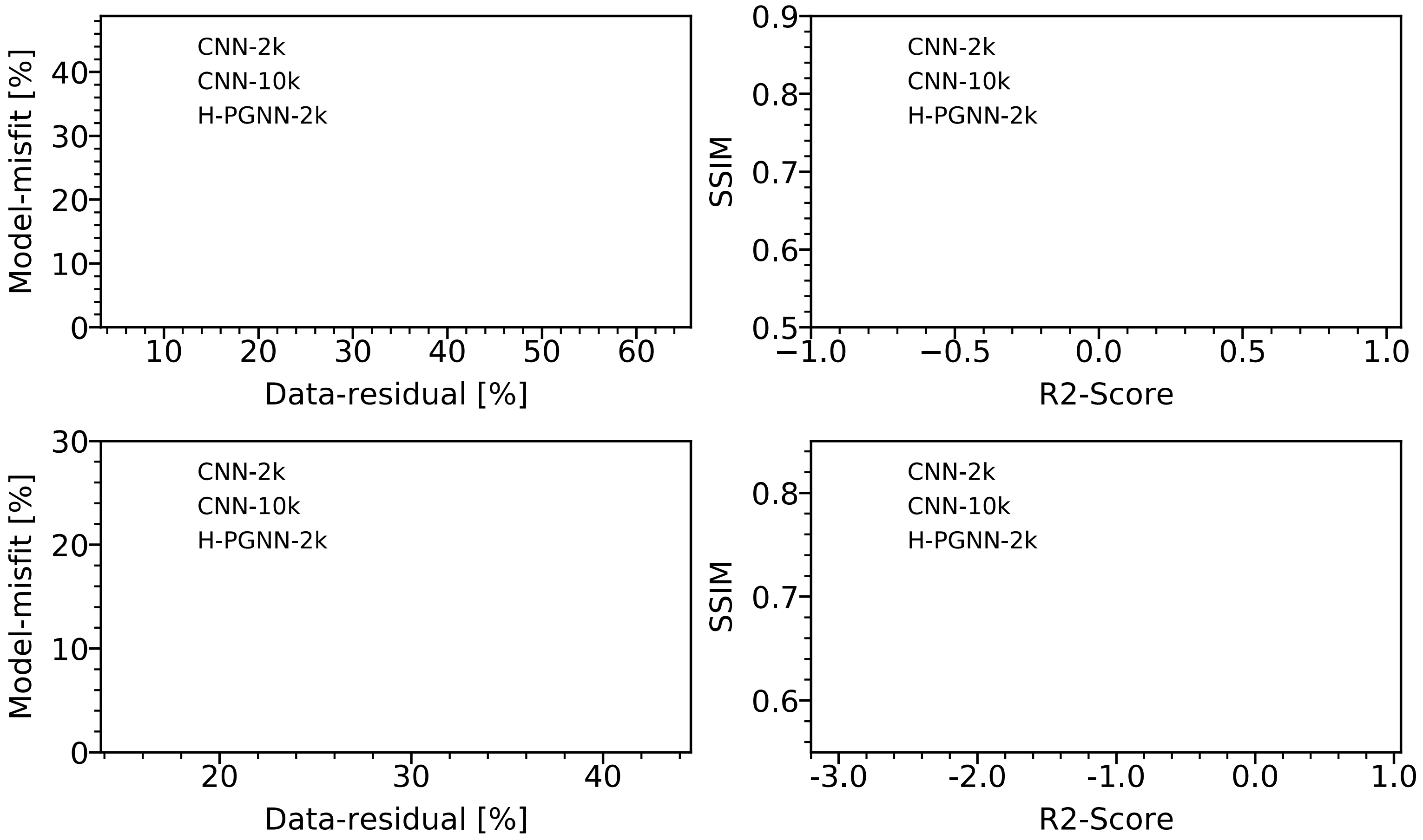
Numerical Examples: Test Results



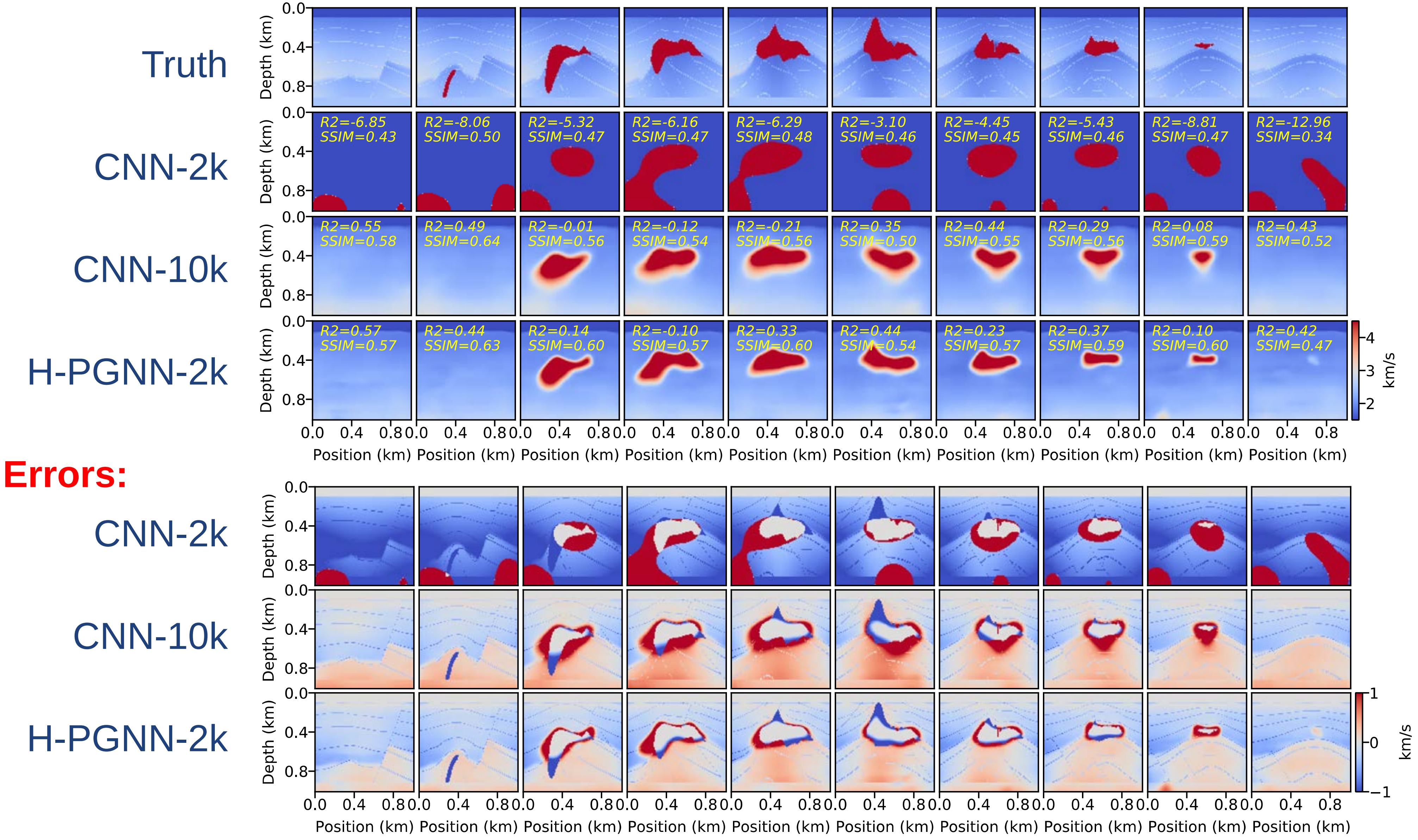
Data residual



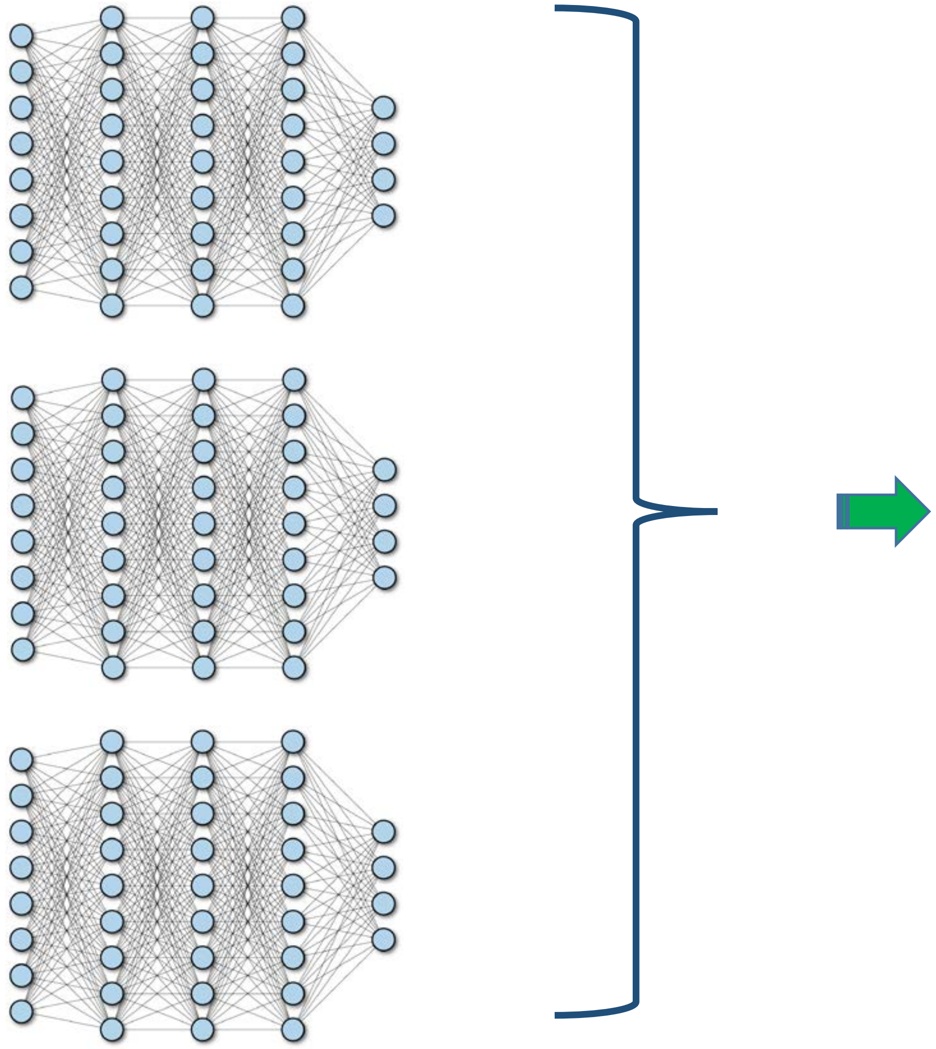
Error Analysis

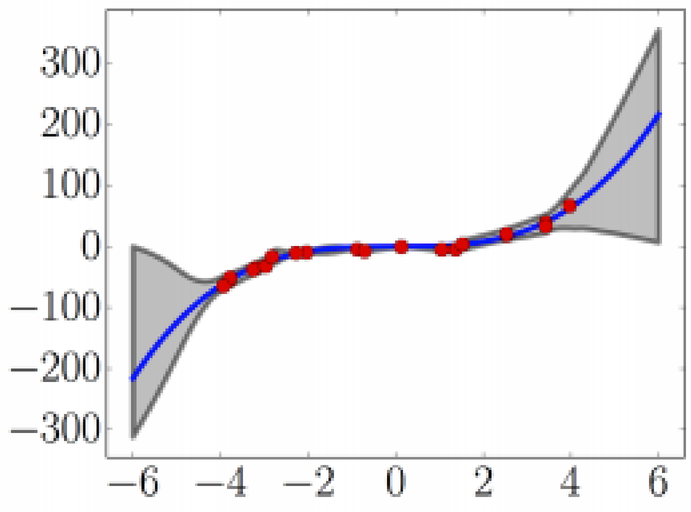


Generalization Ability

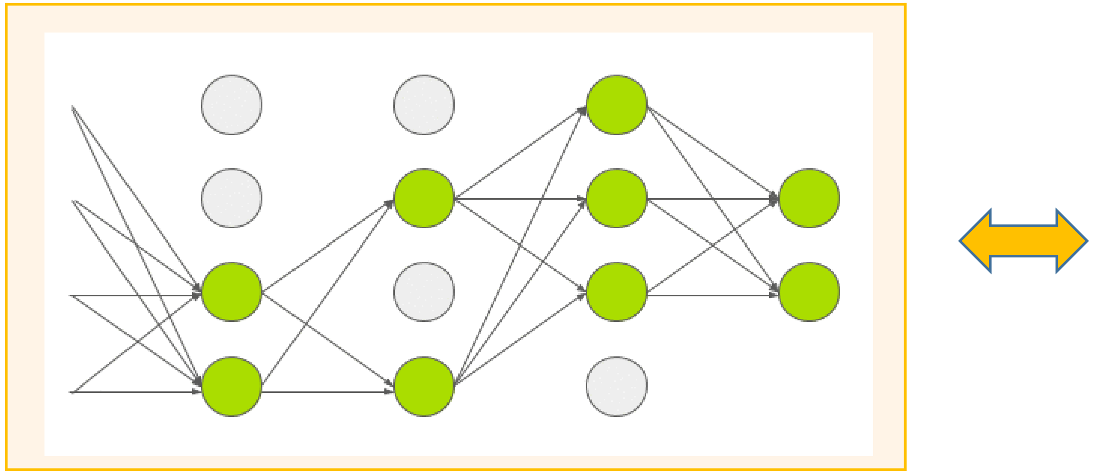


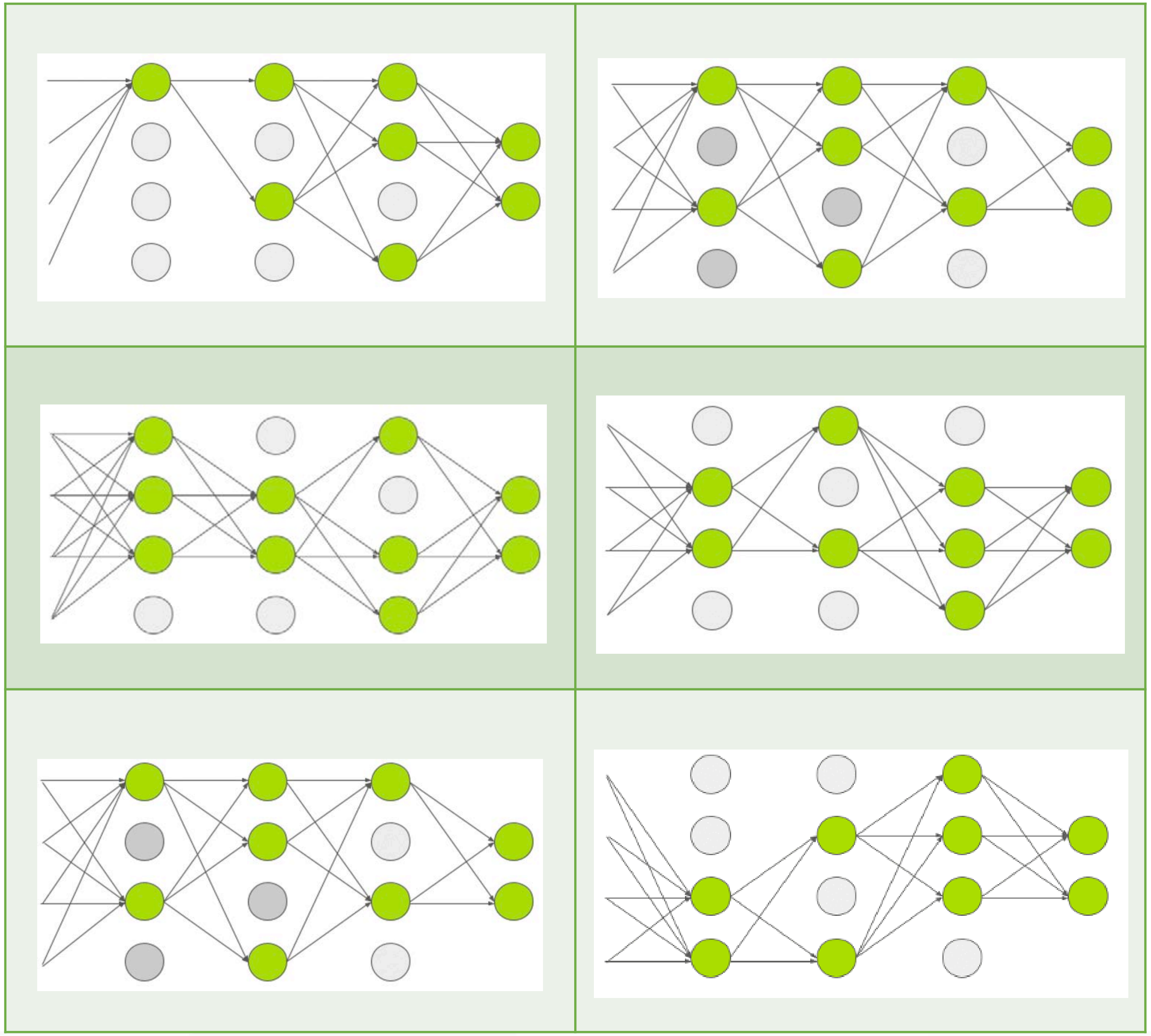
Uncertainty Analysis: Deep Ensembles





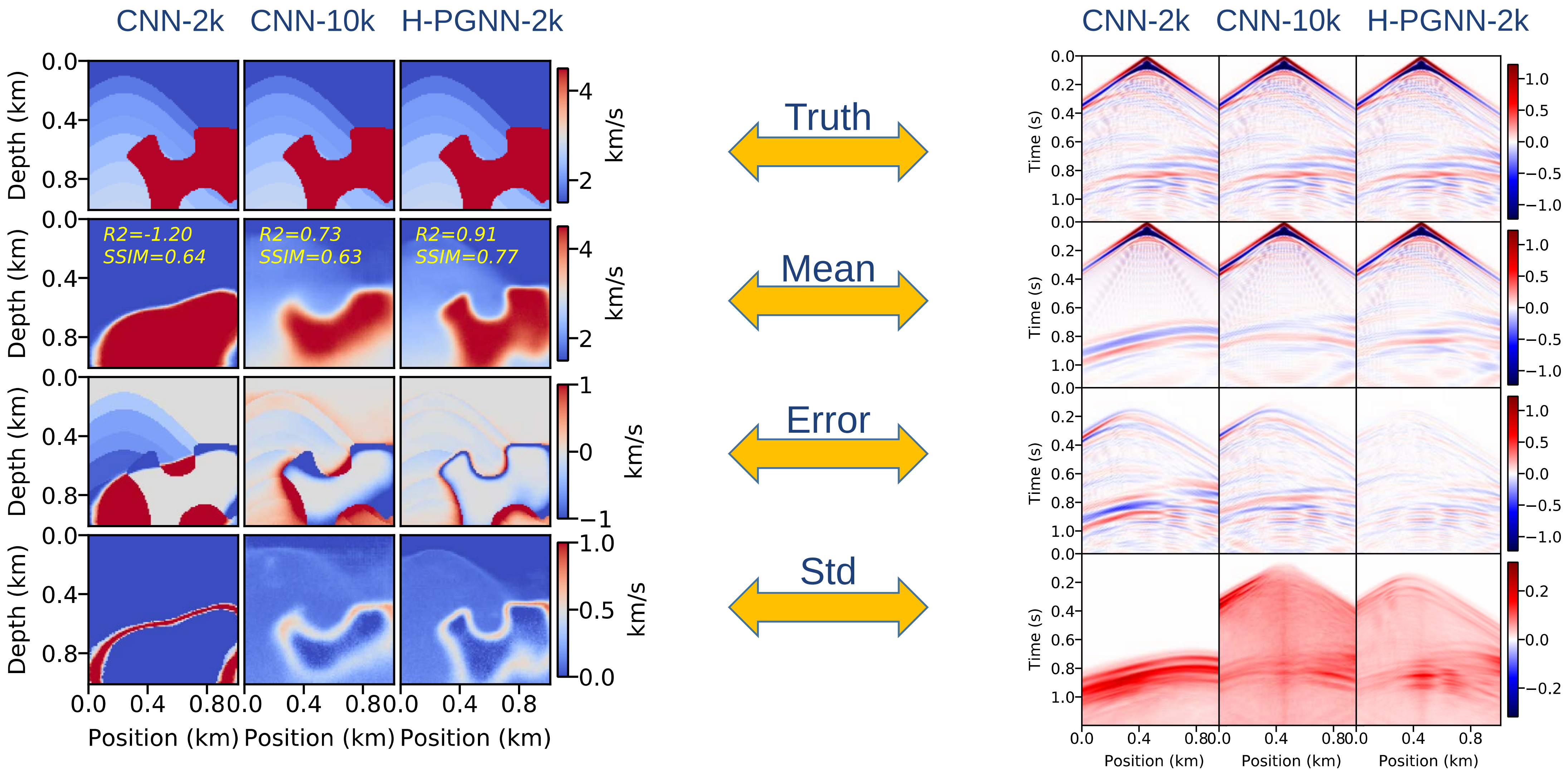
Uncertainty Analysis: Dropout



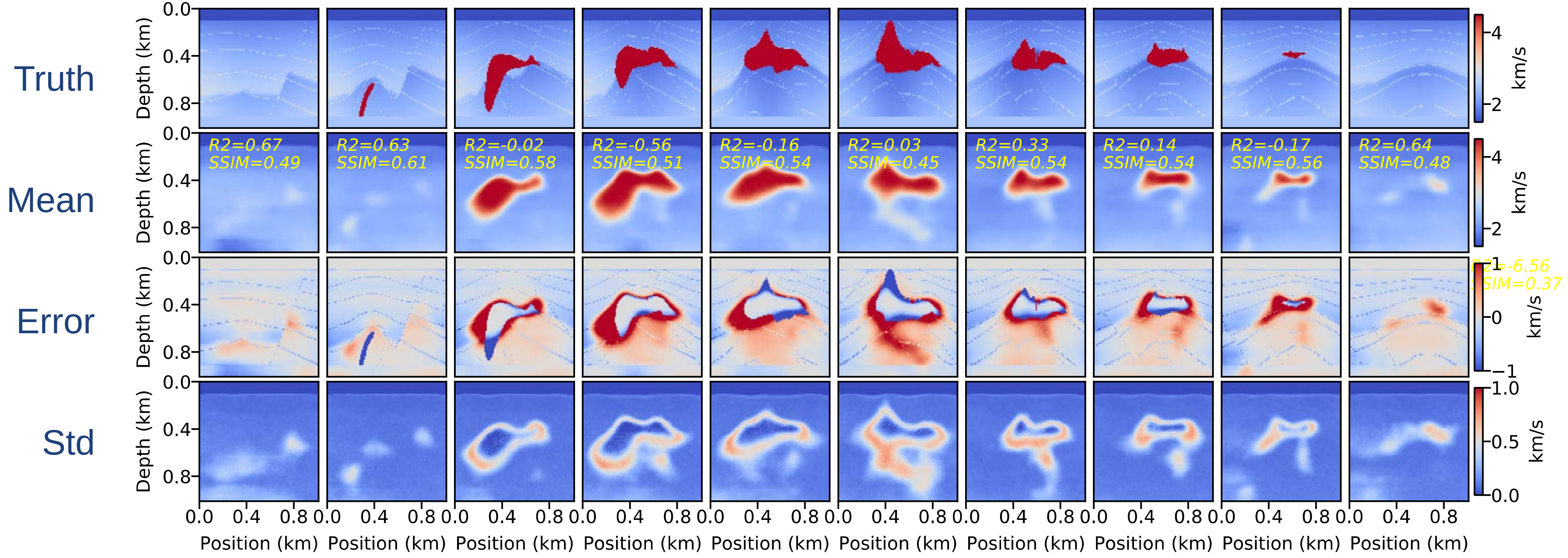


Gal and Ghahramani, 2016

Uncertainty Analysis: Dropout



Uncertainty Analysis: SEG salt Models



Conclusions

- We propose a hybrid physics-guided neural network for seismic inversion by simultaneously reducing both model misfit and physics-based data residual.
- The trade-off parameter selection is essential and analyzed.
- We perform error and uncertainty analysis of deep neural network predictions.
- Given the same training dataset, the physics-guided network outperforms the fully data-driven network:
  - Higher accuracy (smaller model misfit)
  - Obey Physics laws (smaller data residual)
  - Lower uncertainty
  - Less anomalies

Acknowledgements

- Roar Supercomputer (Penn State)
- CREWES Sponsors
- NSERC

Thank you!

Q & A