IMPERIAL

RAPIDS

Rapid AI-Powered Image-to-Dynamic Simulations

RAPIDS team: Aniket Joshi, Donghu Guo, Nathalie C Pinheiro, Yueyan Li, Hakancan Ozturk, Anthony Tran, Boyang Chen, Claire Heaney, Fangxin Fang, Christopher C Pain

Outline

RAPIDS
Rapid AI-Powered Image-to-Dynamic Simulations

- CNN-based generative surrogate fluid flow model
- Domain decomposition with accelerator embedded
- Multi-GPU training and inference capability
- Diffusion architecture embedding underway

RAPIDS
Background

- Need for a general-purpose physics-driven generative surrogate model capable of handling any scenario.
- Flexibility in publicly available AI for science software limited by scope. RAPIDS expands the scope in a highly extensible and modular codebase.

RAPIDS

Structural architecture

RAPIDS

Structural architecture

Flow Field Prediction System Architecture

Models

UNet
Dynamic 2D/3D, skip connections, residual option

UNet++
Nested connections, deep supervision

Data Handling

ChunkedFlowFieldDataset
Memory-mapped arrays, on-demand chunks

HDF5 Support
Multiple simulation files, format conversion

Training Pipeline

FlowFieldPredictor
Warmup, checkpointing, multi-simulation

Accelerated Training
Distributed training, GPU optimization

Prediction Modes

Implicit Mode
Predict t+2 using t and t+1 states

Autoregressive Mode
A → B_pred → C_pred → D_pred (prediction chain)

Masking Strategies

Boundary
Preserve edges

Progressive
Iterative refinement

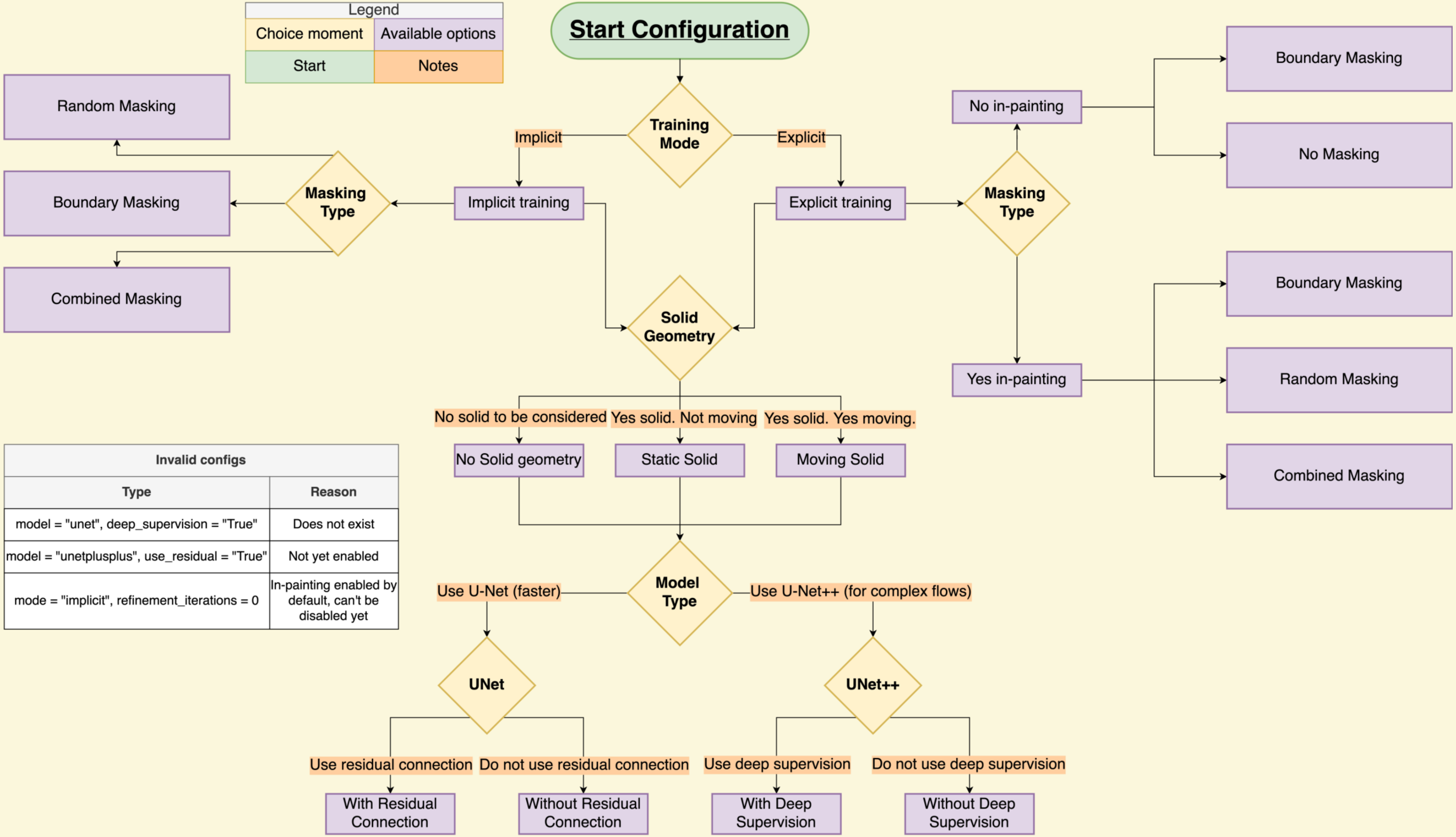
Random
Stochastic masking

Special Features

Solid Handling
Static & moving solid boundaries

Unified 2D/3D
Dimension-agnostic processing

Key Focus: Iterative refinement with solid boundary handling for accurate fluid flow prediction



Invalid configs

Type	Reason
model = "unet", deep_supervision = "True"	Does not exist
model = "unetplusplus", use_residual = "True"	Not yet enabled
mode = "implicit", refinement_iterations = 0	In-painting enabled by default, can't be disabled yet

RAPIDS
Structural architecture

- Informed choices are the goal.
- Prioritising efficiency– NumPy arrays to memmap to HDF5. RAPIDS handles all.
- Models dockerised (optional) for interoperability.
- Foundational software for foundational AI models.
- Object oriented with extensive documentation, support, and flexibility.
- MINTY codebase– **M**aintainable, **I**ntuitive, **N**eat, **T**estable, **Y**ours.

Feature	NumPy Array	memmap	HDF5
Fits in RAM	Yes	No (not required)	No (not required)
Access speed	Fast (all in memory)	Slower (disk reads)	Efficient (partial loading)
Compression	No	No	Yes
Hierarchy	No	No	Yes (groups/ datasets)
Use Case	Small to medium data	Huge flat arrays	Huge dataset, archives, simulations

RAPIDS

Capabilities

Models

- Classic Vanilla U-Net (with and without residual connections)
- U-Net++ (with and without deep supervision)

RAPIDS
Models

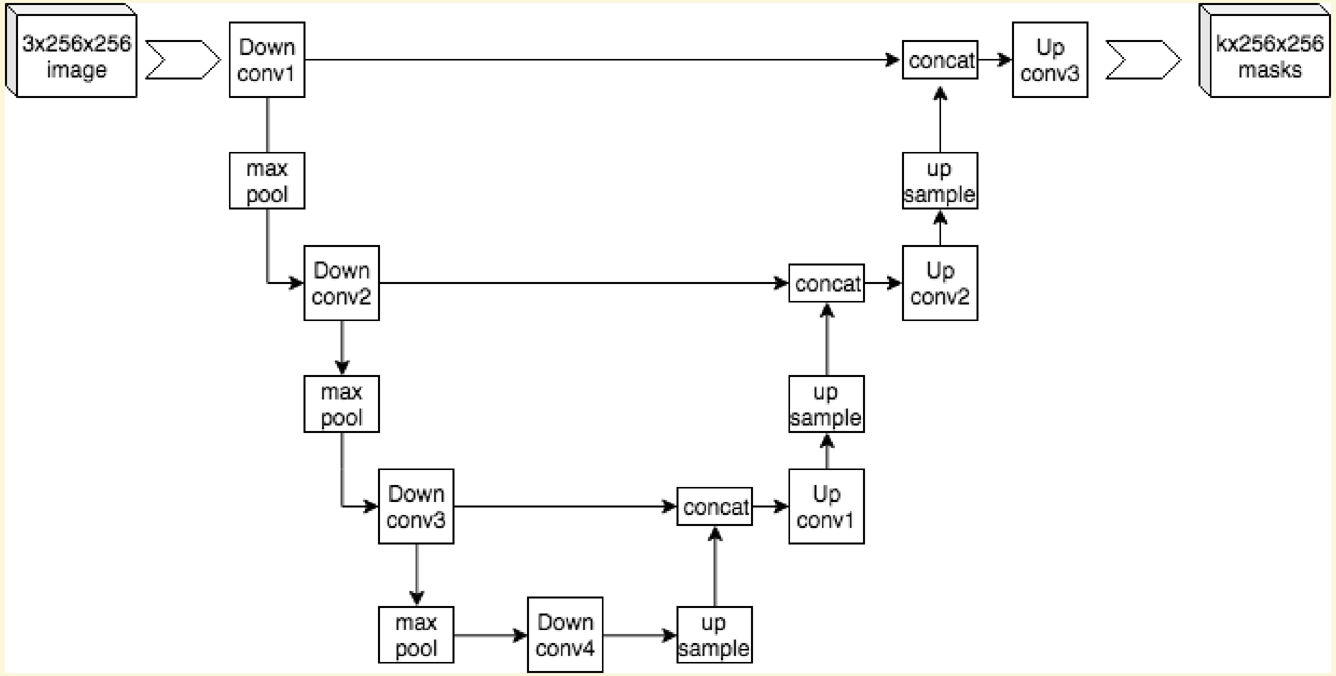


Fig. 1. UNet [1]

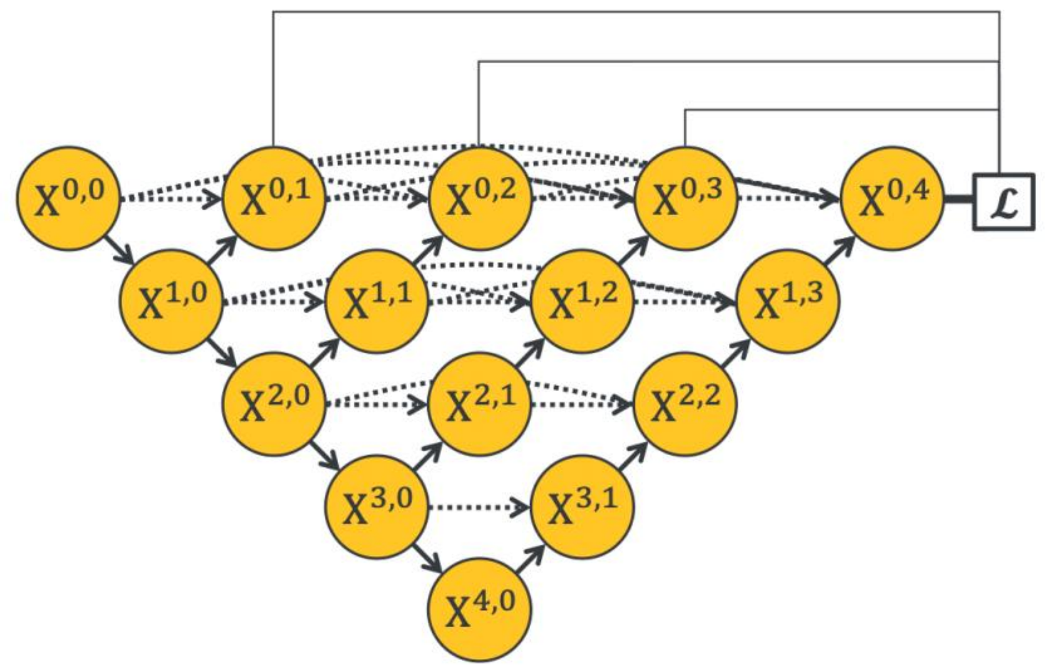
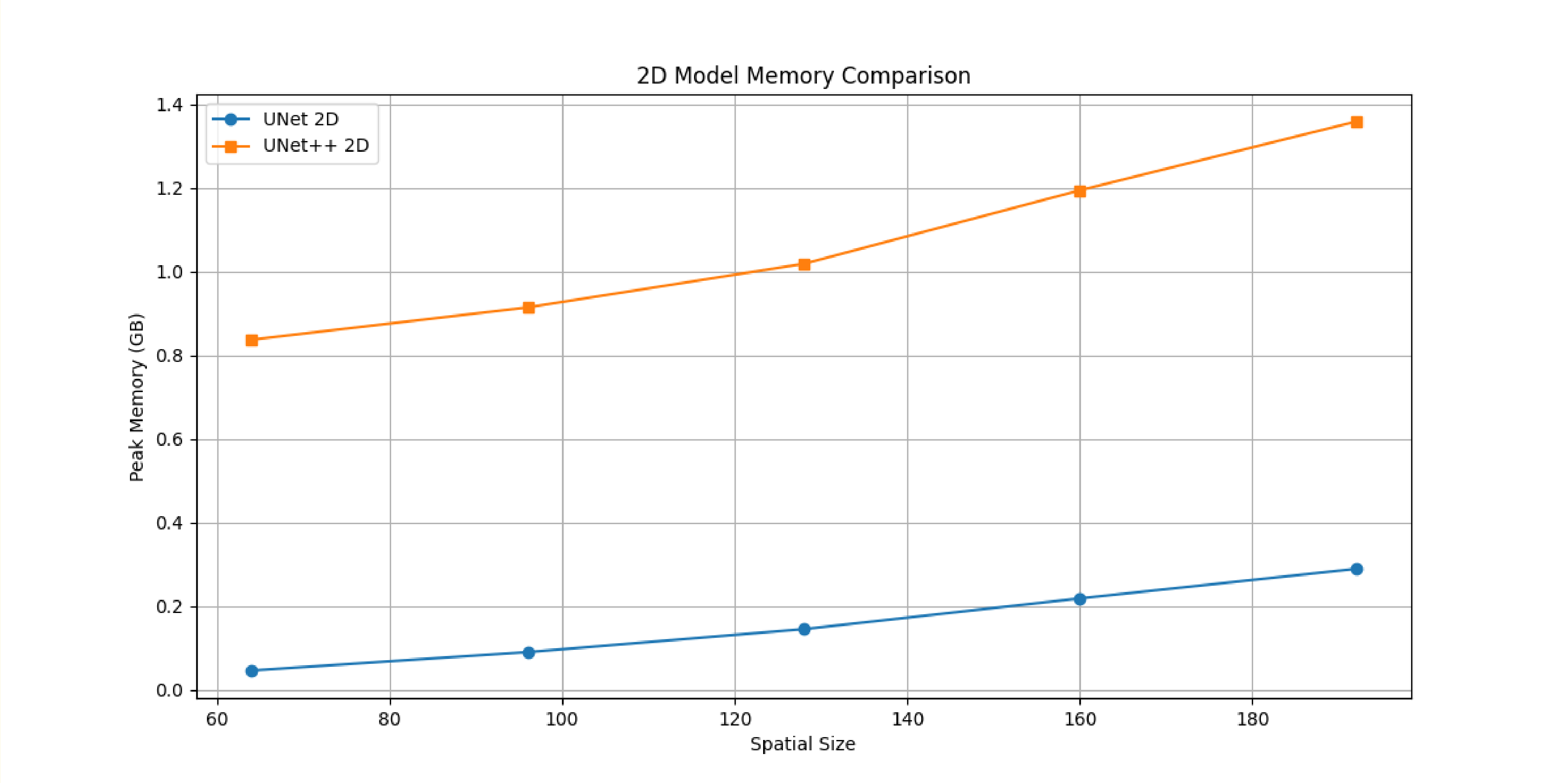


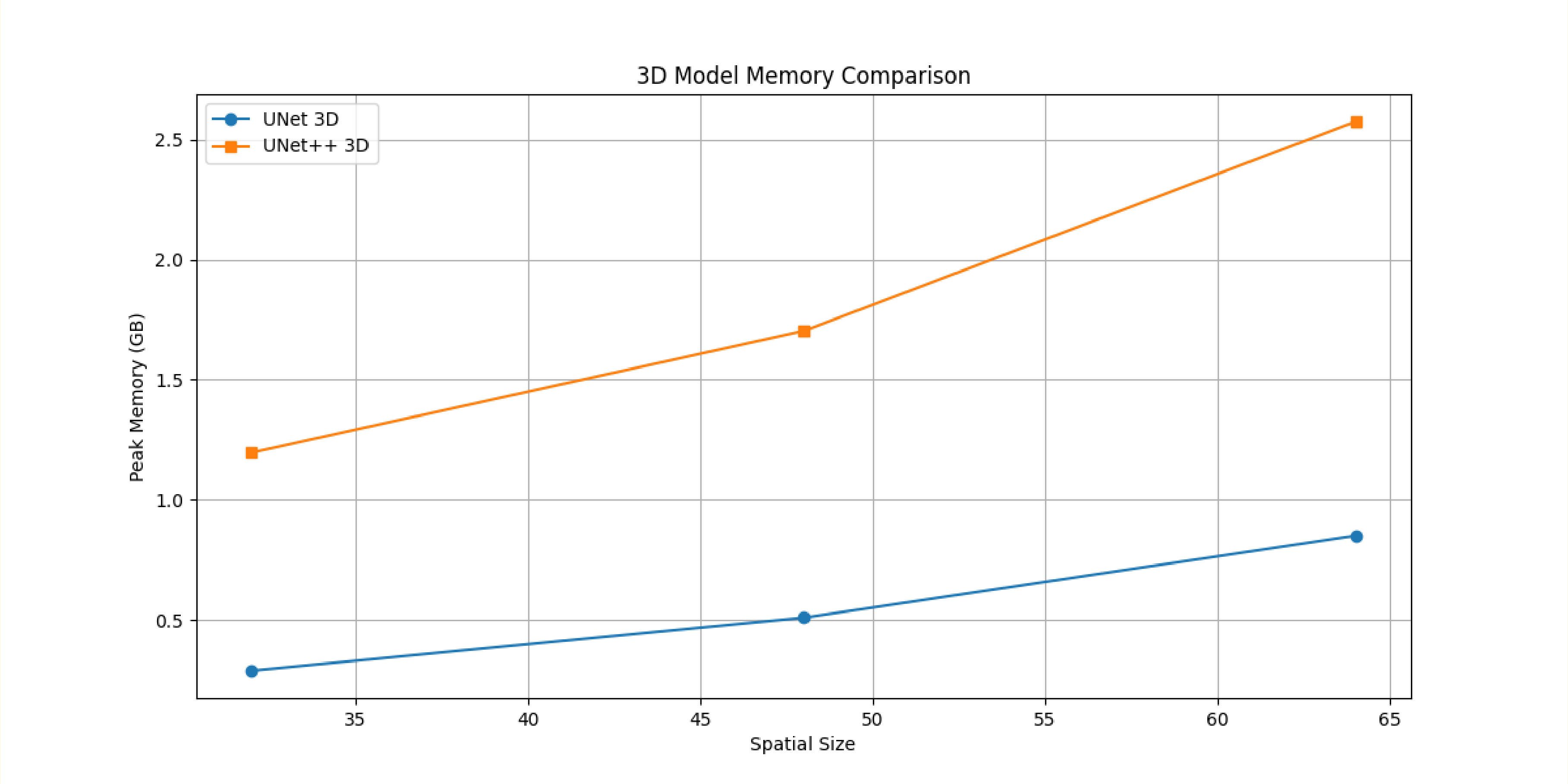
Fig. 2. UNet++ [2]

[1] https://en.wikipedia.org/wiki/U-Net, [2] https://arxiv.org/abs/1807.10165

RAPIDS
Models



RAPIDS
Models

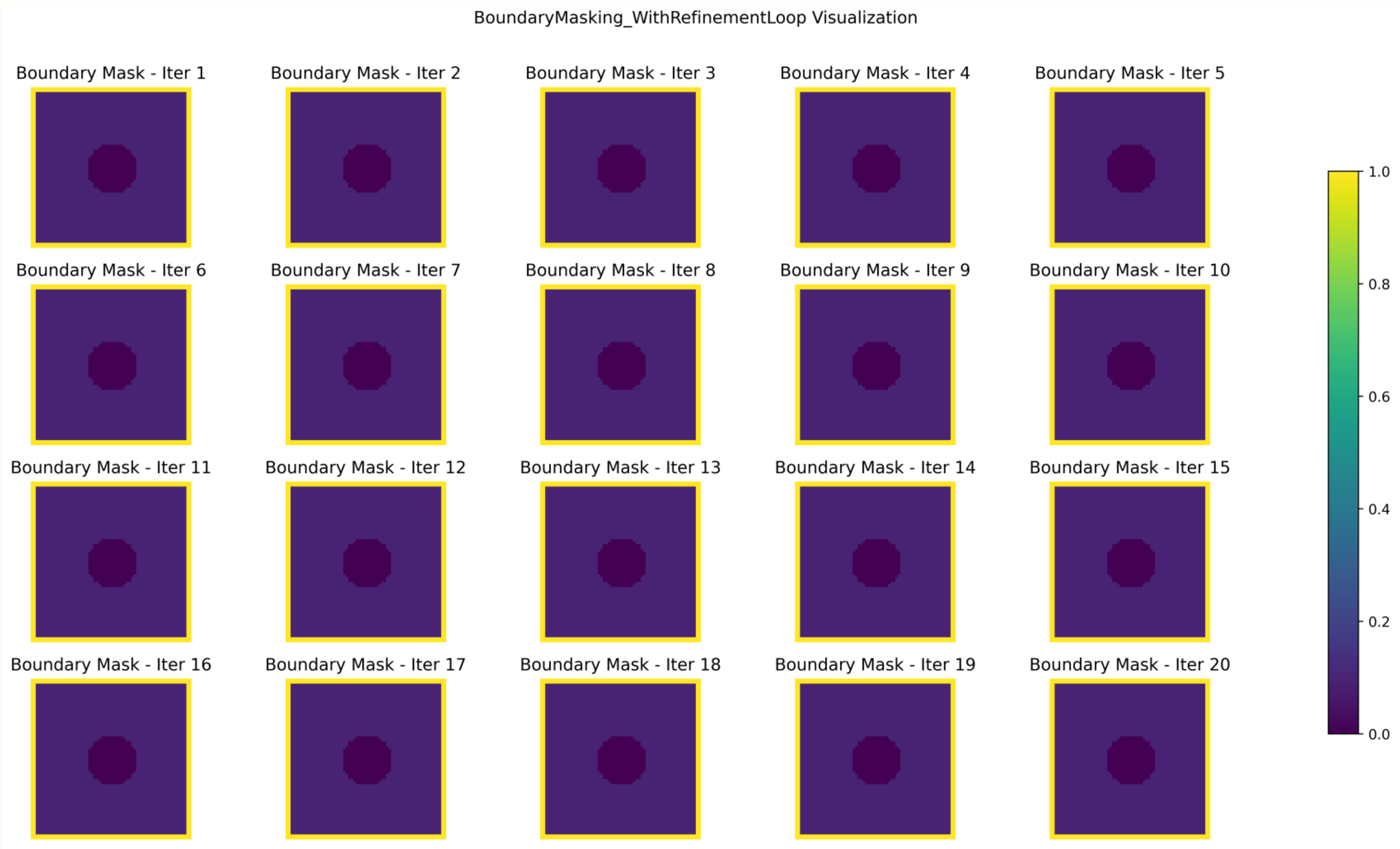


RAPIDS

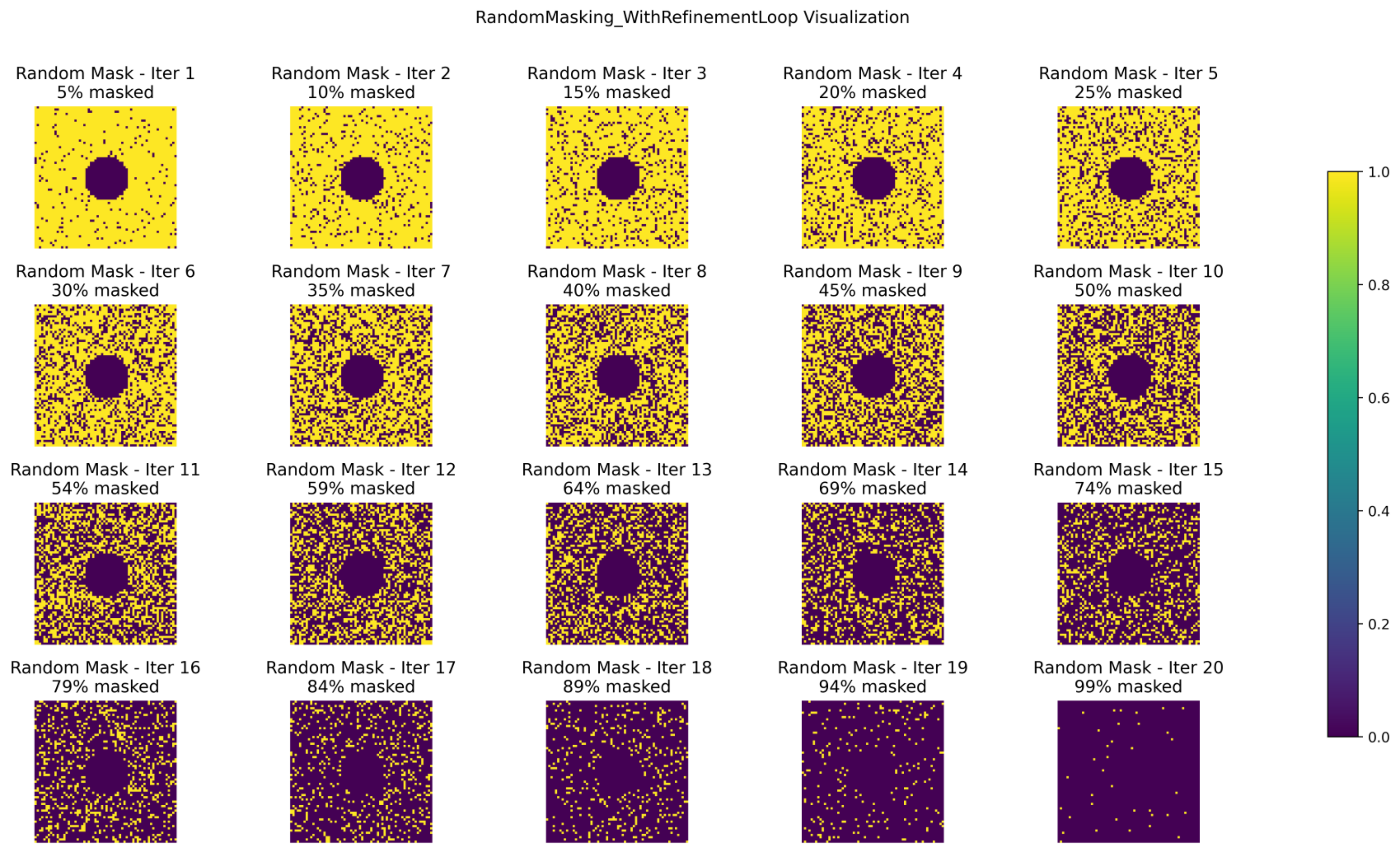
In-painting refinement with masking

- **Iterative Convergence**: Multiple refinement steps allow predictions to progressively converge toward physically viable solutions.
- **Boundary-Aware Generation**: Each refinement step can reinforce boundary constraints, teaching the model to generate interior regions that respect these physical constraints.
- **Error Recovery**: Multiple refinements during inference provide opportunity to correct initial prediction errors, critical for long-range temporal predictions.

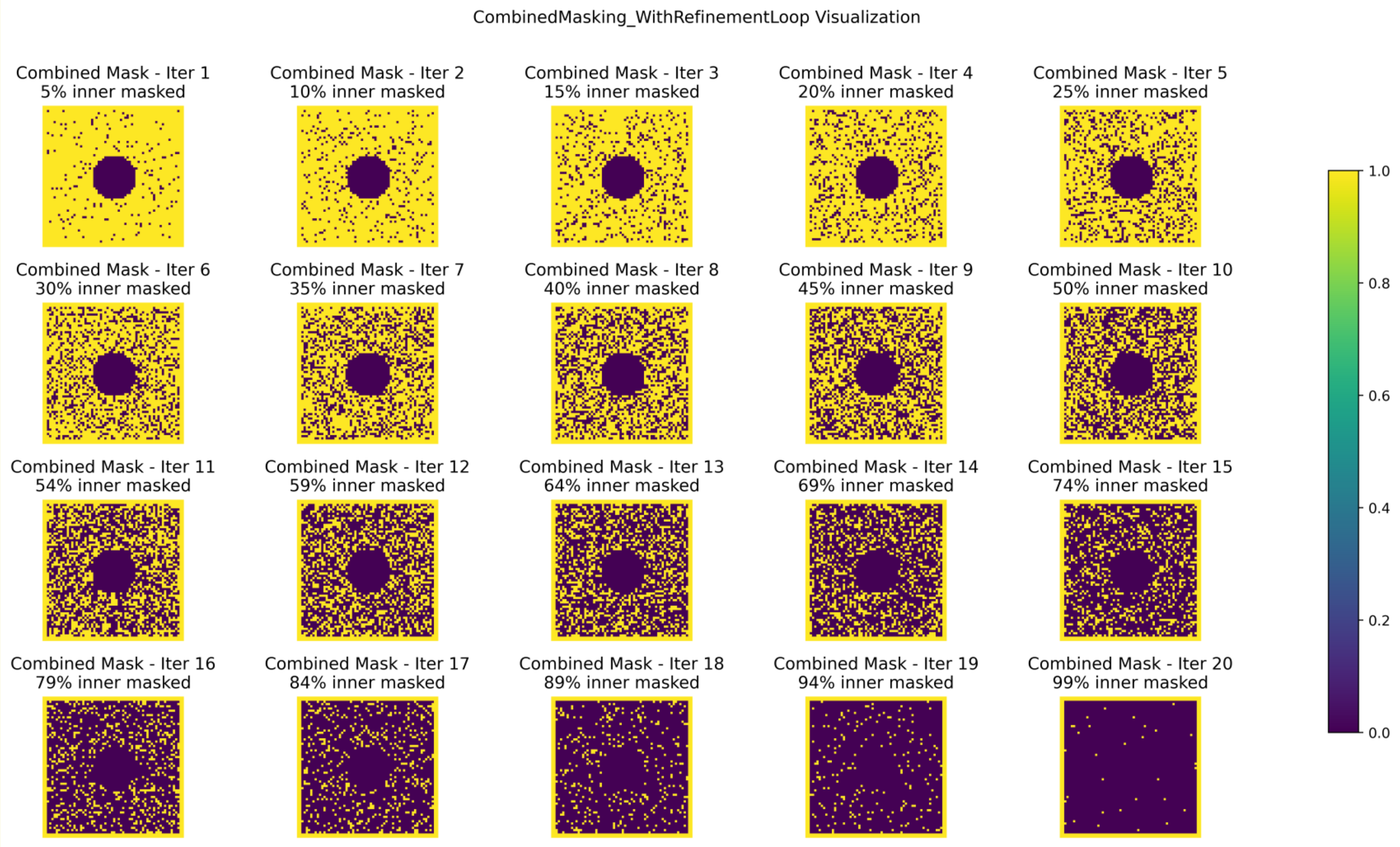
RAPIDS
Boundary Masking



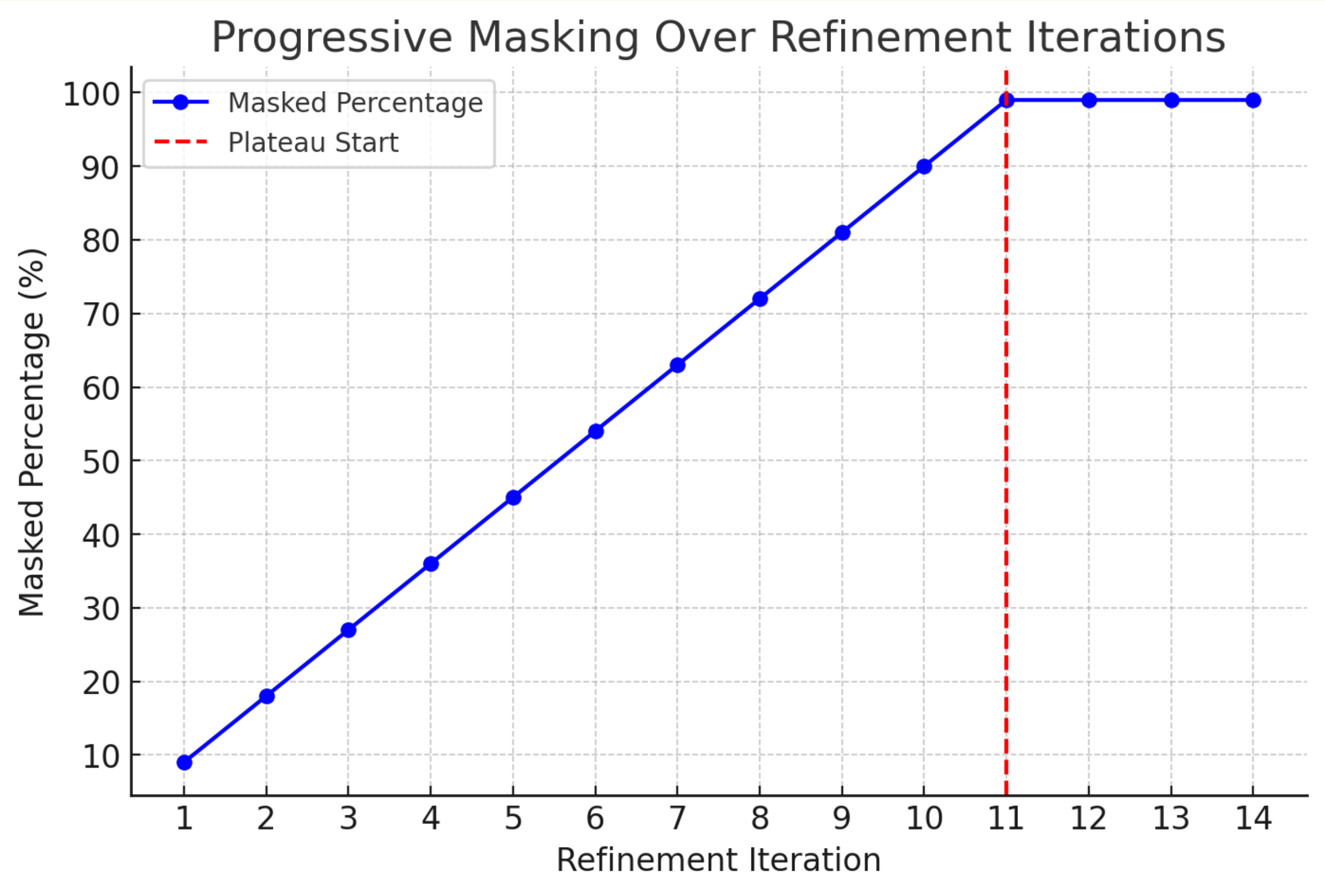
RAPIDS
Randomised Masking



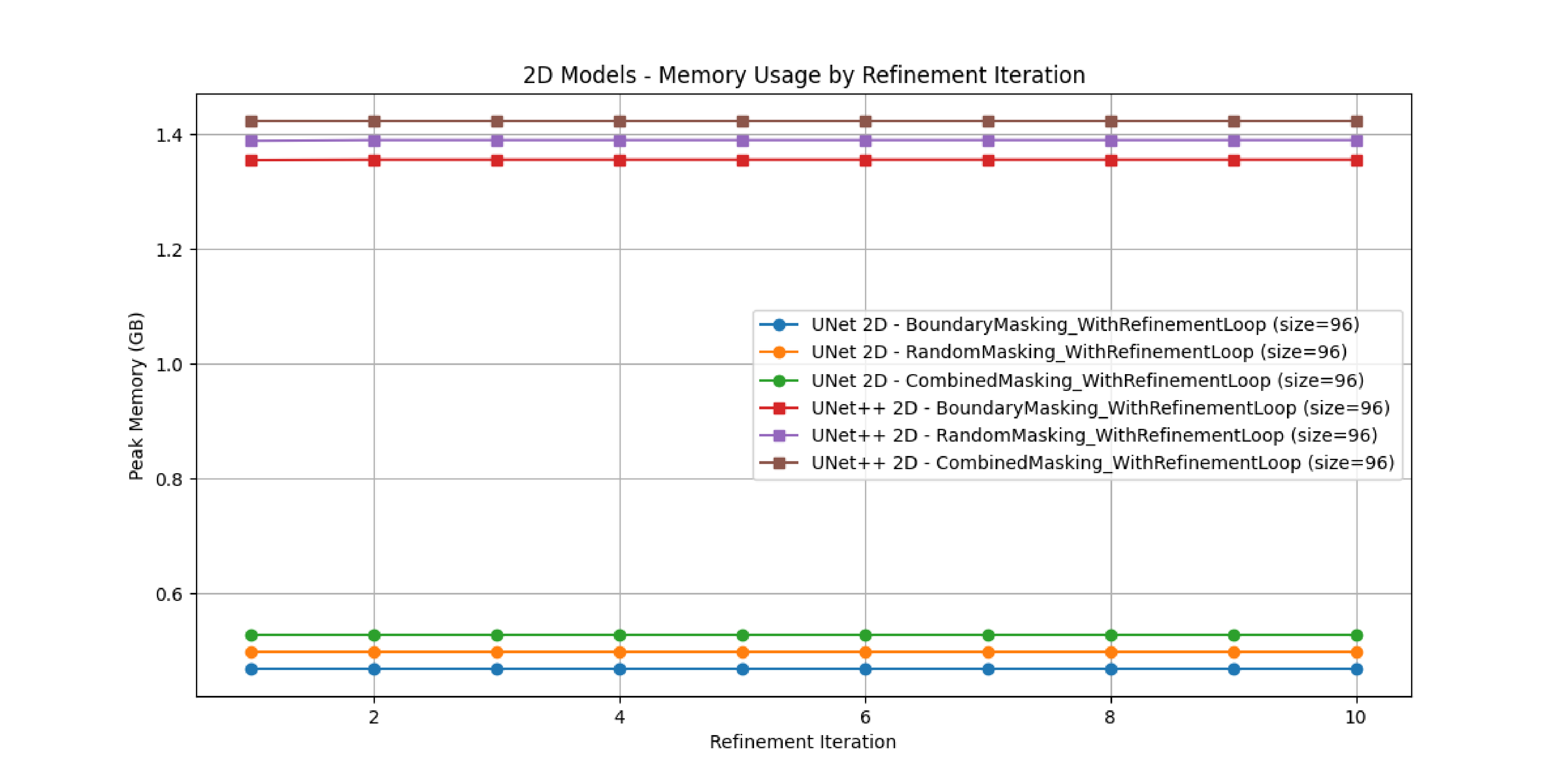
RAPIDS
Combined Masking



RAPIDS
Randomised and Combined Masking

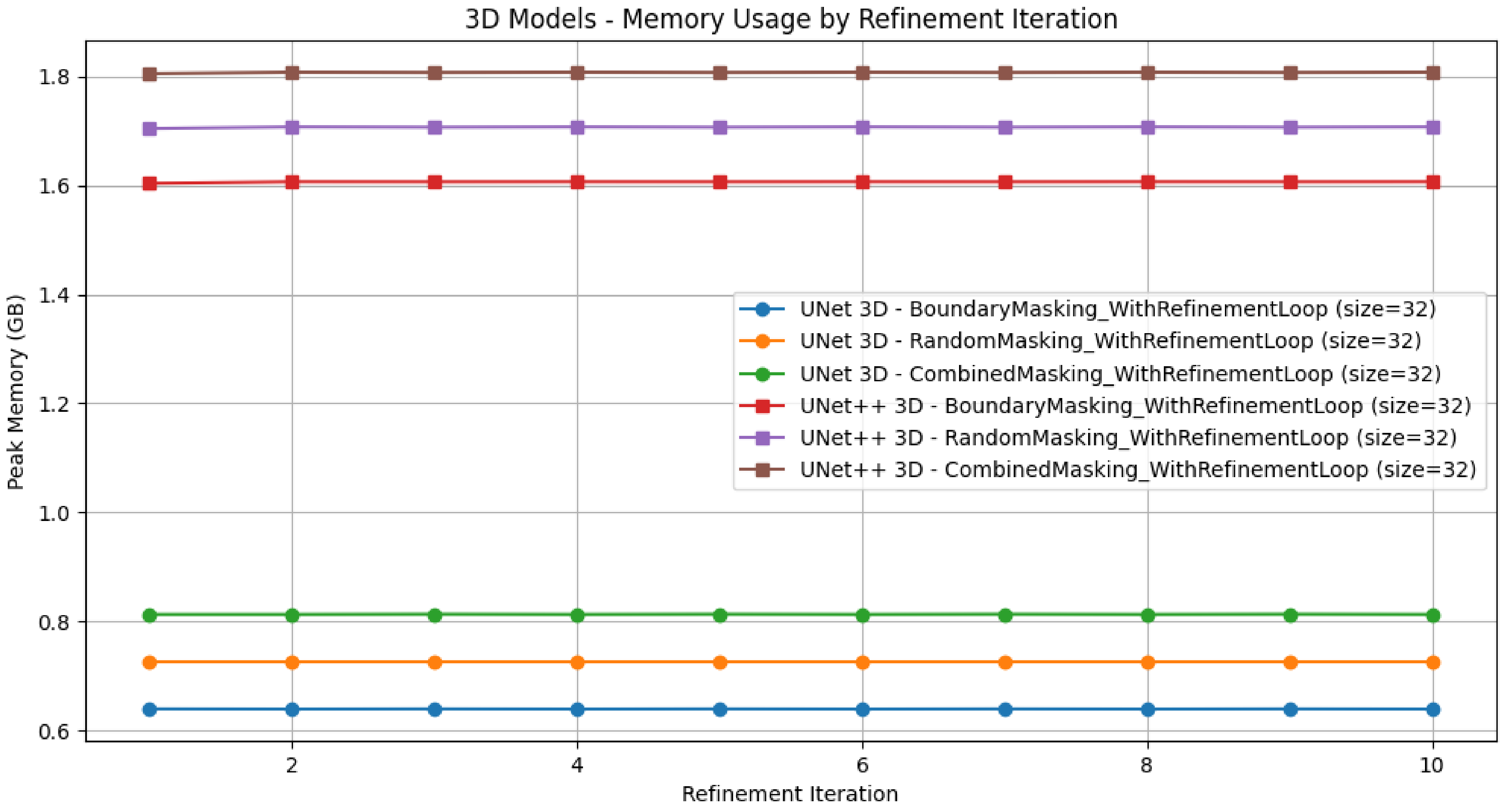


RAPIDS
Models with masking



We care about the end result– how is the loss calculated affects the memory usage

RAPIDS
Models with masking



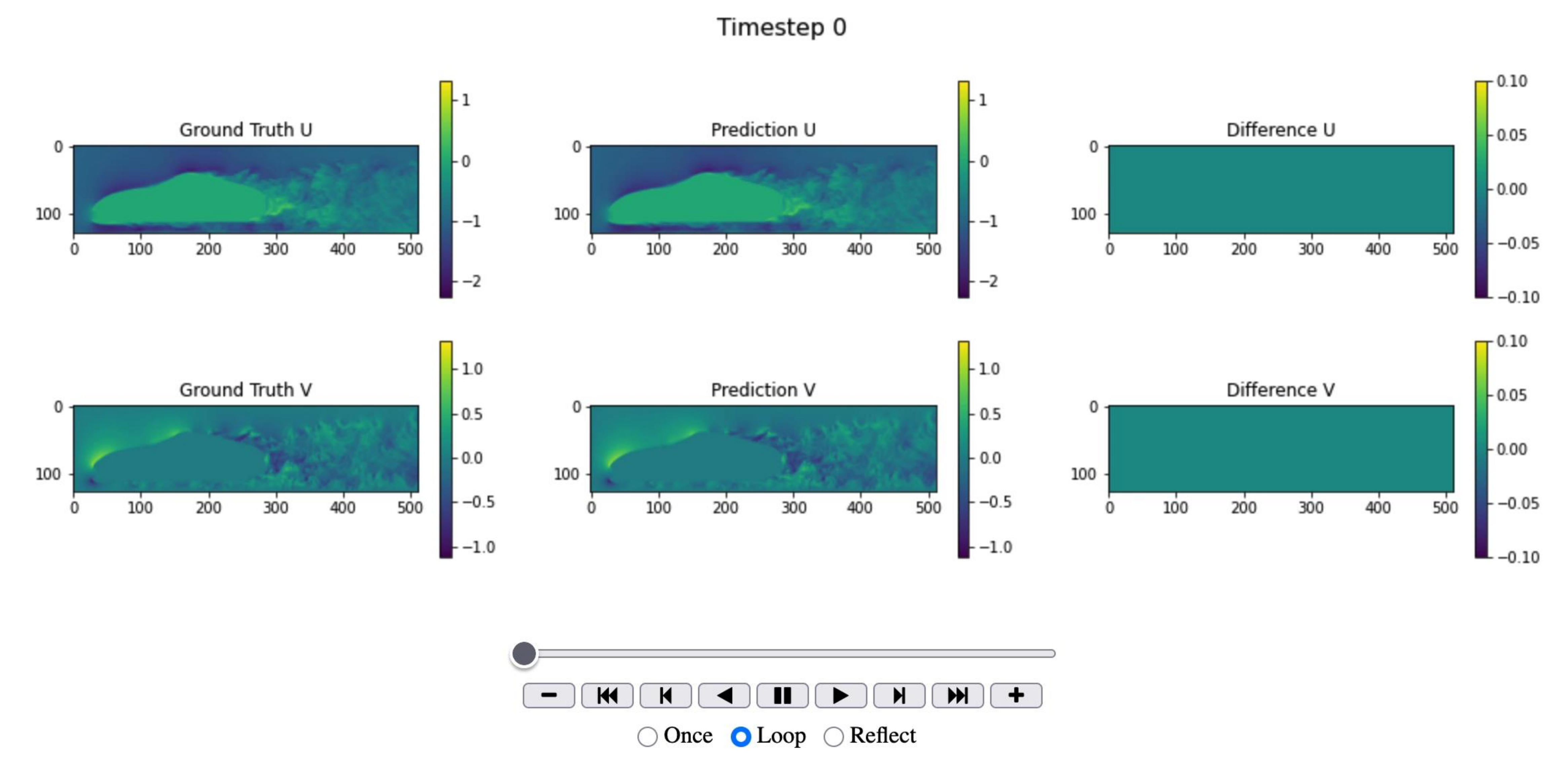
RAPIDS
Solid monitoring

- Is the solid moving? Building vs Aeroplane.
- Handle the temporal sequence of solid images same as the fluid image sequence.
- Computational overhead increases because you add one extra channel (or two), depending upon the architecture.

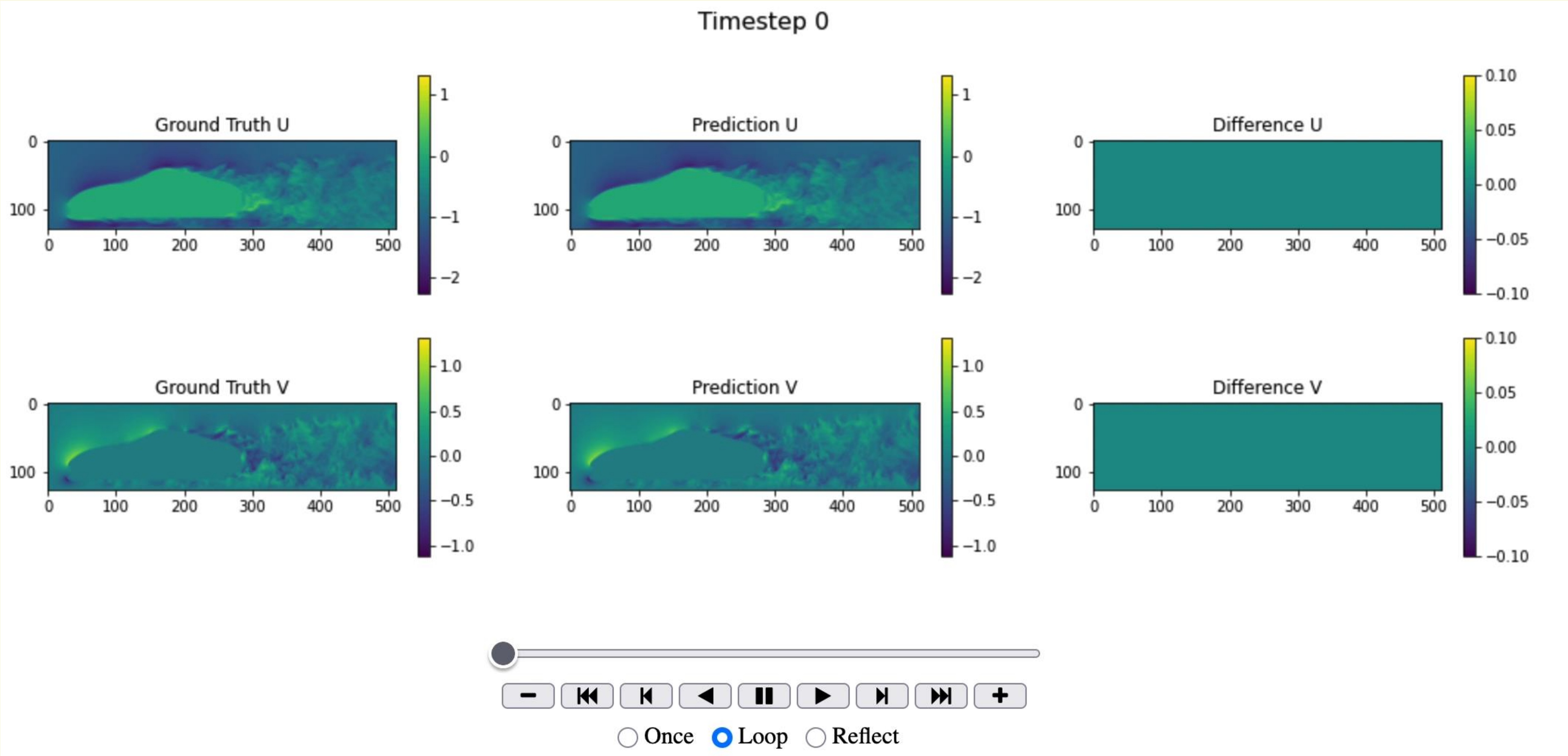
RAPIDS

Demonstrations

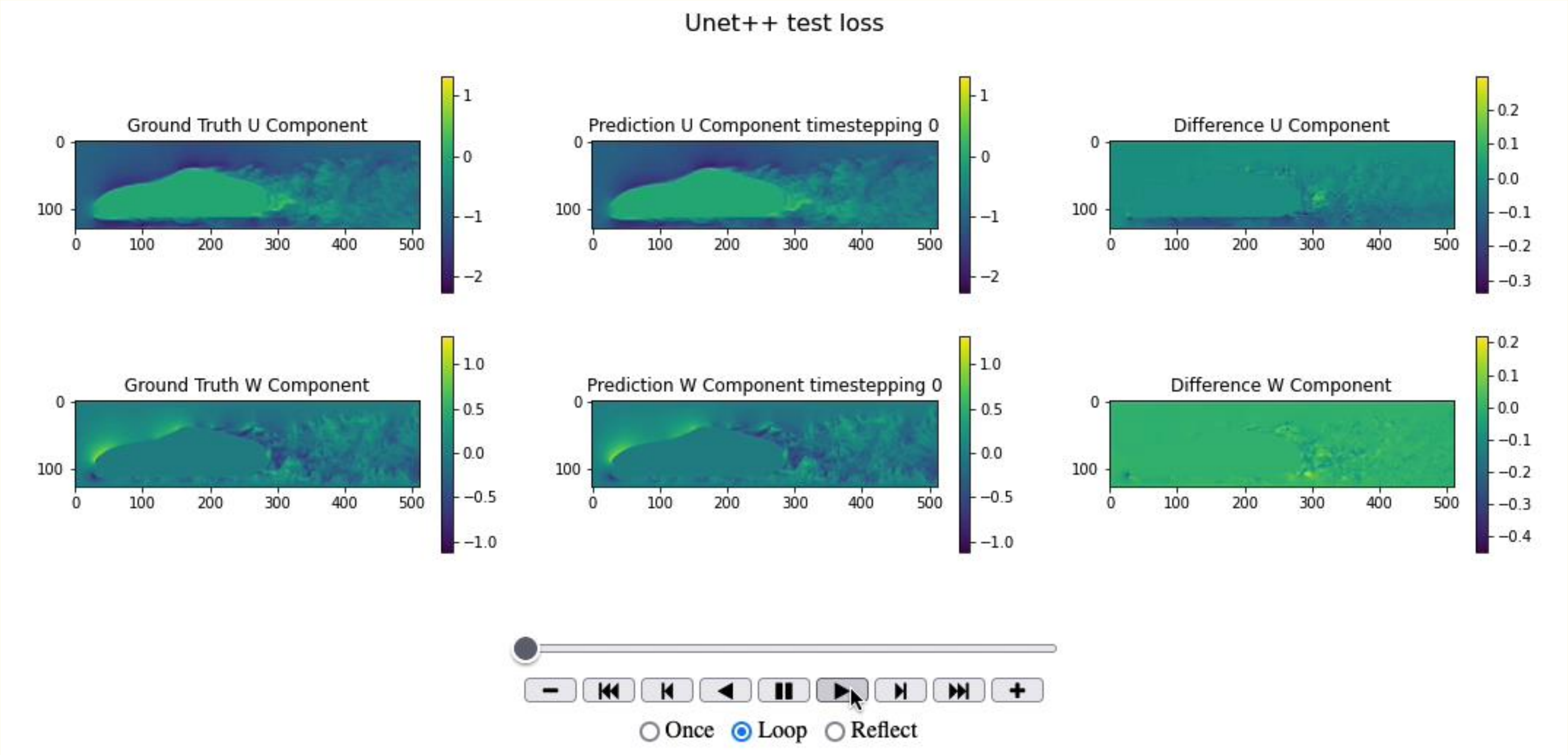
RAPIDS
Car – No solid considered



RAPIDS
Car – UNet, Combined Masking

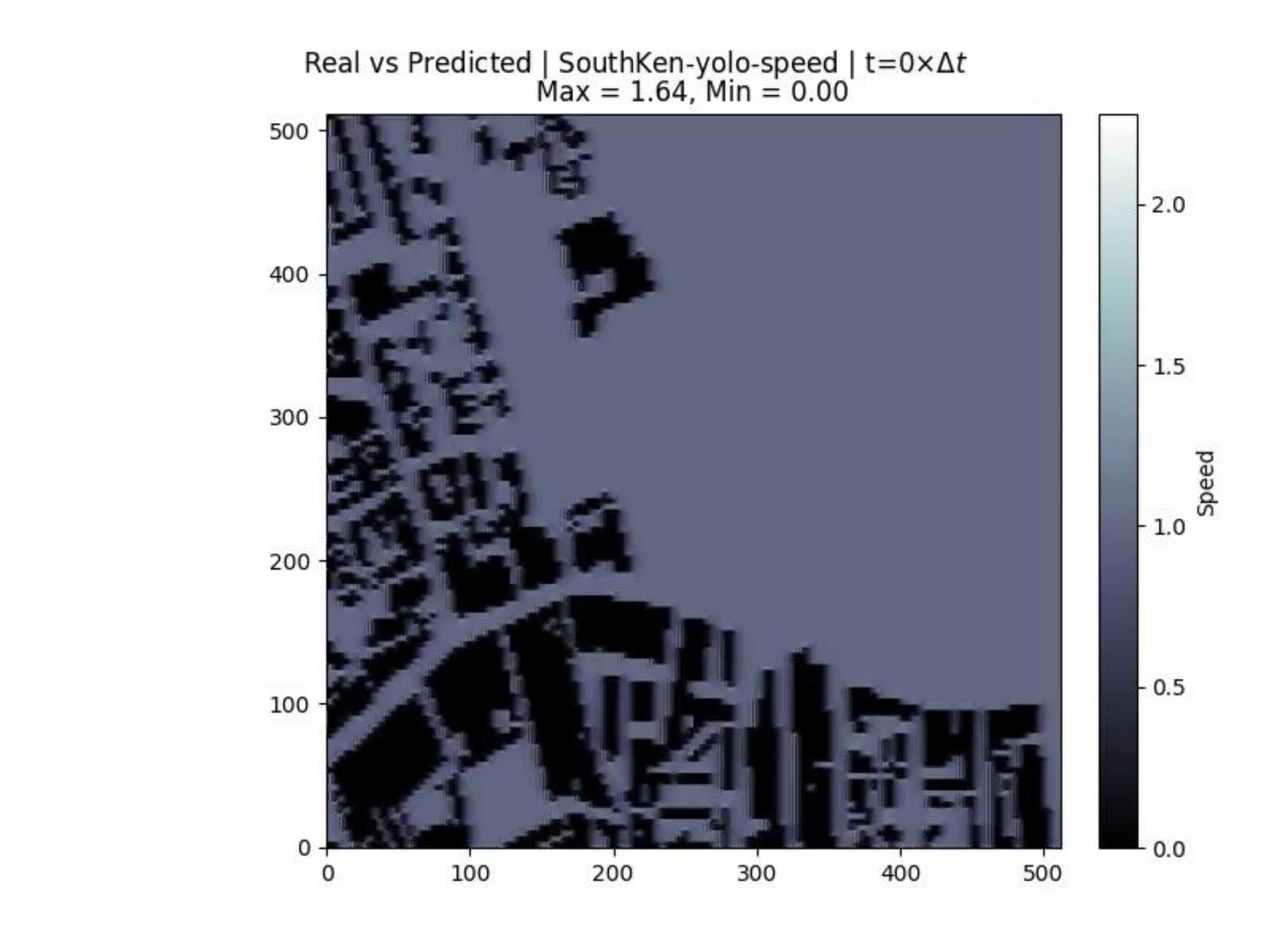


RAPIDS
Car – UNet++, Boundary Masking



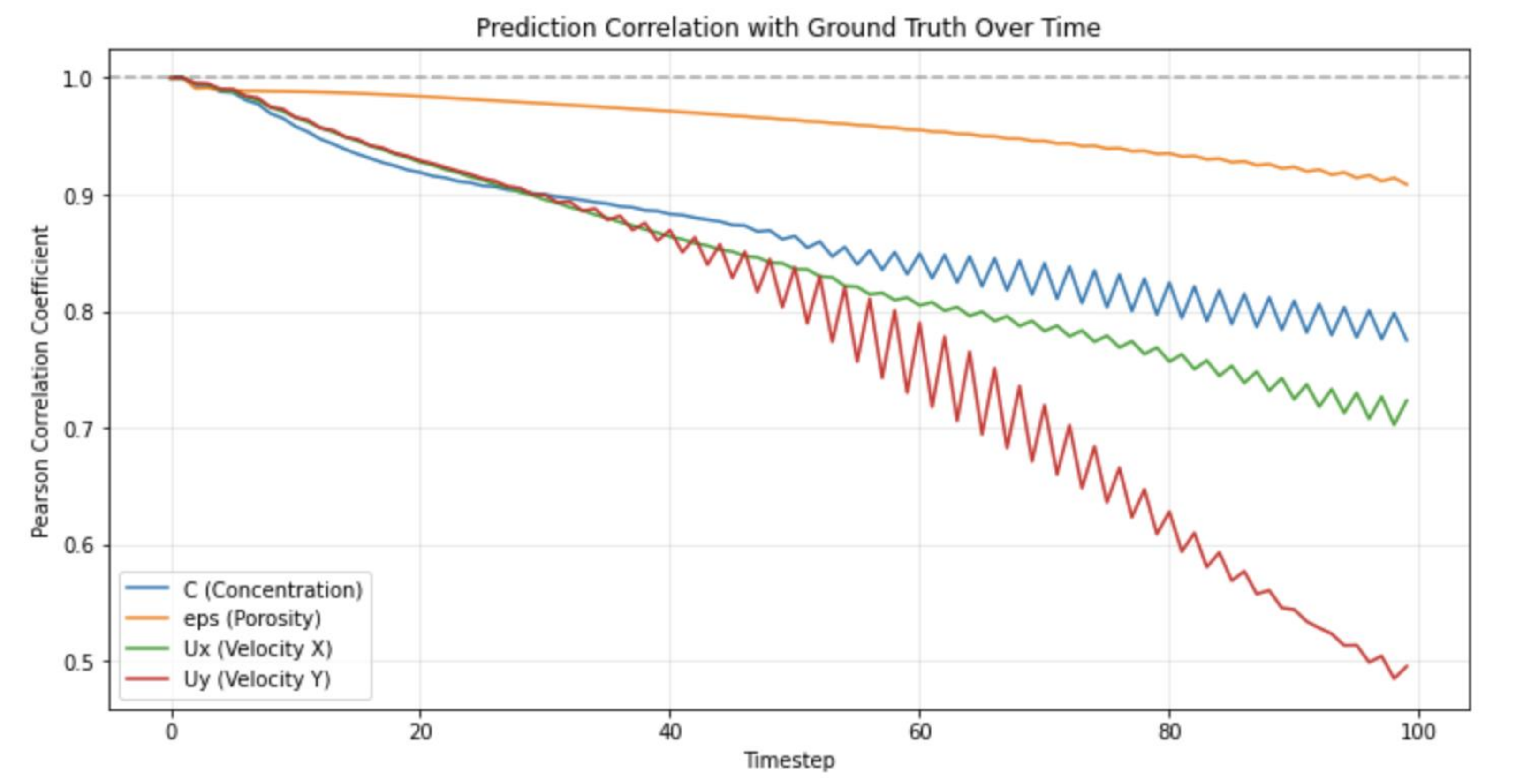
RAPIDS

Going further than a car– South Kensington



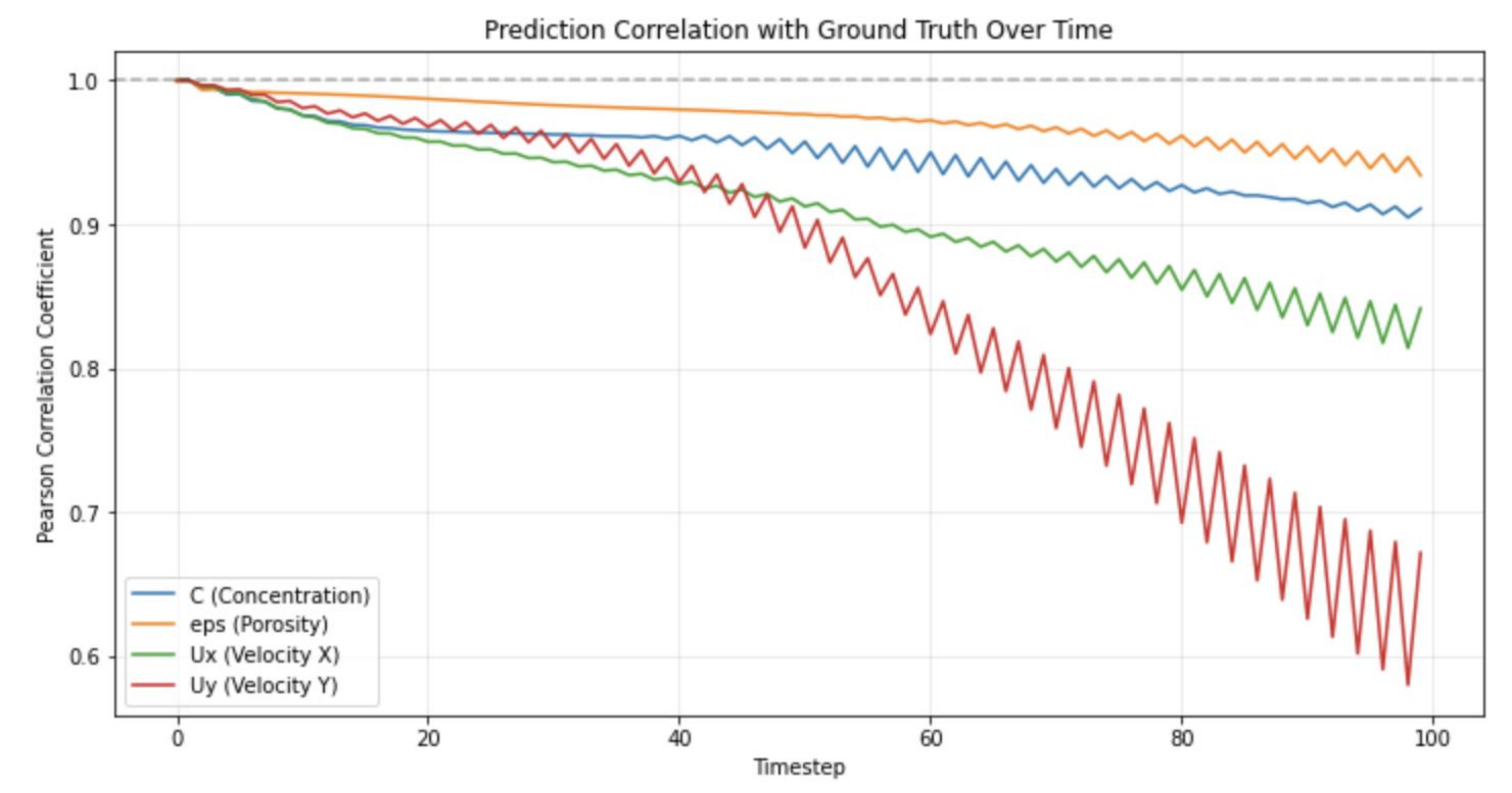
RAPIDS

CO2 sequestration– UNet++, implicit, BoundaryMask, StaticSolid, Ep 15



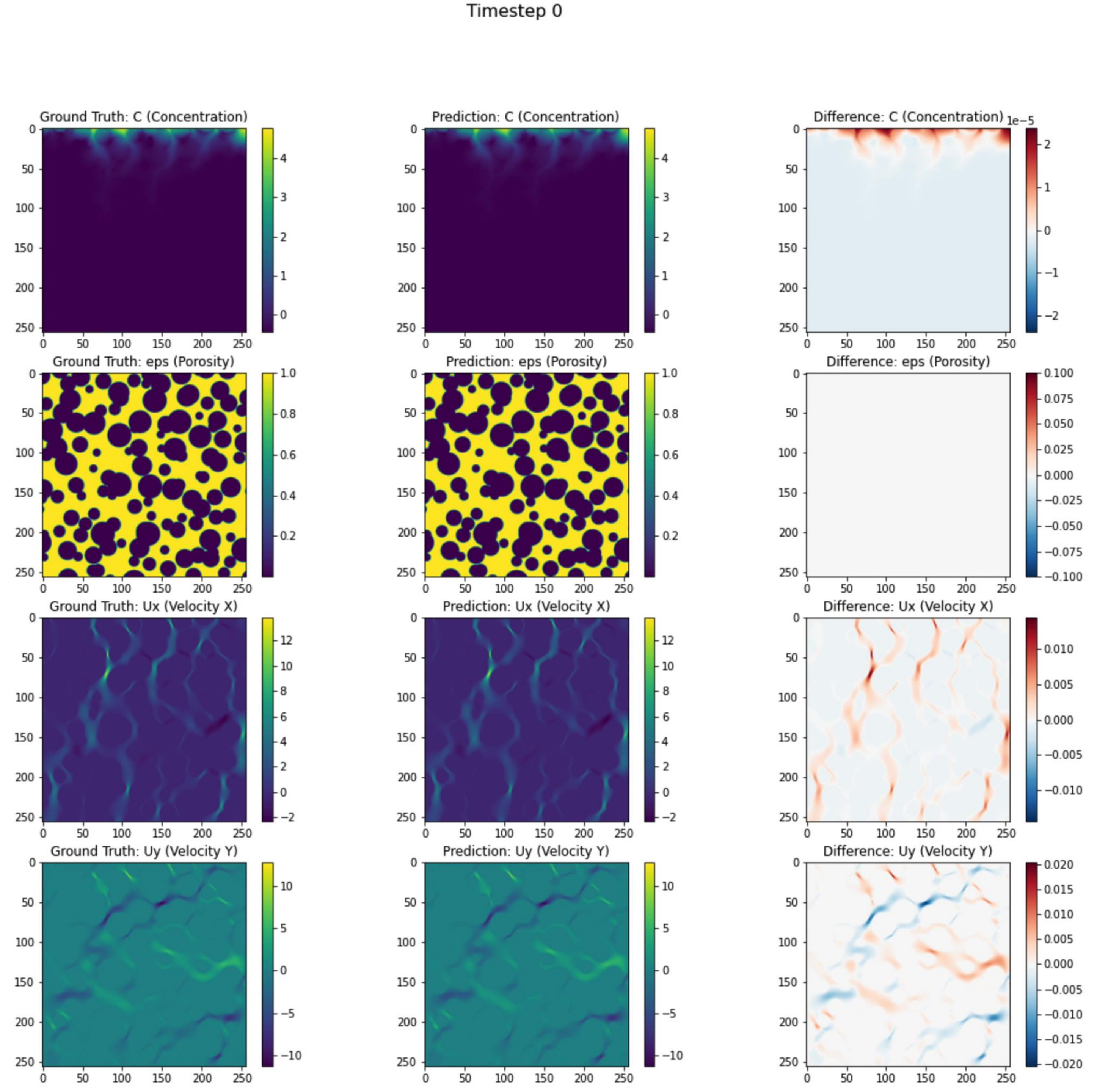
RAPIDS

CO_2 sequestration– UNet++, implicit, BoundaryMask, StaticSolid, Ep 20



RAPIDS

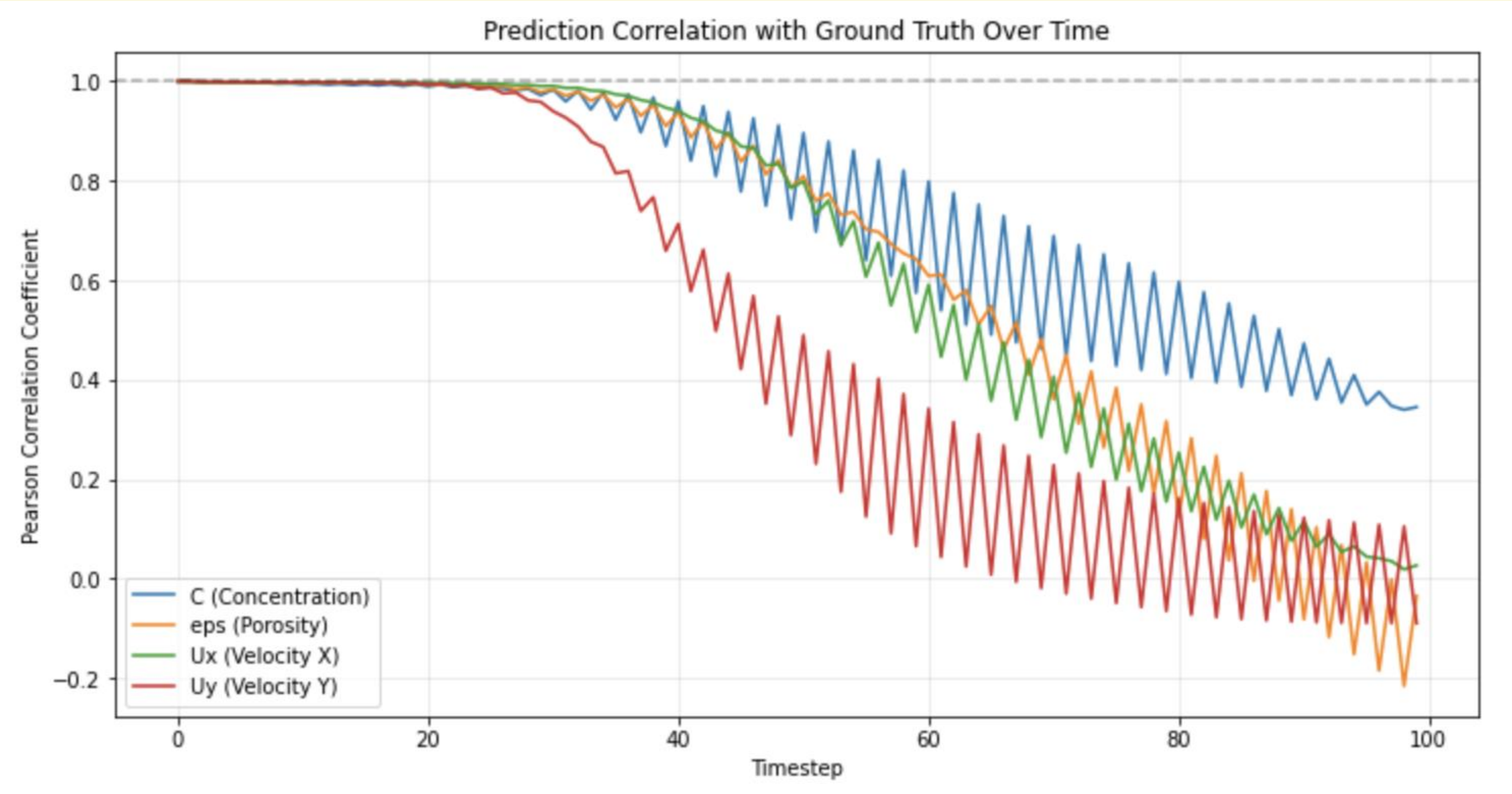
CO2 sequestratio



Ep 20

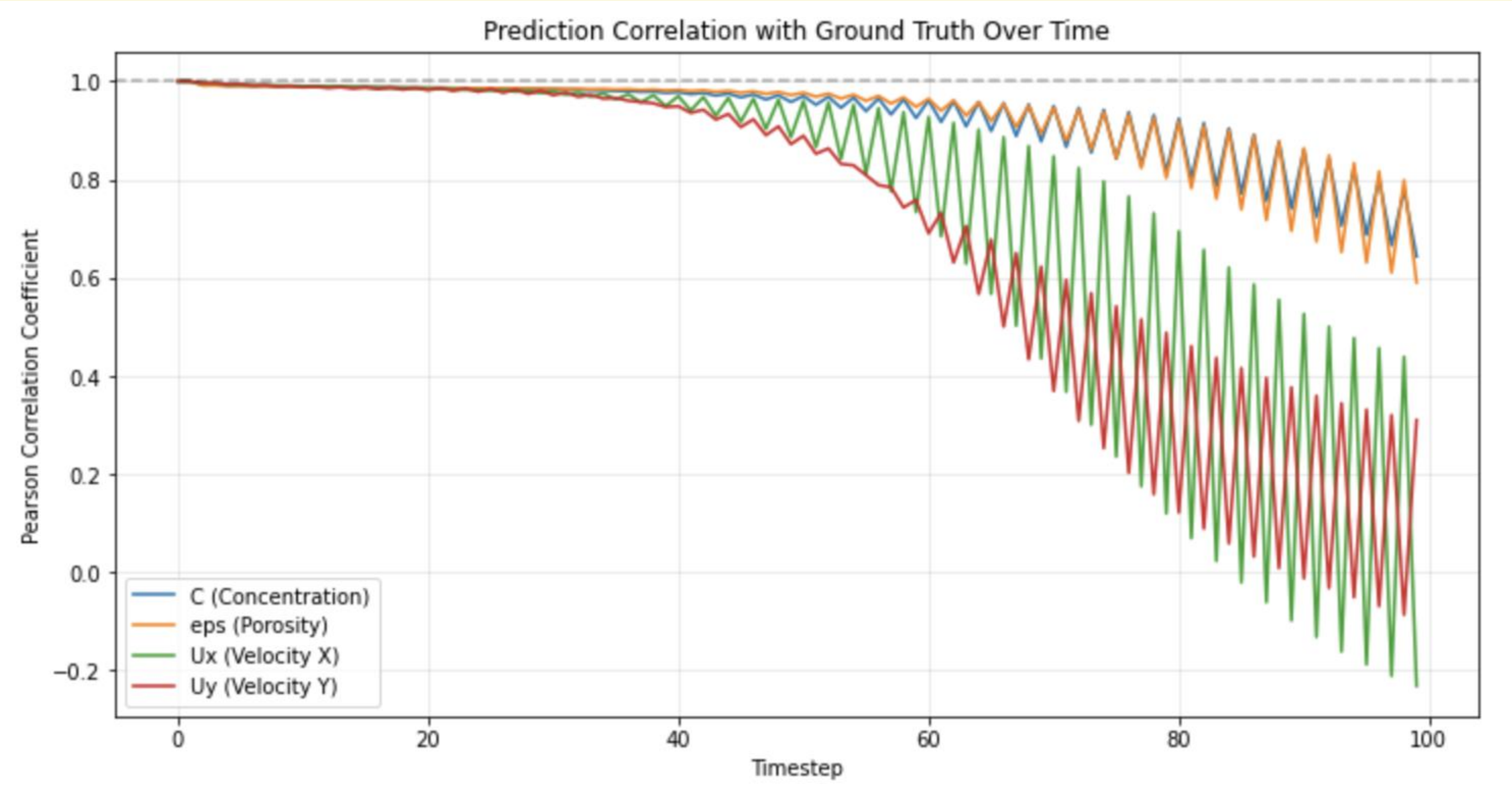
RAPIDS

CO_2 sequestration– UNet++, implicit, BoundaryMask, StaticSolid, Ep 100



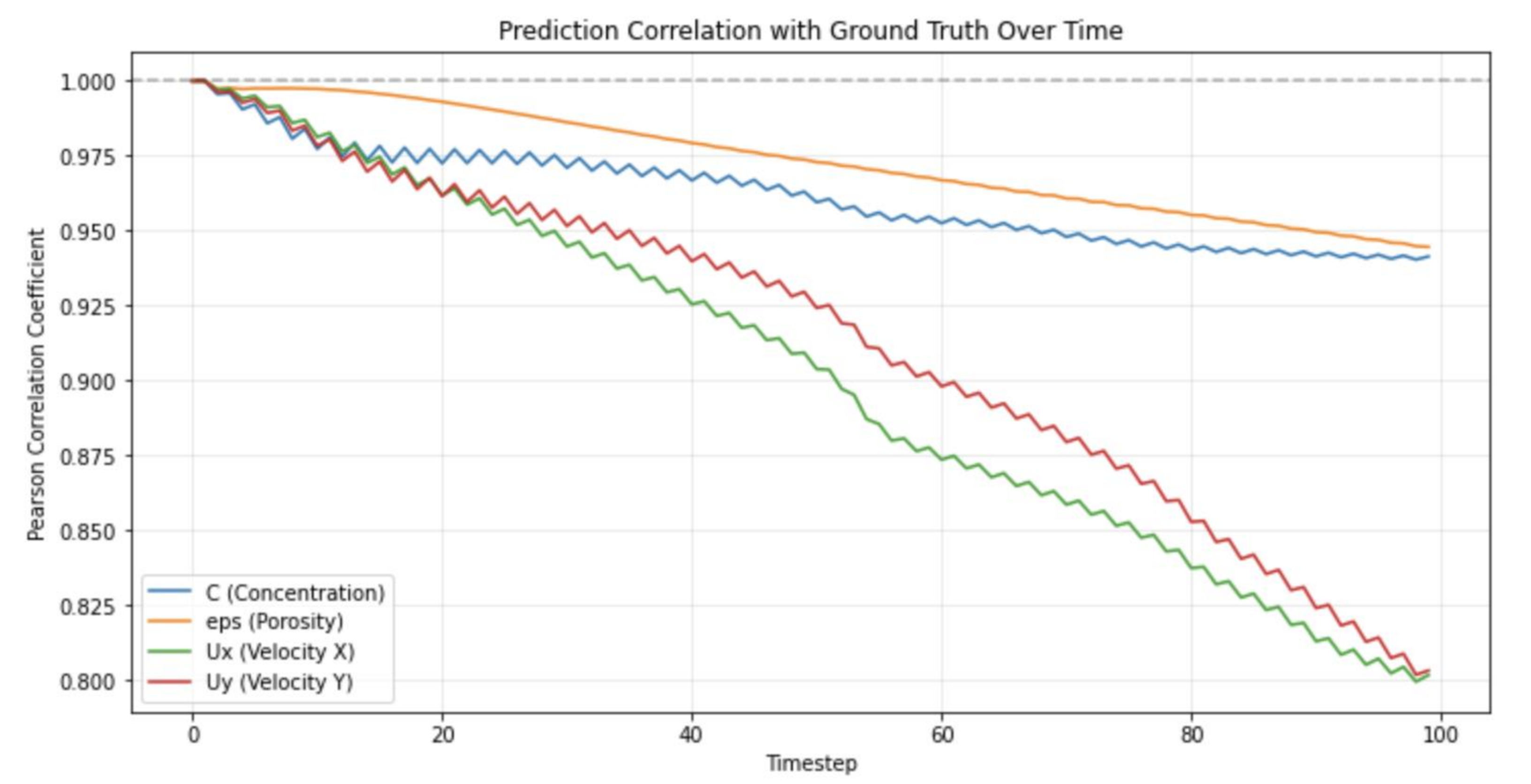
RAPIDS

CO2 sequestration– UNet++, implicit, BoundaryMask, MovingSolid, Ep 20



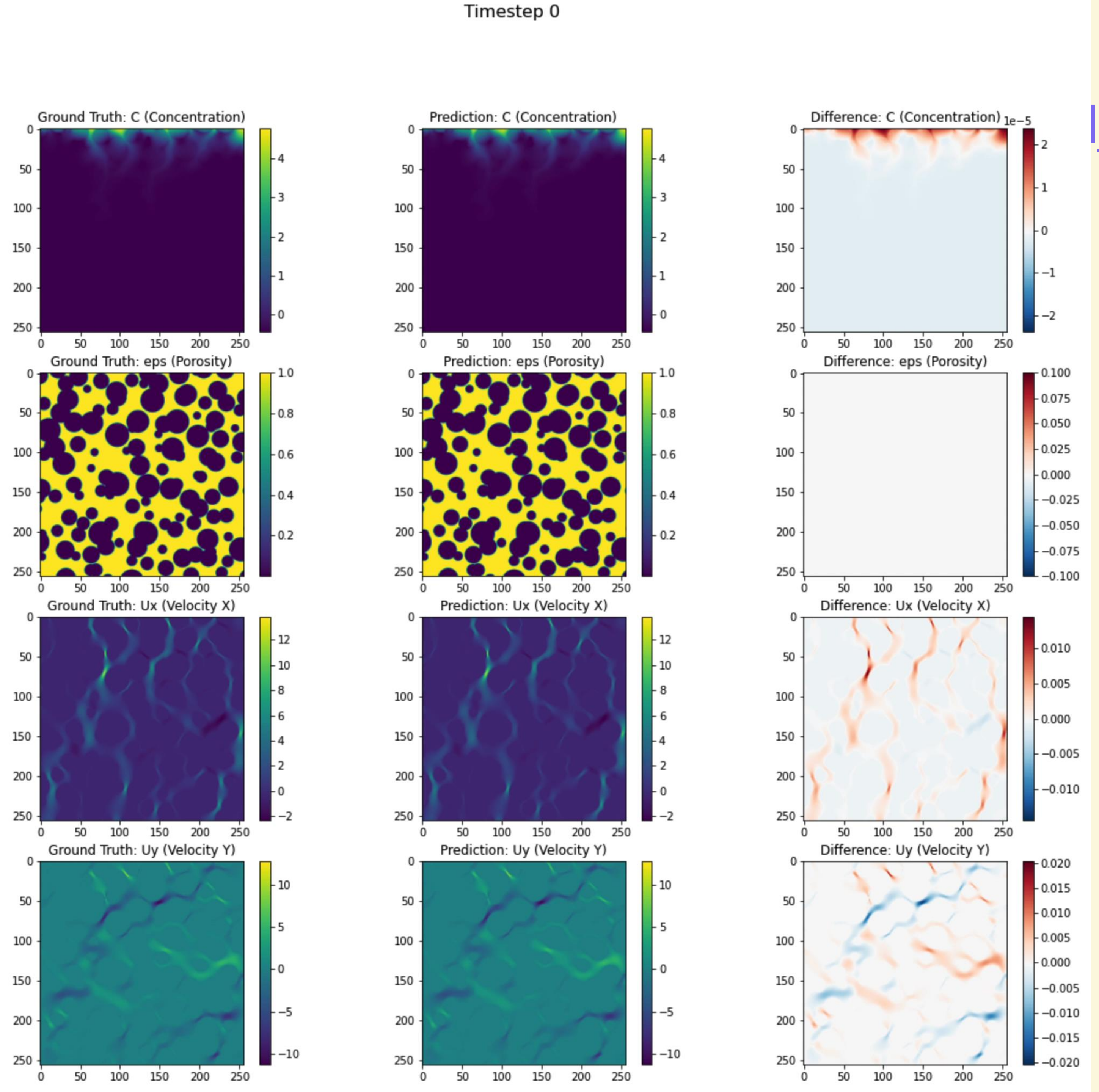
RAPIDS

CO2 sequestration– UNet++, explicit, BoundaryMask, MovingSolid, Ep 50



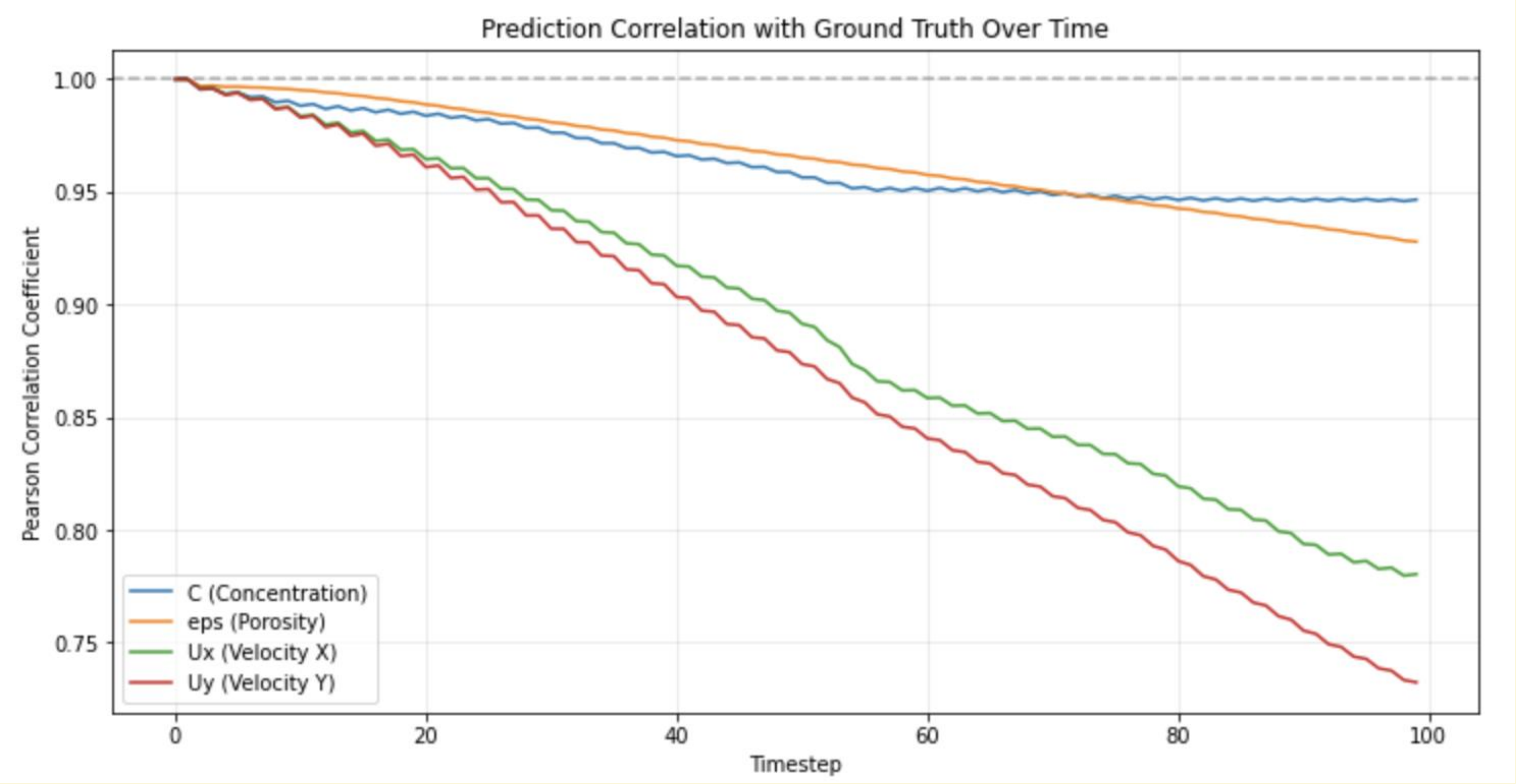
RAPIDS

CO2 sequestratio l, Ep 50



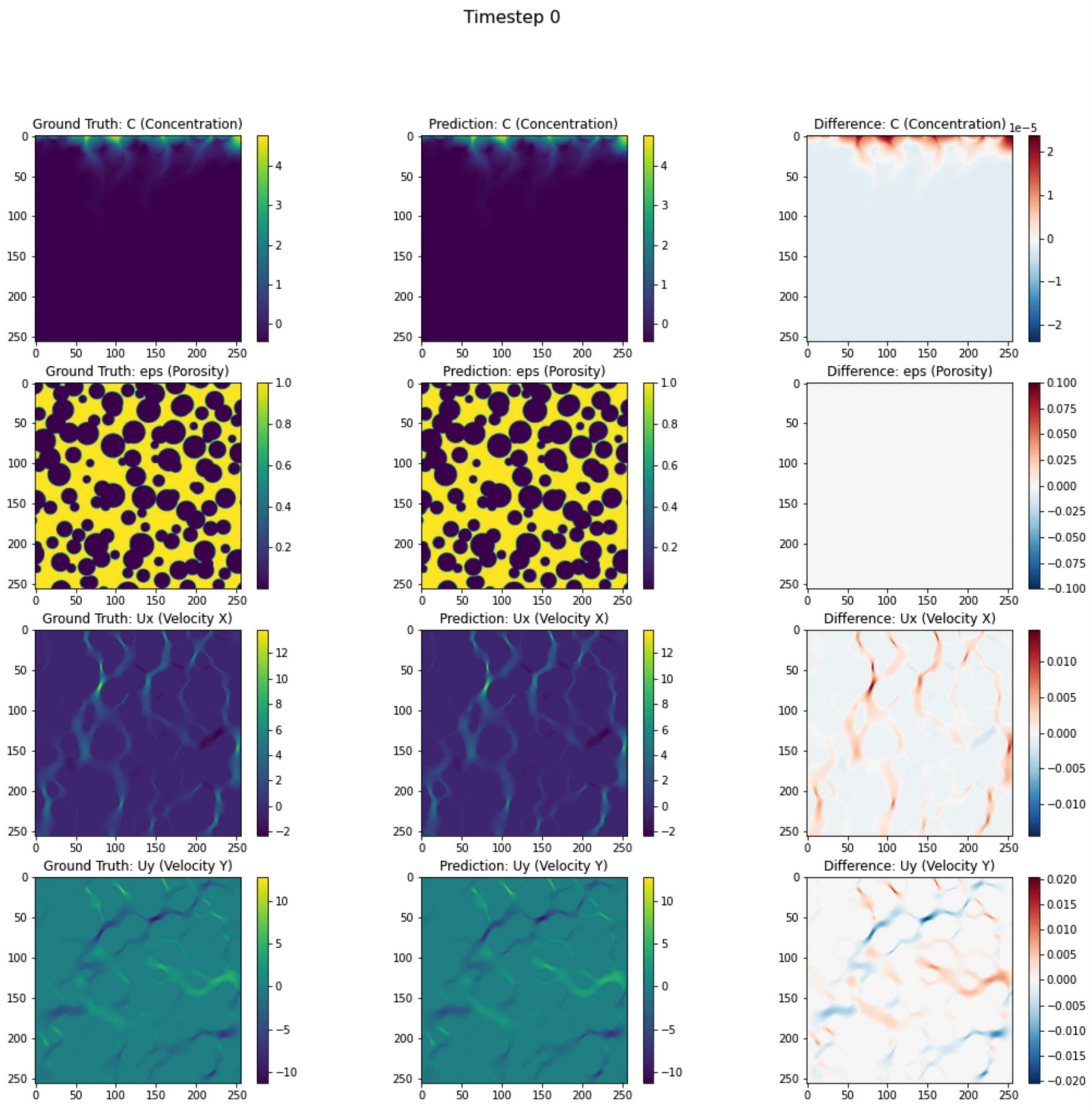
RAPIDS

CO2 sequestration– UNet++, explicit, BoundaryMask, StaticSolid, Ep 20



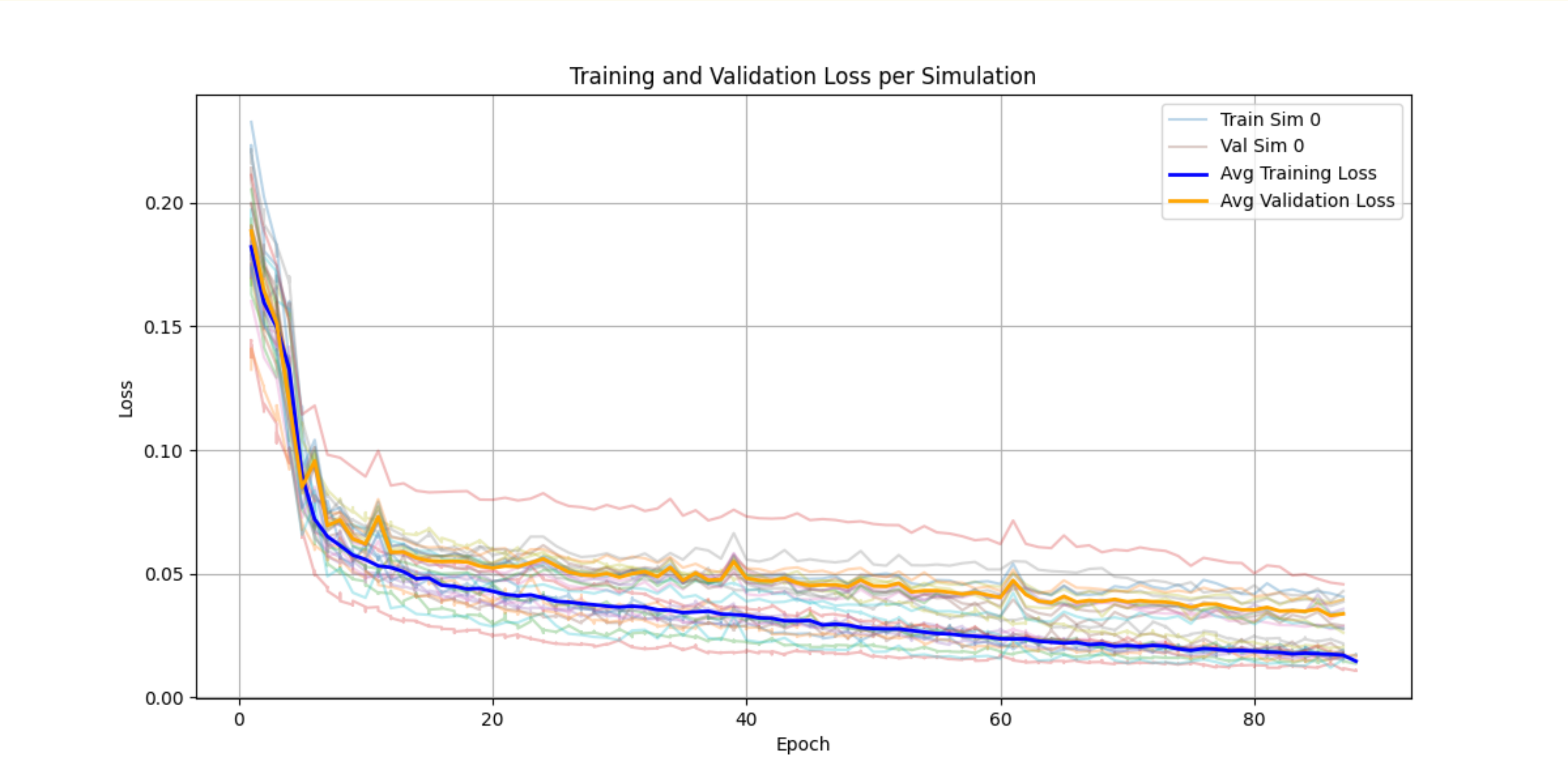
RAPIDS

CO2 sequestratio Ep 20



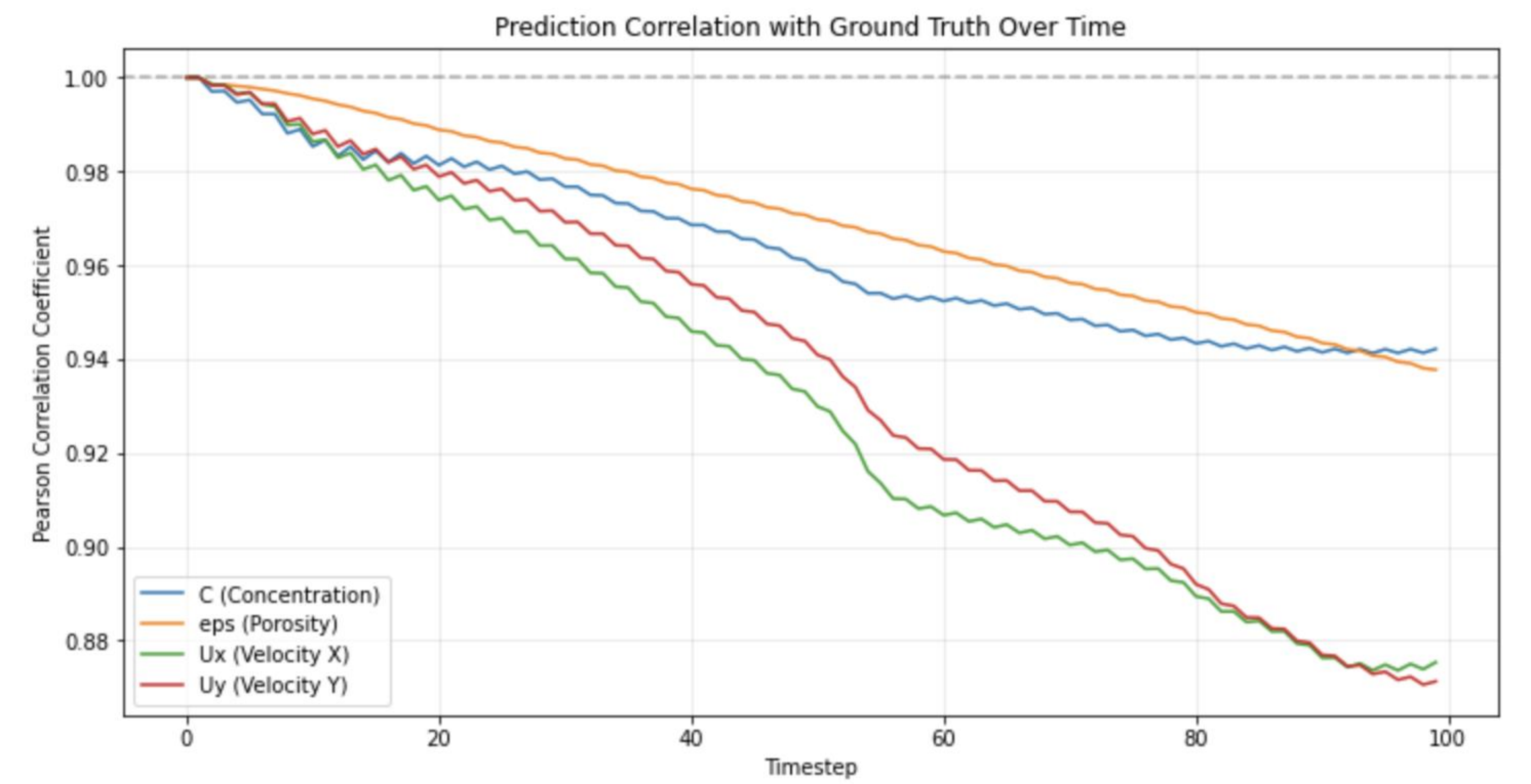
RAPIDS

CO2 sequestration– UNet++, explicit, BoundaryMask, StaticSolid



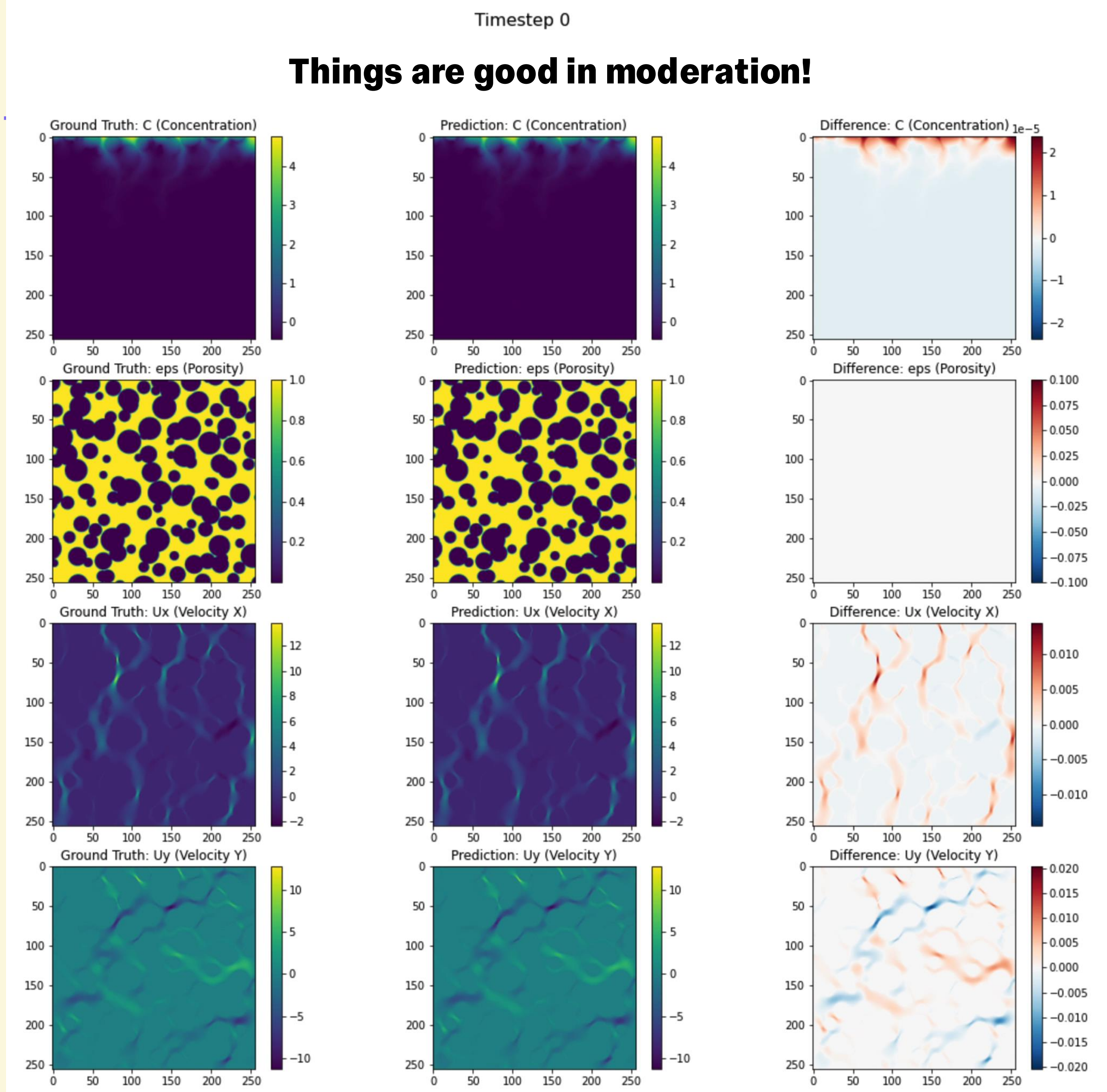
RAPIDS

CO_2 sequestration– UNet++, explicit, BoundaryMask, StaticSolid, Ep 90



RAPIDS

CO2 sequestra … d, Ep 90

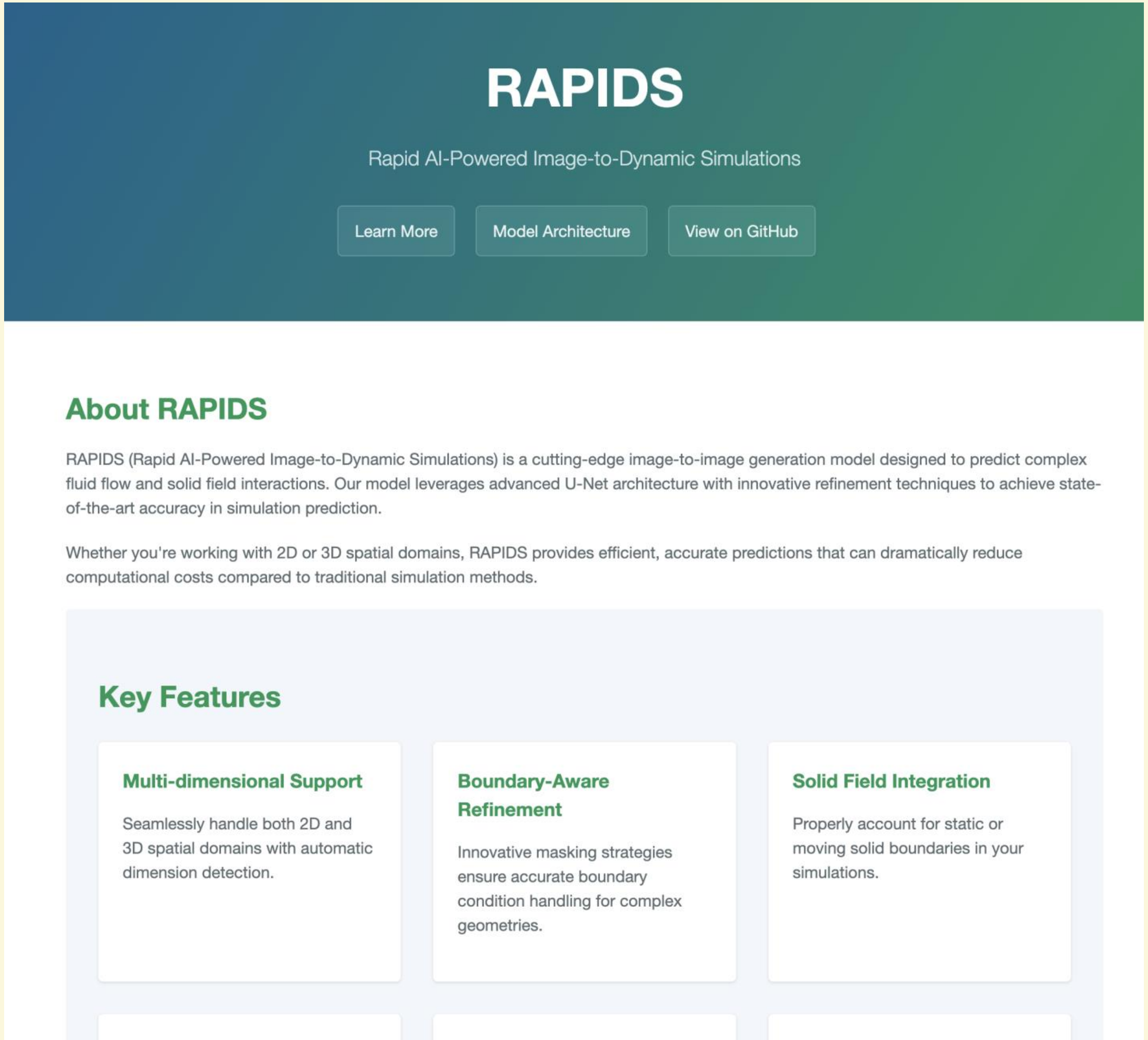


RAPIDS

Future

RAPIDS
Future work

- Stress testing of full package with unit tests.
- Website publication and design with diverse 2D and 3D demonstrations.
- Research write-up with new methodologies.
- Integration with NVIDIA's Modulus architecture.
- Integration of Diffusion architecture.
- Pre-trained foundational model dissemination.



https://imperialcollegelondon.github.io/RAPIDS

Thank you

For RAPIDS
By AMCG