Towards Agentic Foundation Models: Digital Twin Solutions for Virtual Sensing in Inaccessible Locations

Syed Bahauddin Alam
Appointed Committee Member, National Academies of Sciences, Engineering, and Medicine (NASEM)
Named National AI Leader, UIUC's White House National AI Action Plan (2025)
National Academy of Engineering (NAE) Frontier of Engineering 2026

MARTIANS: Machine Learning & ARTficial Intelligence for Advancing Nuclear Systems Lab
Nuclear, Plasma & Radiological Engineering
National Center for Supercomputing Applications
University of Illinois Urbana-Champaign

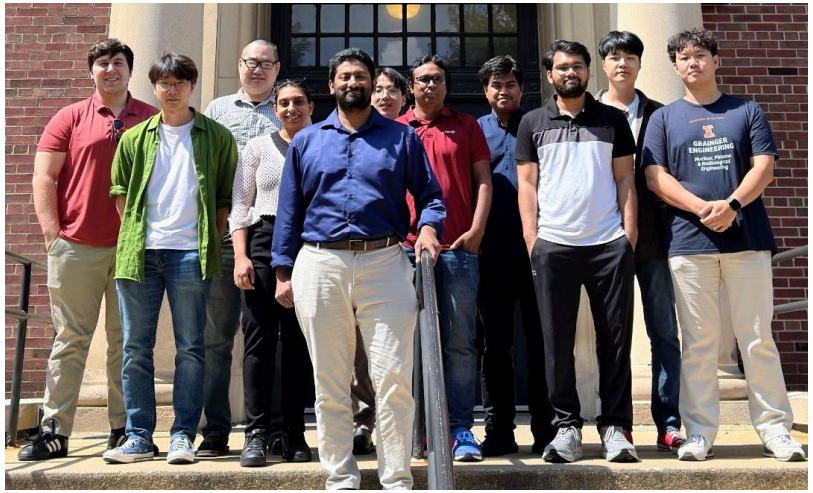
MARTIANS LAB

University of Illinois Urbana-Champaign

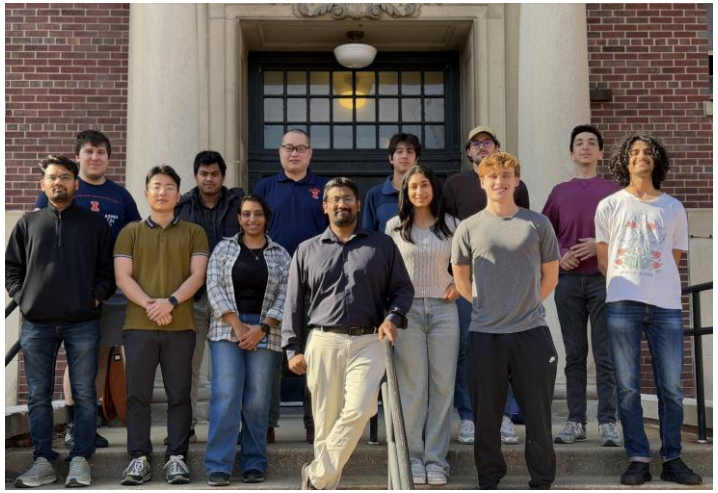
SCAN TO VISIT OUR LAB

martianslab.web.illinois.edu

2026



2025



2024



Thanks to my fabulous family, who literally did all the work and earned the recognitions:

- NSF Early Career Award 2026
- DOE Distinguished Early Career Award 2025
- American Nuclear Society Ted Quinn Early Career Award 2026
- NAE Frontier of Engineering 2026
- Featured in MIT Tech Review and Bloomberg
- Committee Members on *Foundation Models for Scientific Discovery* at the NASEM (Nationally 11 appointed)
- Named National AI Leader in UIUC's official response to the White House AI Action Plan (2025) [5 selected from UIUC]
- U.S. Nuclear Regulatory Commission (NRC) Distinguished Faculty Award
- Selected as 1 of ~100 Global Leaders by the Simons Foundation for advancing *Foundation Models for Science*.
- ANS 40 Under 40
- Global Supercomputing Award 2025 (first time for Grainger UIUC)

AI/ML is extremely promising in Engineering



Courtesy: Prof. Seid Koric, NCSA, UIUC

But the classical domain expertise is still needed !

What is a Digital Twin

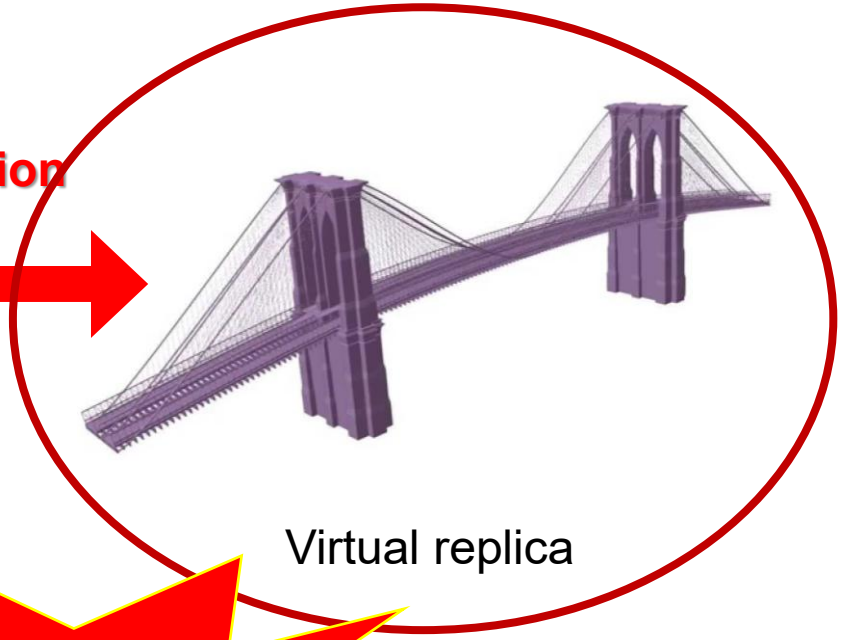
A digital twin is a virtual replica of a physical system

and is always '**in sync**'



Physical system

Temporal synchronization



Virtual replica

What is inside the virtual replica?

Courtesy: Prof. Souvik Chakraborty, IIT Delhi, Fulbright Prof, UIUC (2022)

What is a Digital Twin in Engineering?

Definition:

"A digital twin (DT) is a **virtual representation** of an object or system that **spans its lifecycle**, is **updated from real-time data**, and uses **simulation, machine learning and reasoning to help decision-making**." (IBM)

Benefits in Engineering Plants (by U.S. NRC 2021):

- **Increased operational efficiencies**
- **Enhanced safety and reliability**
- **Reduced errors**
- **Faster information sharing**
- **Better predictions**

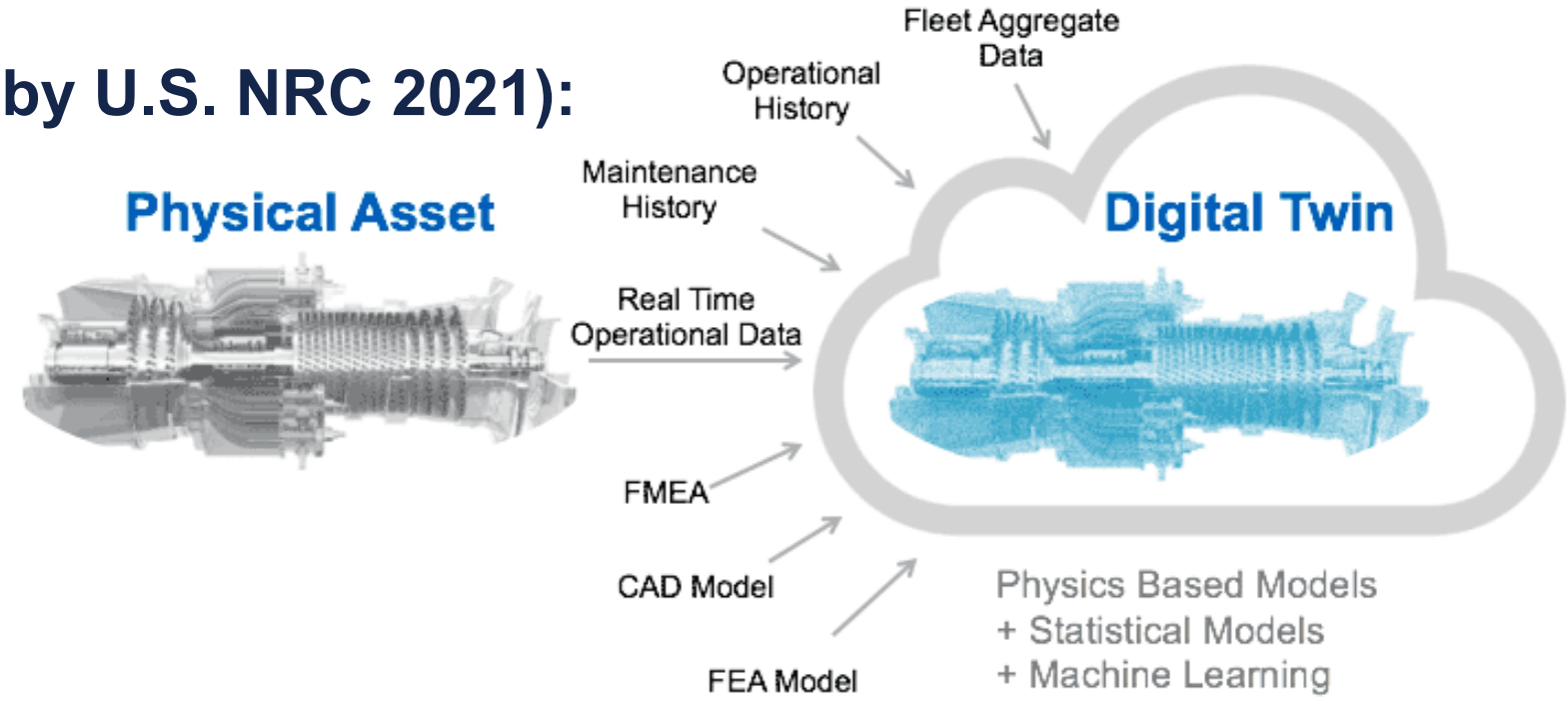
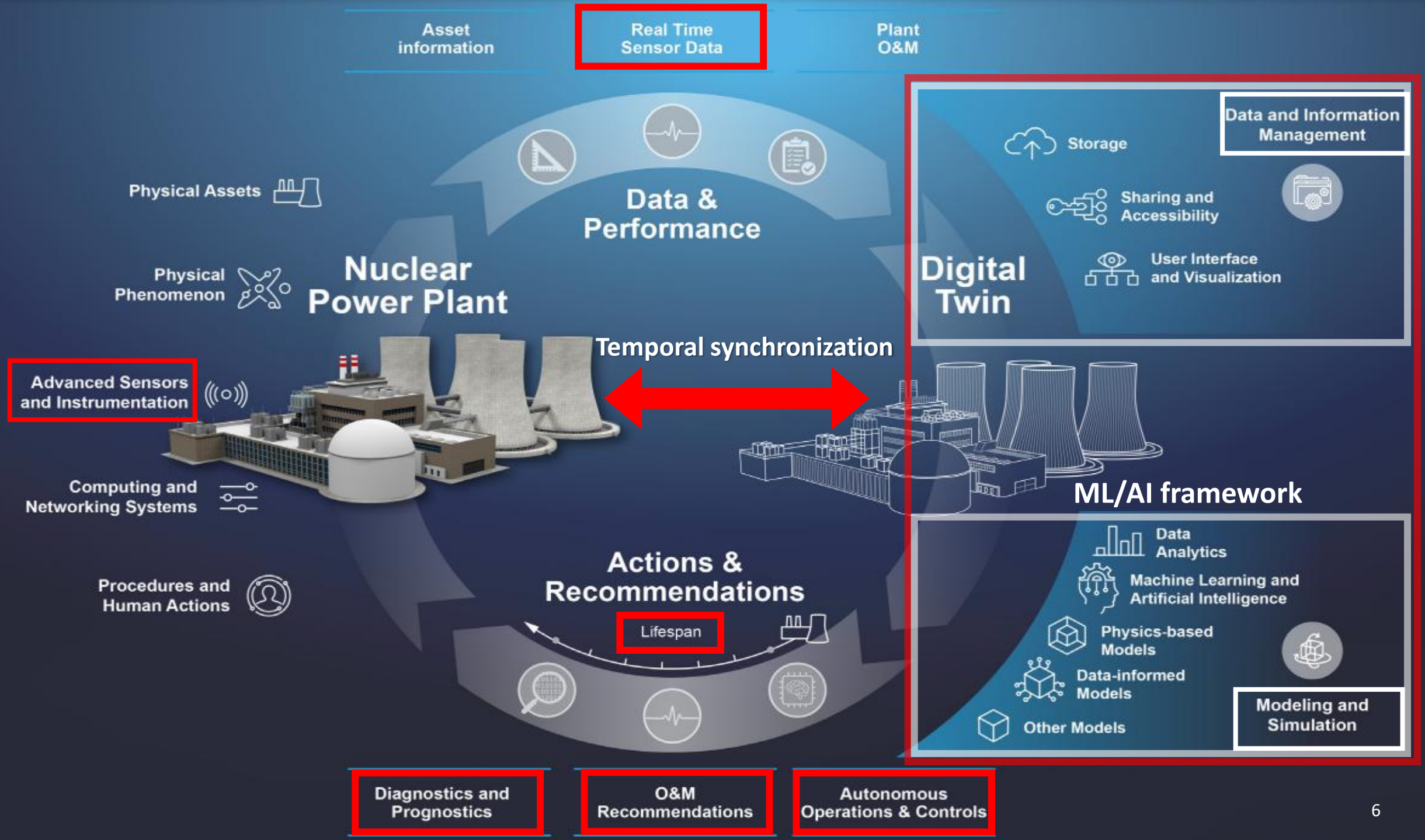


Figure courtesy: GE Research

Key DT Components: (V. Yadav, 2021)

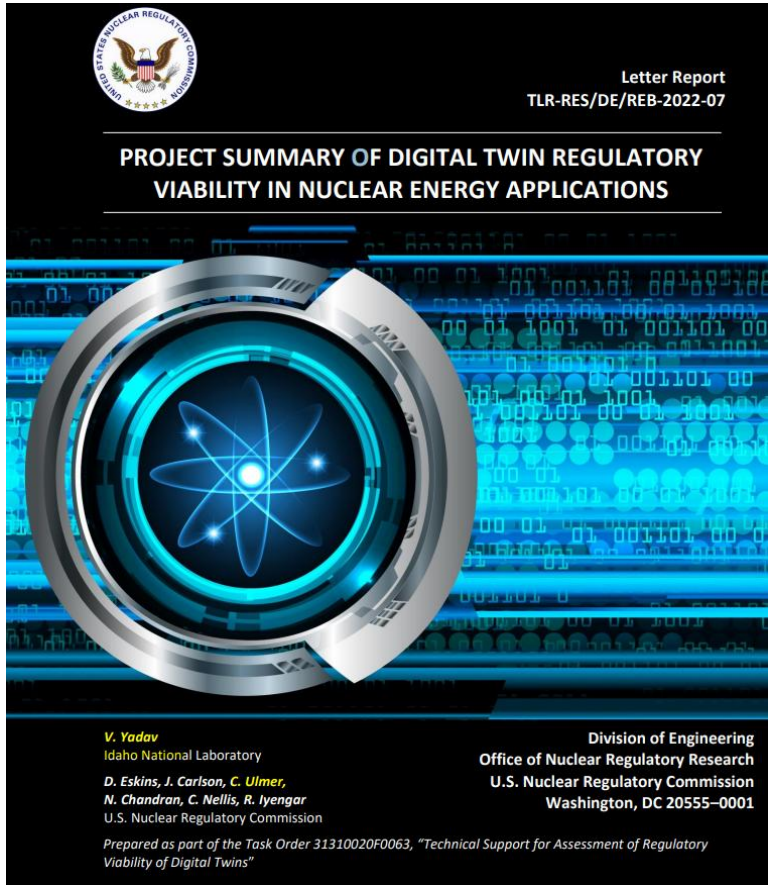
1. **Advanced Sensors and Instrumentation**
2. **Modeling and Simulation**
3. **Data and Information Management**

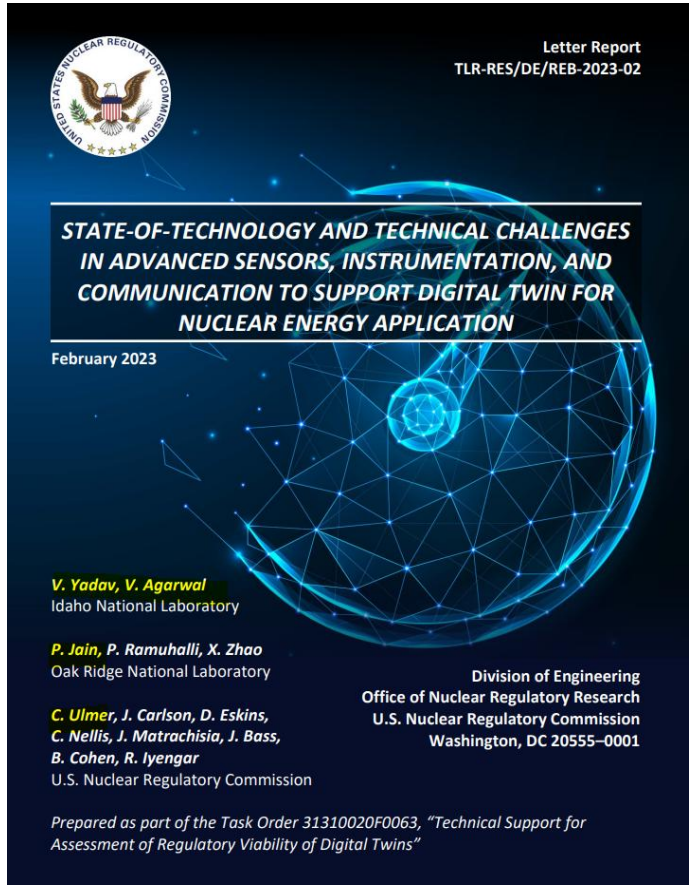
- **V., Yadav, V. Agarwal, P. Jain et al** (2021). Technical Challenges and Gaps in Digital-Twin-Enabling Technologies for Nuclear Reactor Applications. U.S. Nuclear Regulatory Commission.
- General Electric. https://www.ge.com/research/offering/digital-twin-creation


Asset information
Real Time Sensor Data
Plant O&M
Data & Performance
Physical Assets
Physical Phenomenon
Nuclear Power Plant
Advanced Sensors and Instrumentation
Computing and Networking Systems
Procedures and Human Actions
Temporal synchronization
Actions & Recommendations
Lifespan
Diagnostics and Prognostics
O&M Recommendations
Autonomous Operations & Controls
Digital Twin
Data and Information Management
Storage
Sharing and Accessibility
User Interface and Visualization
ML/AI framework
Data Analytics
Machine Learning and Artificial Intelligence
Physics-based Models
Data-informed Models
Other Models
Modeling and Simulation

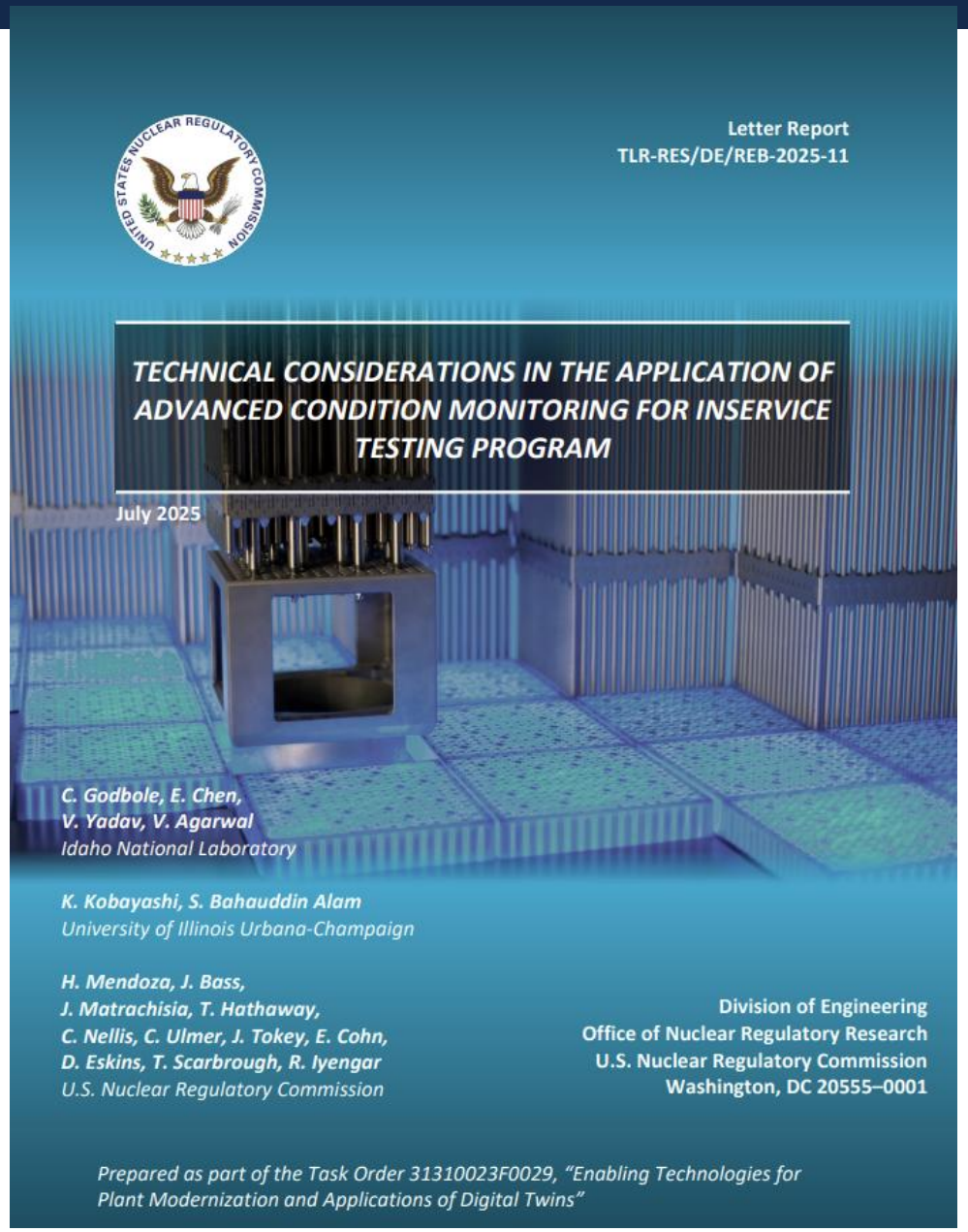
U.S. DOE & U.S. NRC Recommendations







Technical Challenges and Gaps Directed by U.S. DOE/NRC



1. Data Analytics

- ❑ ML algorithm selection
- ❑ Physics-based and Data-informed modeling
- ❑ Faster prediction: real-time applicability

2. Integration of Heterogeneous Data: Data Fusion

- ❑ Experimental data
- ❑ Synthetic or simulated data

3. Treatment of Noisy or Erroneous data (AI/ML training data)

- ❑ Processing noisy data, data containing erroneous observations
- ❑ Lack of data, missing data, outliers, and inconsistencies in data

4. Uncertainty Quantification

- ❑ Uncertainty quantification with concerns in Point #3
- ❑ Verification and validation

5. Real-Time Update: on-the-fly temporal synchronization

6. Explainable AI (XAI)

The Real Elephants in the Room: for Engineering Systems

1. Intelligence is resource-constrained: AI is not designed that way

Engineering platforms operate with **limited power, compute, and connectivity**

Most AI assumes cloud access and abundant resources

→ **Deployable & Energy-Efficient AI is a feasibility requirement, not an optimization**

2. Sensor failure is the norm, not the exception

sensors degrade, drift, fail, or are cyber-manipulated

Most monitoring collapses when data is missing or corrupted

→ **Physical AI is required to infer truth beyond sensors**

3. Inaccessible regions are treated as uncertainty/unknown

Subsea pipelines, risers, and internals cannot be directly sensed

“Unknown” is wrongly treated as “unobservable.”

→ **Physical AI must convert inaccessibility into inference**

4. Risk evolves faster than human workflows

Weather, aging, and cyber threats change continuously

Static models wait for human intervention

→ **Agentic AI is required to adapt and decide in real time**

5. Risk scales across fleets, intelligence does not

Current solutions are **asset-by-asset**

Failures and cyber threats propagate system-wide

→ **Foundation Models → Scale generalization**

6. Cyber risk is treated as IT, not inference

Attacks target **data integrity**, not just control systems

Models assume data is honest

→ **Agentic Physical AI reason about trust, not just prediction**

7. Digital Twins mirror systems: systems that think

Most Digital Twins are descriptive and data-driven

They fail when data degrades or disappears

→ **Digital Twins must be powered by Agentic Physical AI**

White House Genesis Mission: *A National Mission to Accelerate Science Through Artificial Intelligence (2025)*

NASEM study: *Foundational Research Gaps and Future Directions for Digital Twins (2024)*

NASEM study: *Implications of Artificial Intelligence-Related Data Center Electricity Use and Emissions: A Workshop (2025)*

Vision/Issues/Solutions?

01

From Seeing Systems → To Inferring the Unseeable

Enable virtual sensing (Physical AI) to reconstruct flow, stress, corrosion, leakage, and integrity where sensors cannot exist or have failed. Focus on inaccessible regions.

- Virtual sensing models
- Physical AI prototypes
- Uninstrumented region risk reports

02

From Static Digital Twins → To Agentic on/offshore Intelligence

Transform Digital Twins into agentic systems that detect, reason, and adapt under changing threats. Act before operators know there is a problem.

- Agentic Digital Twin platform
- Adaptive risk detection
- Automated response protocols

03

From Asset-by-Asset Models → To Gulf-Scale Foundation Intelligence

Develop Foundation Models for on/on/offshore energy that generalize safety intelligence across platforms, pipelines, and subsea networks. Share risk insights fleet-wide.

- Foundation AI models
- Cross-platform risk insights
- Regional safety intelligence reports

04

From Cloud-Dependent AI → To Energy-Aware Edge Intelligence

Design energy-efficient AI that runs on-device, survives intermittent connectivity, and prioritizes inference under strict power budgets. Edge AI protects when cloud fails.

- Edge AI prototypes
- Energy efficiency benchmarks
- Offline inference solutions

05

From Cybersecurity as IT → To Cyber-Physical Reasoning

Unify cyber and physical inference, enabling AI to reason about data trust, spoofing, and deception in real time. Address data integrity attacks.

- Cyber-physical inference engine
- Real-time data trust analytics
- Deception detection systems

nature communications

Article | Open access | Published: 19 September 2026

Virtual sensing to enable real-time monitoring of inaccessible locations & unmeasurable parameters

Kazuma Kobayashi, Farid Ahmed, Jaewan Park, Subhankar Sarkar, Seid Koric, Souvik Chakraborty & Syed Bahauddin Alam ✉

Nature Communications (2026) | Cite this article

Online & Real-Time Monitoring Methods

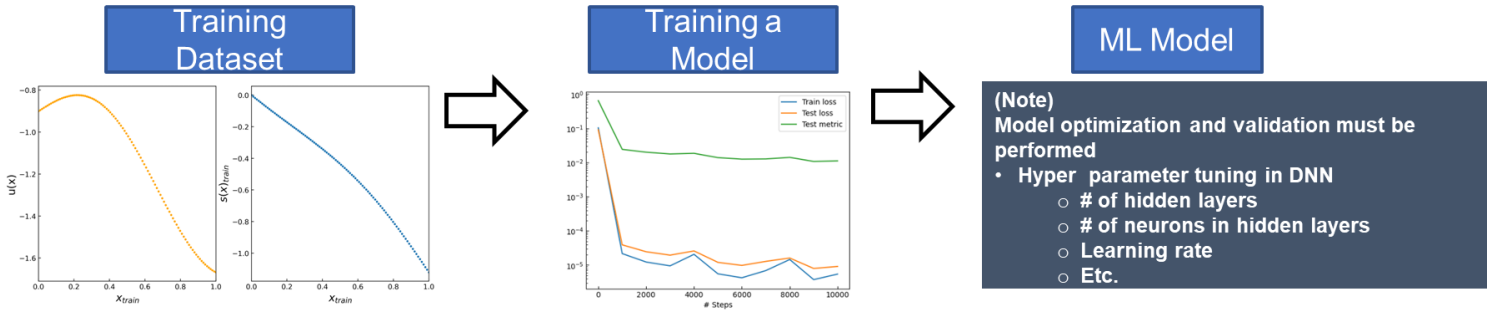
Online & Real-Time Monitoring Methods

Challenges[1] in ROM on Online and Real-Time Monitoring:

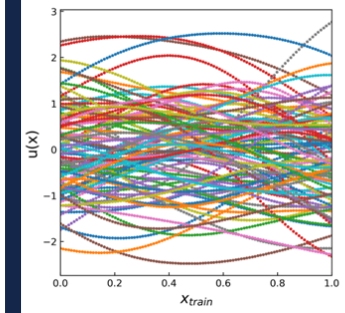
- **Dataset Shift**: ML models trained on specific plant operating conditions may --not generalize well to new, varied operational conditions, leading to decreased model performance and reliability.
- **Frequent Retraining**: Constant retraining (or Transfer learning) of ML models to adapt to new conditions is impractical.

- **Develop algorithms** capable of making accurate predictions in real-time, even in complex environments.
- **Integrate sensor data from physical assets** into digital twin systems for real-time monitoring.
- **Train models** with a diverse set of operational conditions to reduce the need for retraining

What we are doing now



What exactly is it in real life



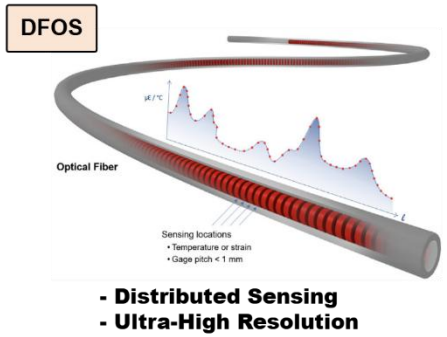
Build **100** Models (or Simulations)

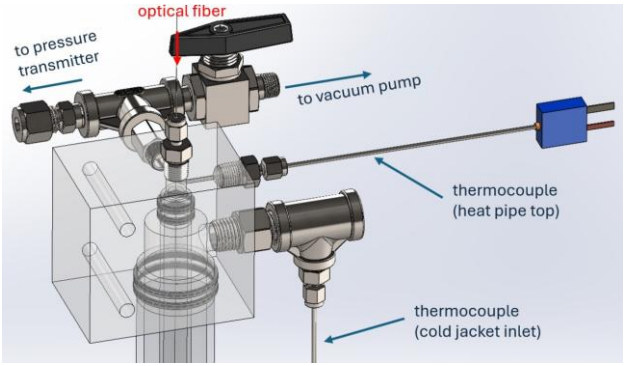
- Kobayashi, K., **Alam, "**Distributed Sensors for Realistic Digital Twin", *PNAS (under review*), 2024
- R. Hossain, **S. Alam**, "Impact of Sensor Degradation on Digital Twin Monitoring," *Nature NPJ Materials Degradation*, 2024.
- Kobayashi, K., **Alam, S. "**Real-time inference to enable digital twin solutions for nuclear Engineering systems". *Sci Rep* **14**, 2024.
- Kobayashi, K., **S.B. Alam**. "Improved generalization algorithms for engineering systems. "Engineering Applications of Artificial Intelligence," 2024.

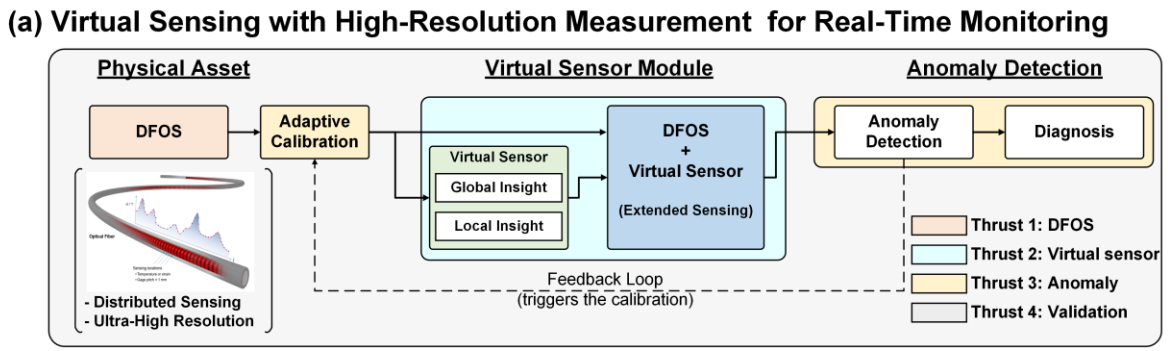
Near-Real Digital Twin for Online Inference for Inaccessible Locations

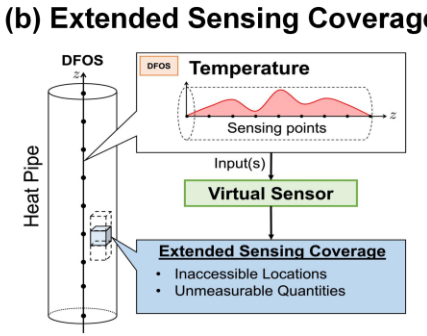
All good! But can algorithms' capture sensors' spatiotemporal information

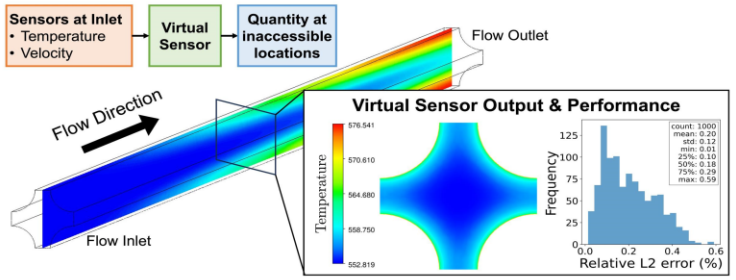
!? Data-Model Fusion – Physics-integrated deep learning
!? Physics-Informed Neural Networks (PINNs)
!? Multi-Fidelity Neural Network (MFNN)
✓ Functional Mapping (Universal approximation theorem) → Spatiotemporal











Scalability? Why and What? benchmark Navier-Stokes

$$u_{\text{inlet}}(t) \mapsto \mathbf{u}(x, y, t)$$

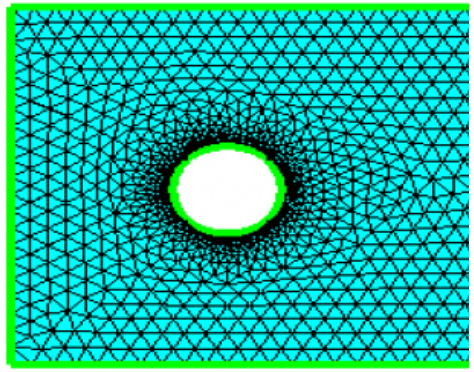
$$\nabla \cdot \mathbf{u} = 0$$

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} = -\frac{1}{\rho}\nabla p + \nu \nabla^2 \mathbf{u}$$

$$u_{\text{inlet}}(t) = U_0(t)$$

(Note) Spatial resolution inherits from a training data (i.e., FEM)

$u_{\text{inlet}}(t)$

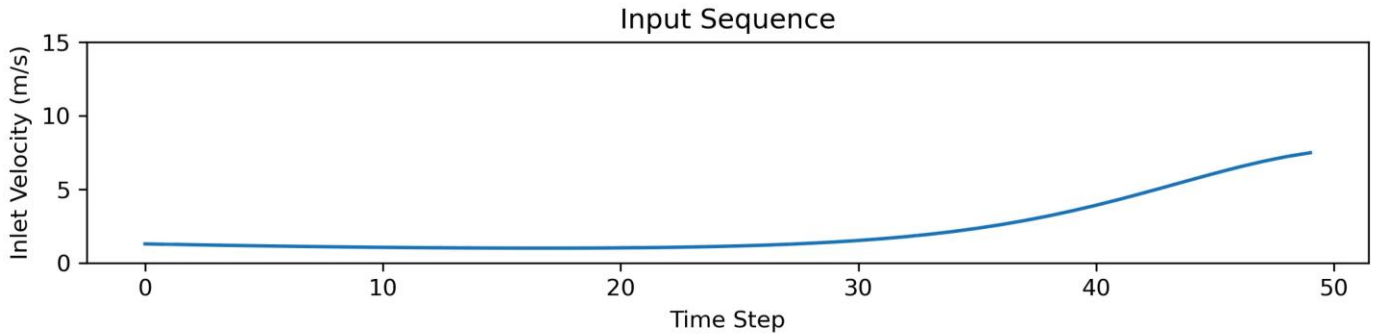


Flow meter at the inlet

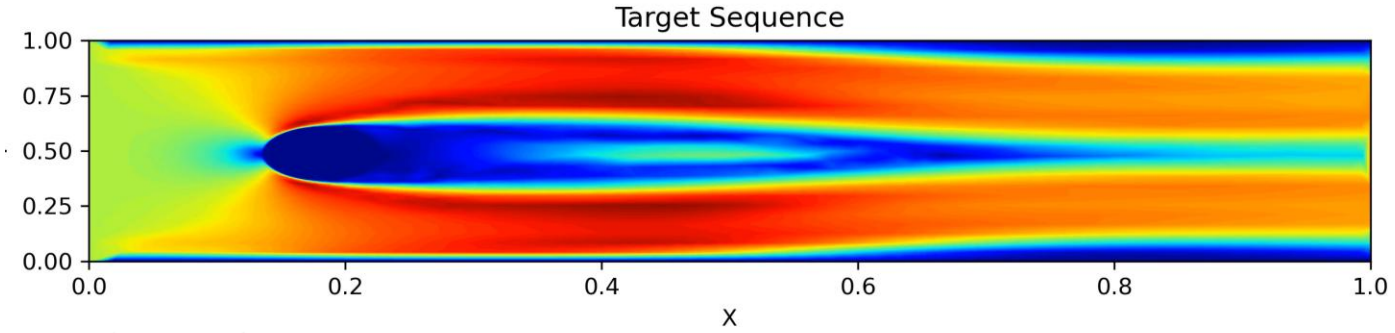
$u_{\text{inlet}}(t)$

Velocity field in the system

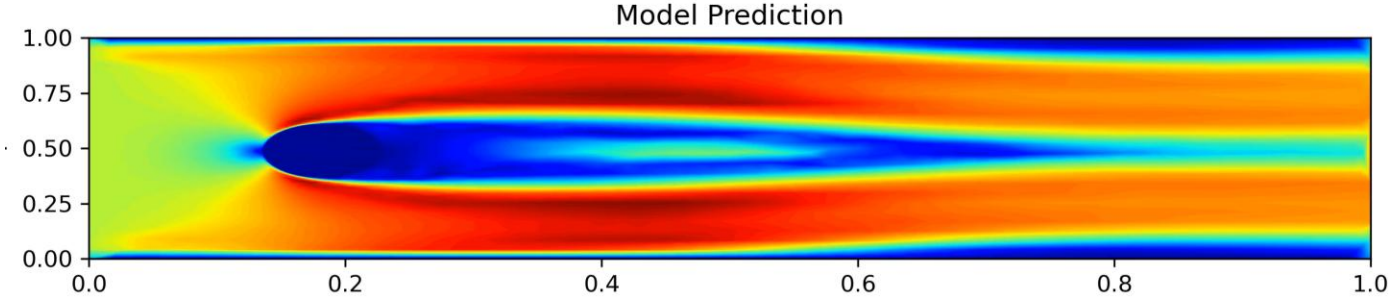
$\mathbf{u}(x, y, t) = (u(x, y, t), v(x, y, t))$

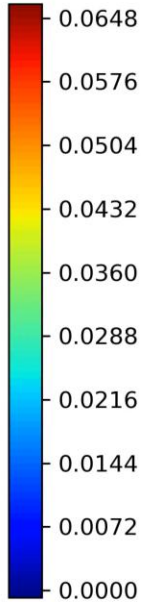


$u_{\text{inlet}}(t)$



$\mathbf{u}(x, y, t)$

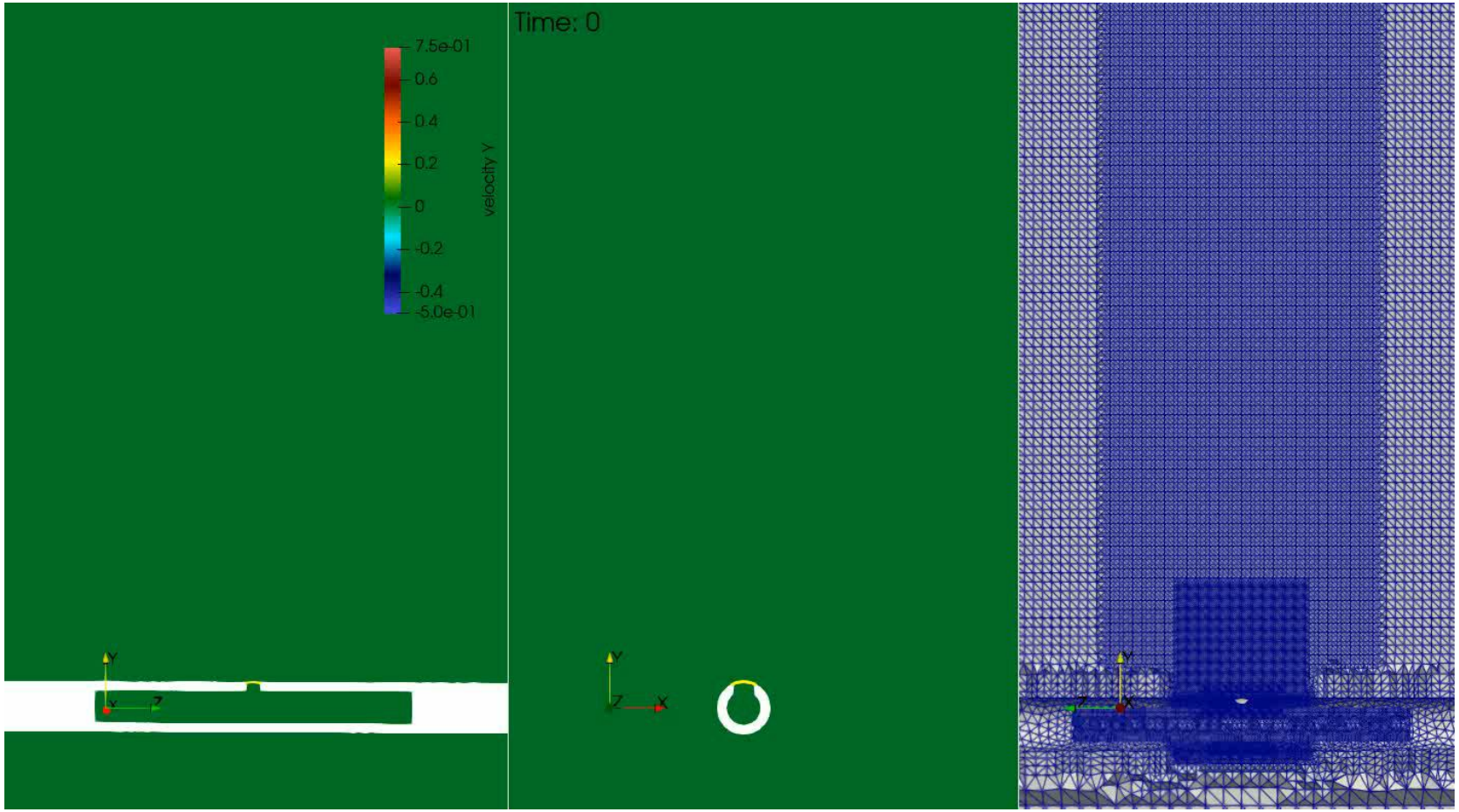




Scalability? Issues or Challenges?

Locally varying velocity component: Air-Water two phase flow

- *Inflow rate : 50ml /min*
- *Recirculation shifts back and forth from center of the bubble to edge*

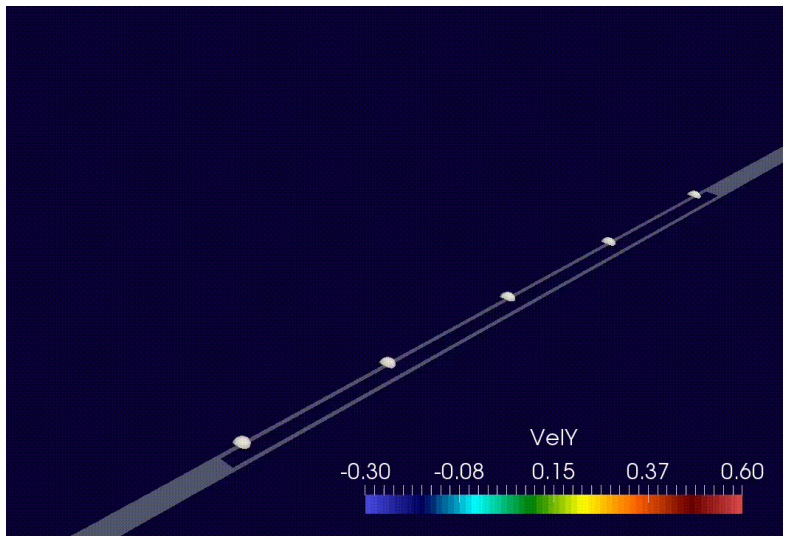


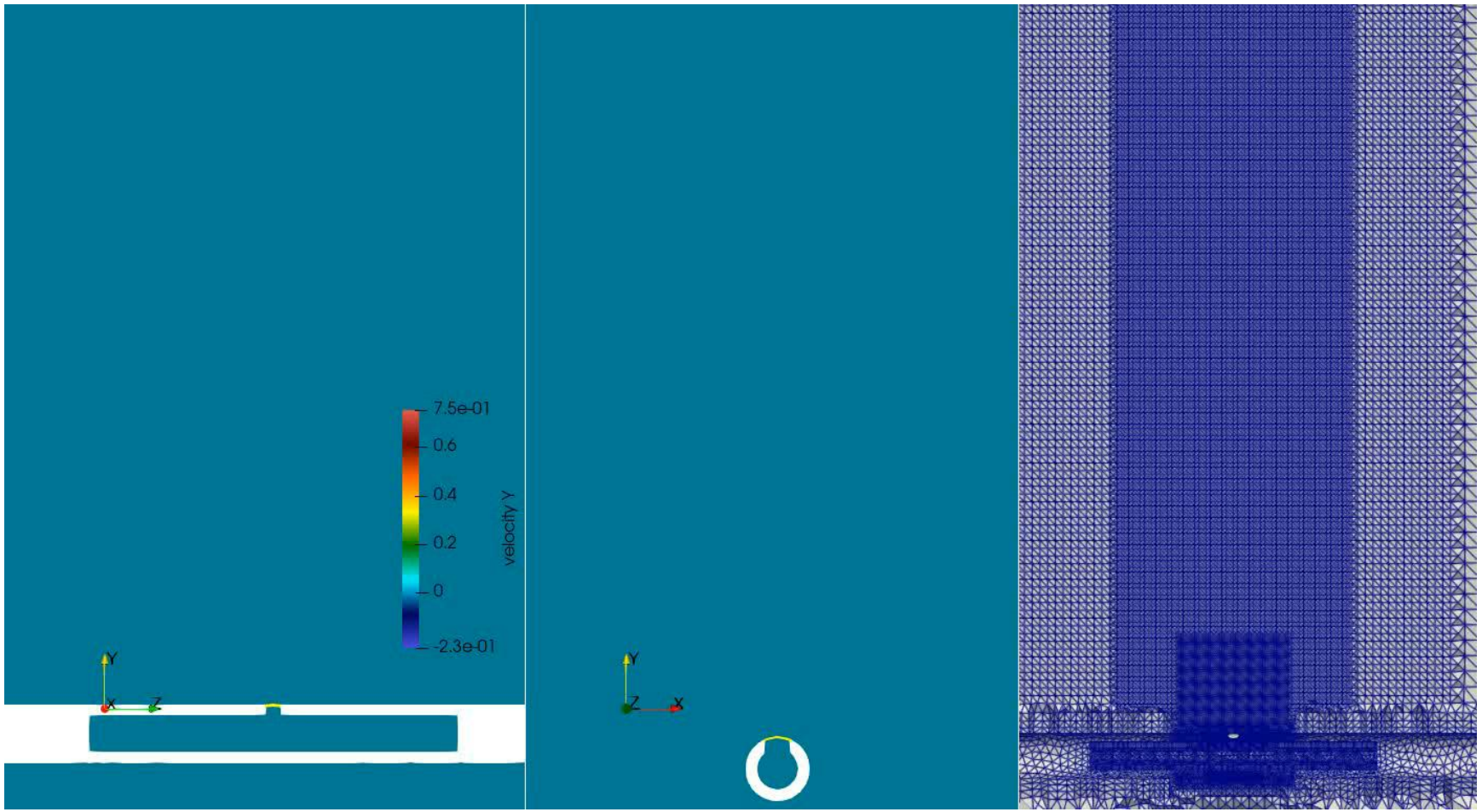
Courtesy: F. Ahmed, ALAM Lab, NPRE/UIUC

Scalability? Issues or Challenges?

Locally varying velocity component: FLiNaK-He two phase flow

- *Inflow rate : 50ml /min*
- *Recirculation shifts back and forth from the center of the bubble to the edge*

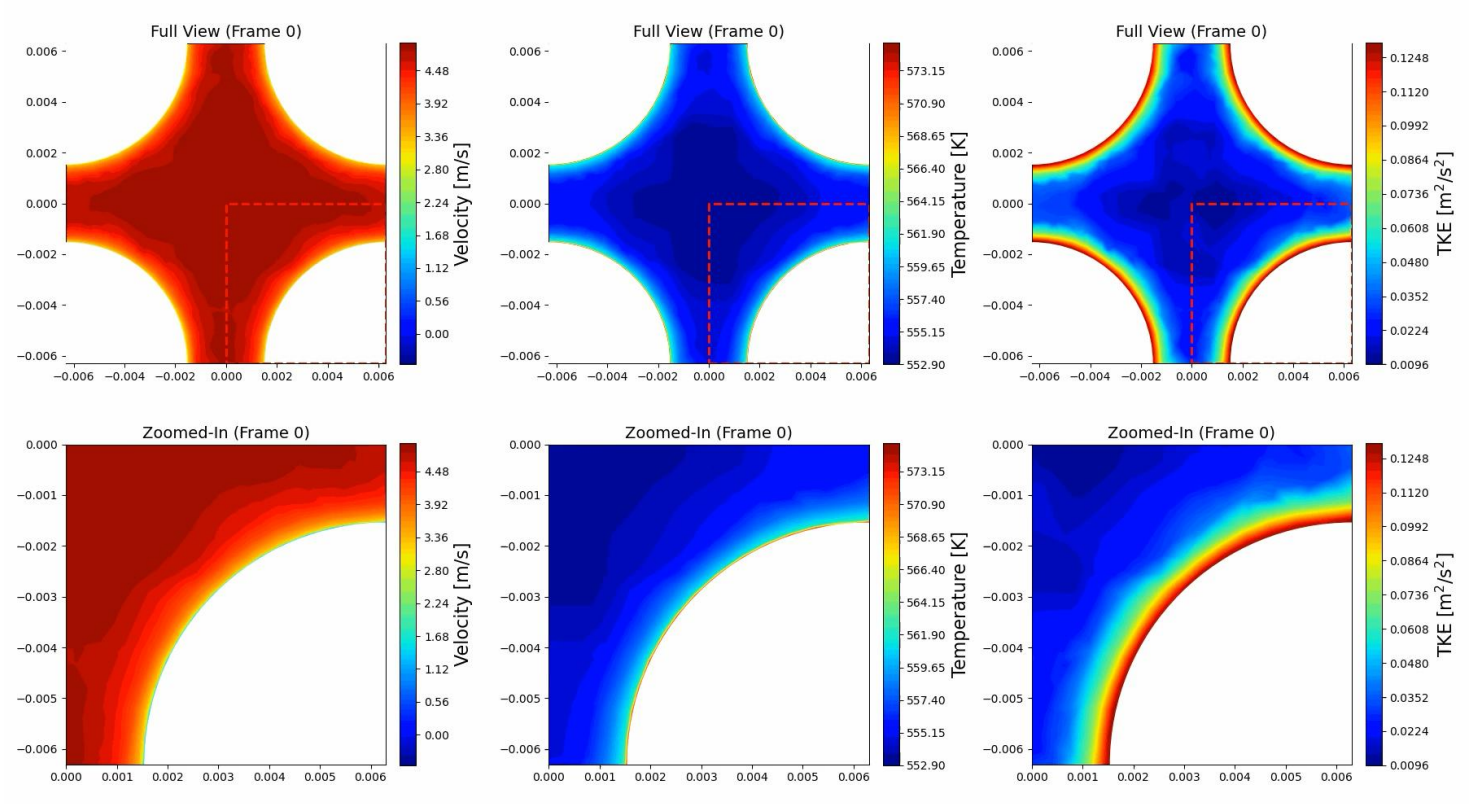


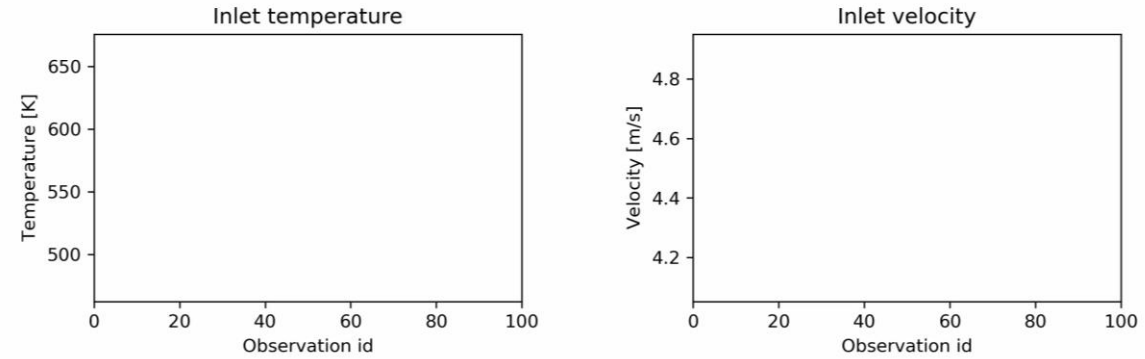


Courtesy: F. Ahmed, ALAM Lab, NPRE/UIUC

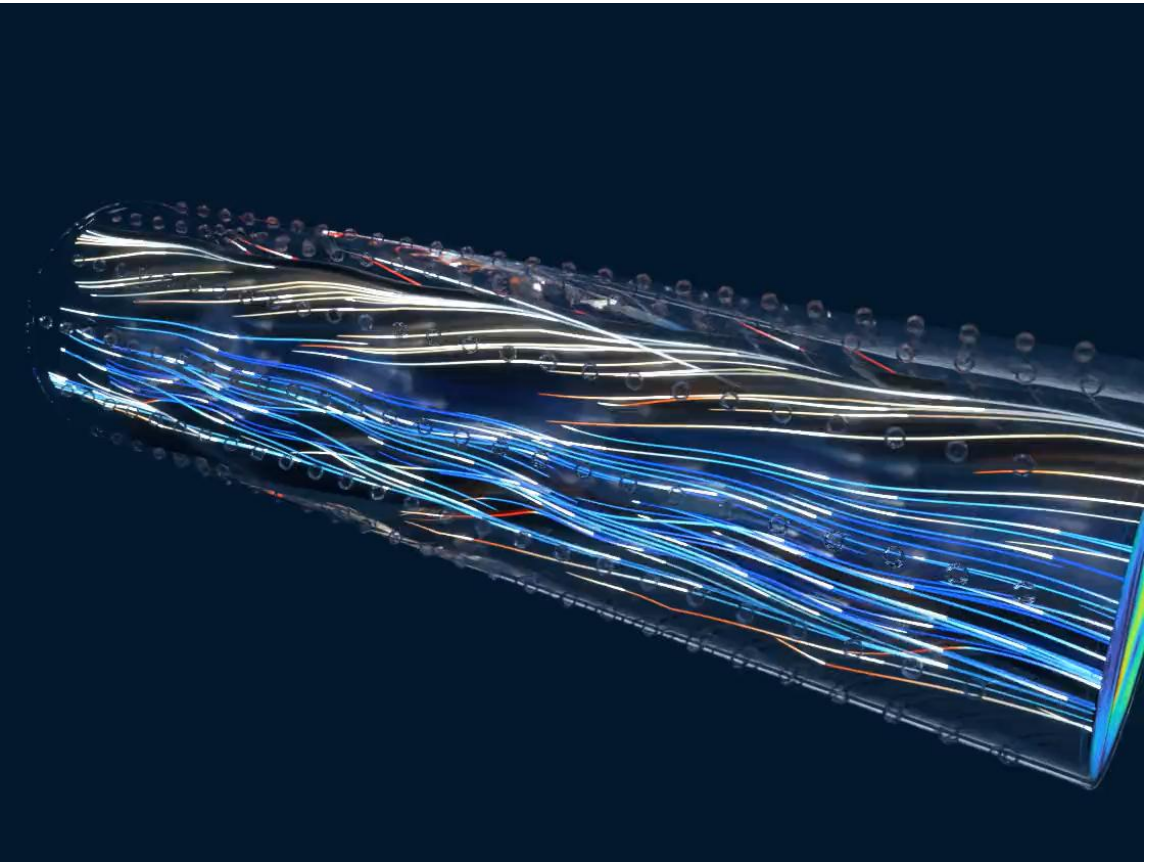
Example: Coolant Channel in Engineering Systems

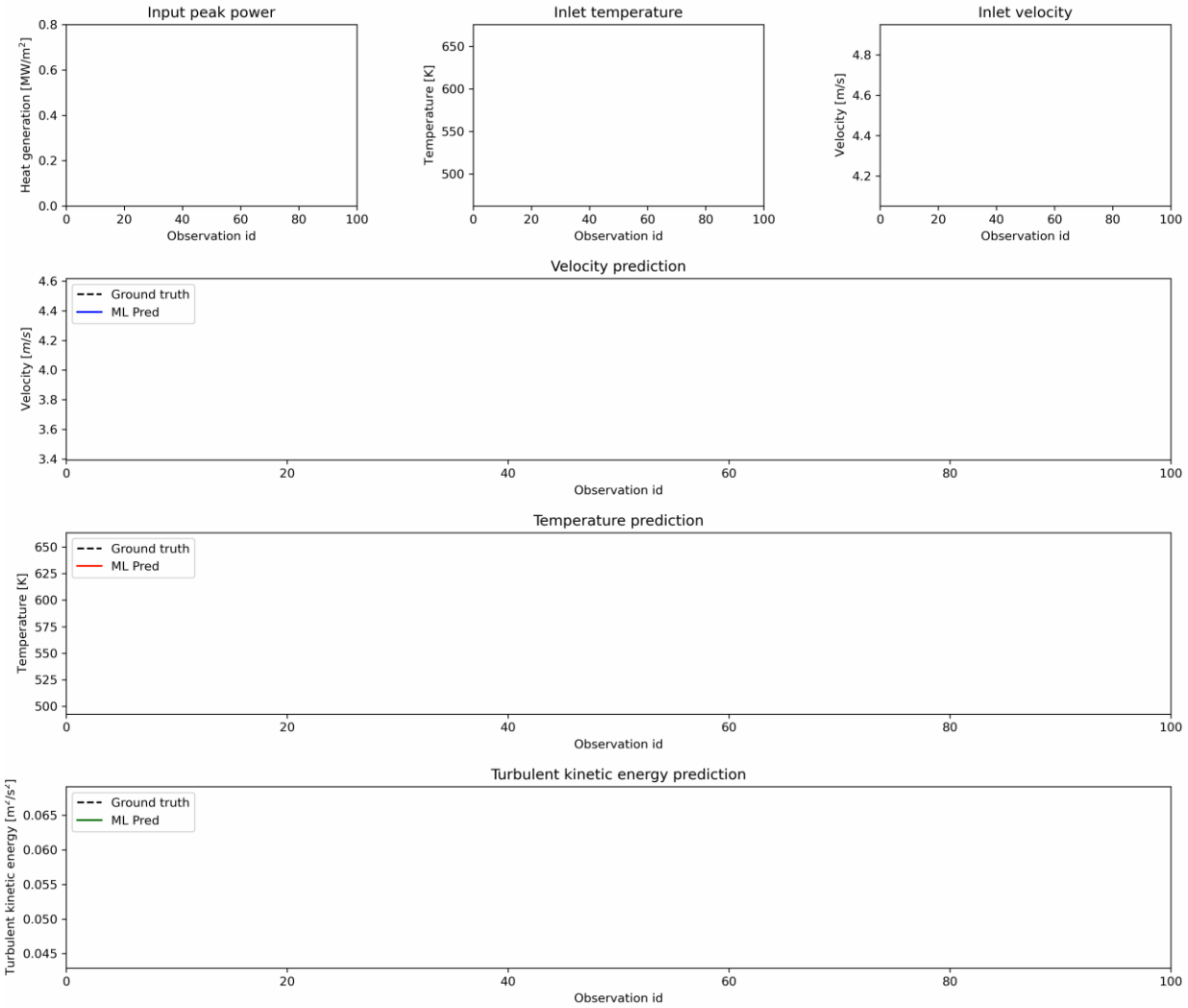




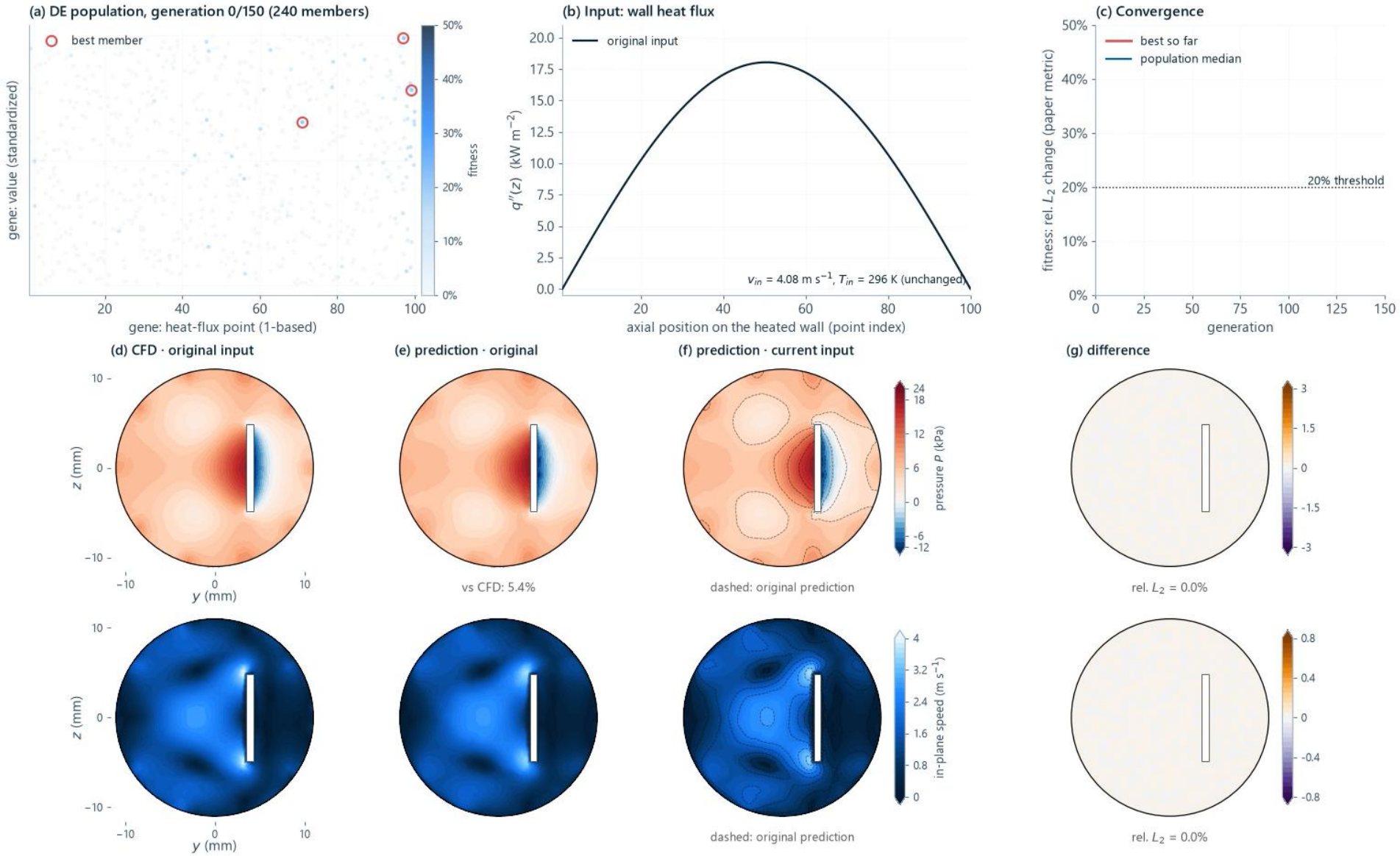


Cooling System



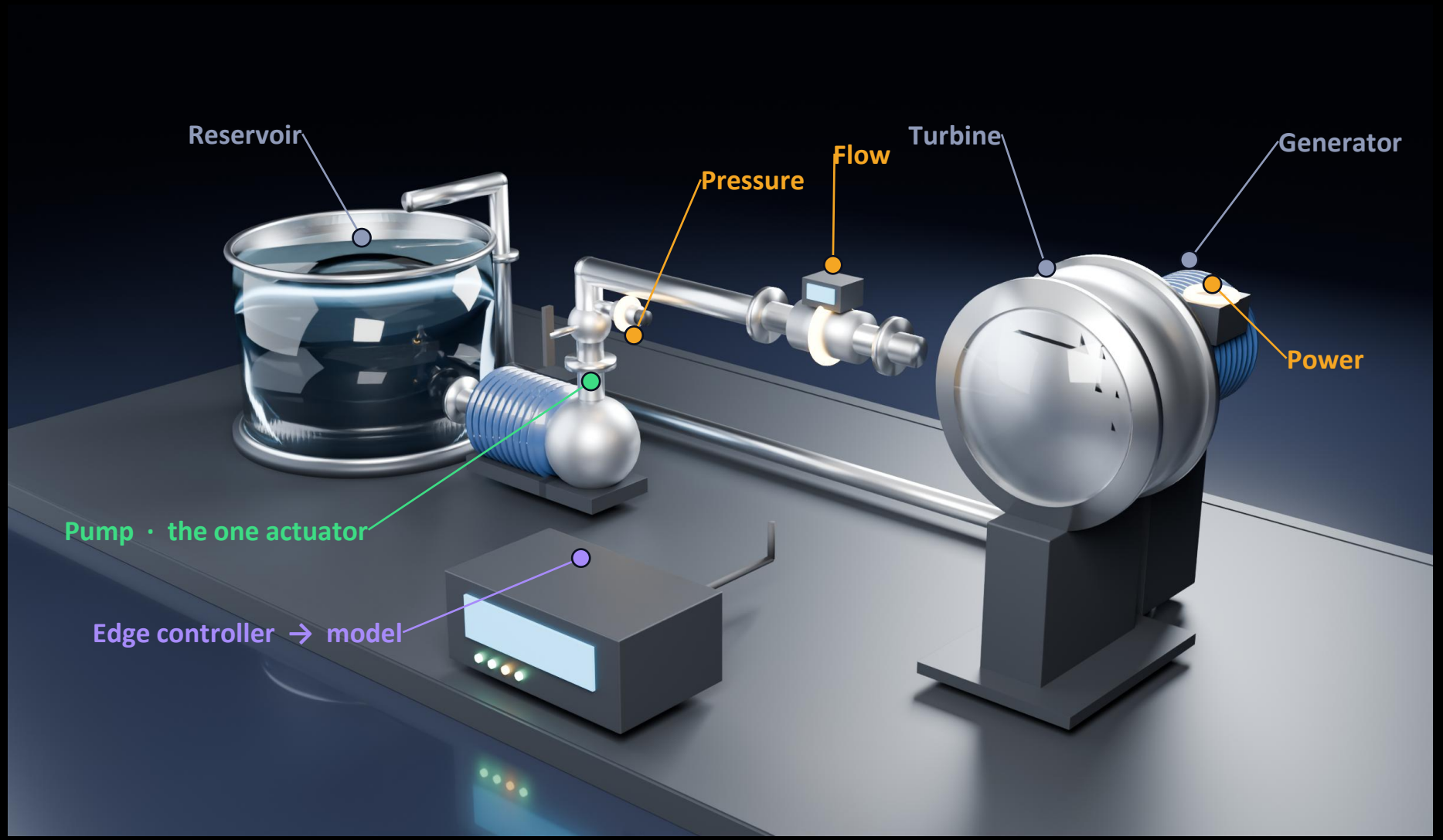


Cyber Attacks and Defense Strategies for DT Models



Based on S. Roy et al., arXiv:2603.22525 (2026). Re-run of the paper's attack method (same settings, seed 1002) on the shipped S-DeepONet, test sample index 2.
The search finds which three values matter; the change shown is the smallest step in that direction that shifts the prediction by 20%.

Agentic AI on aPhysical Testbed



sensors

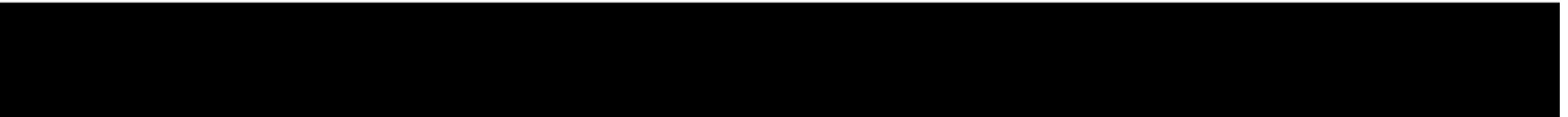
actuator
pump duty cycle

goal
hold a requested power

Multi-Domain LLM Agent: AI Agent for Nuclear Reactor

Agent handles nuclear reactor, acting as a text-based autonomous agent: Receives plant state as text input, Outputs control actions





LLM-Based Nuclear Fuel Assembly Optimization

LLM Agent as a Reactor Optimizer

LLM agents learn loading patterns through

Preference optimization outperforms classical search

.

Results

DPO fitness: 1.637

- Outperforms unconstrained GA baseline (1.834)

GRPO fitness: 1.715

- Superior to GA, with constraint satisfaction

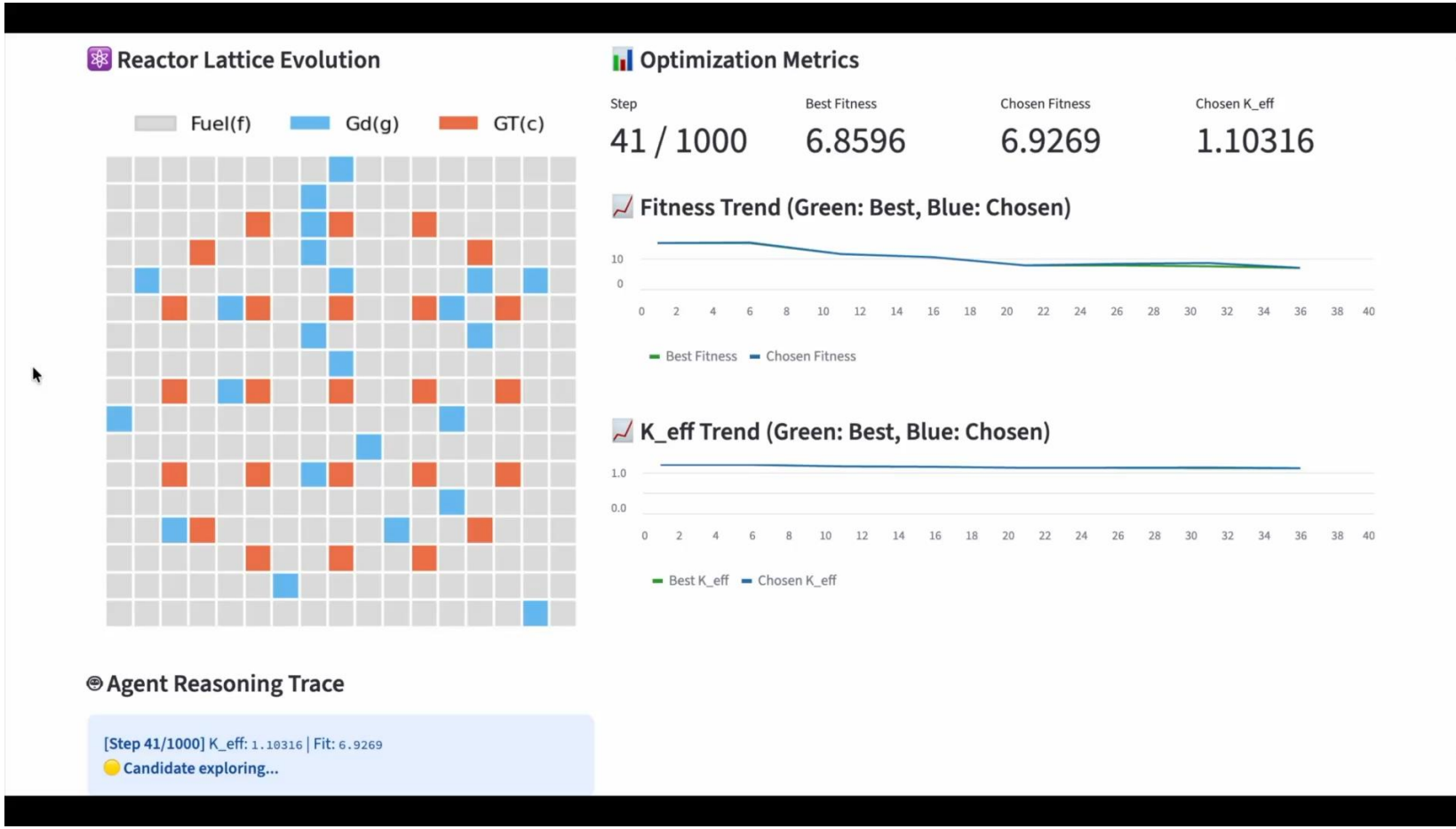
LLM agents satisfy engineering constraints

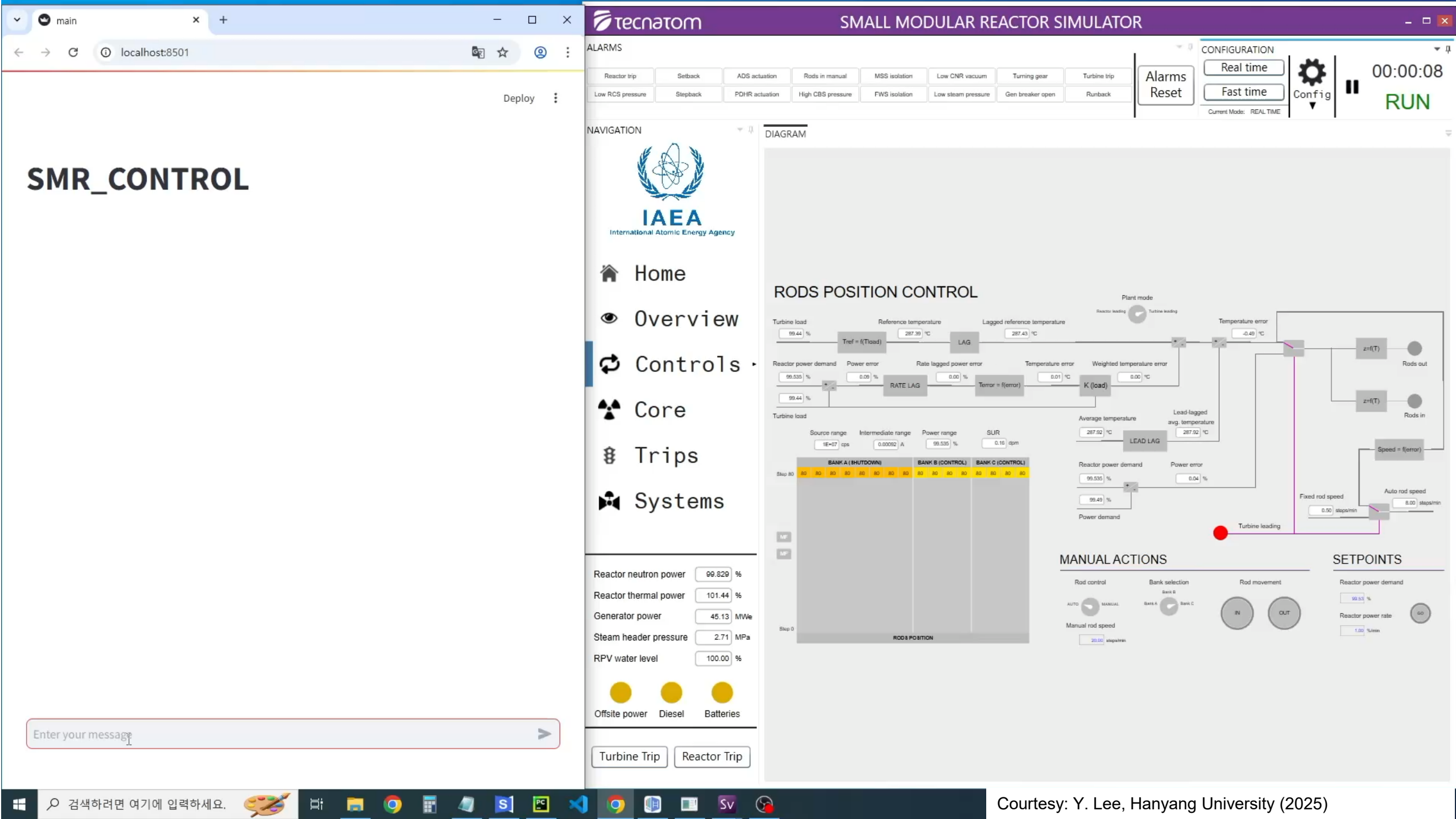
- e.g., exactly 16 Gd pins

Agentic Aspect

The LLM acts as an autonomous design agent:

- Proposes fuel configurations,
- receives fitness feedback from simulation,
- and iteratively improves via preference learning.

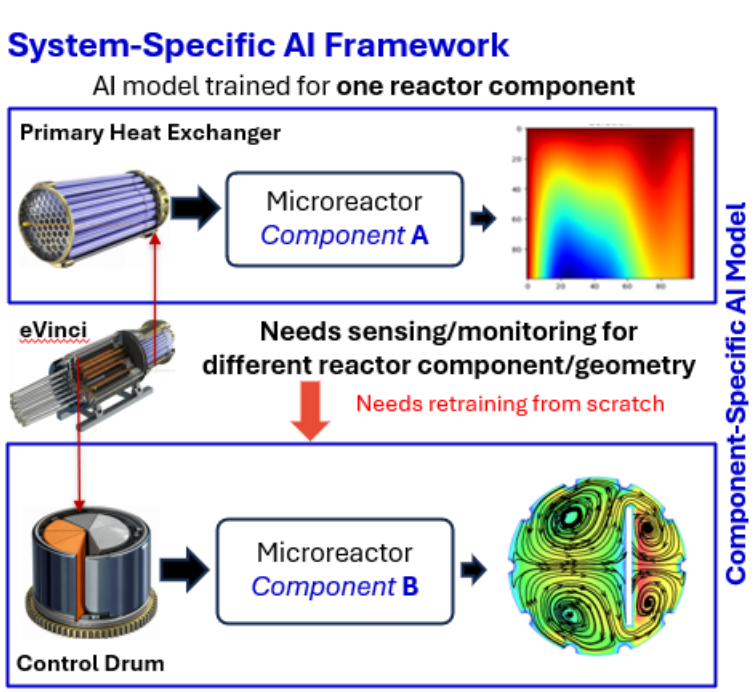




Courtesy: Y. Lee, Hanyang University (2025)

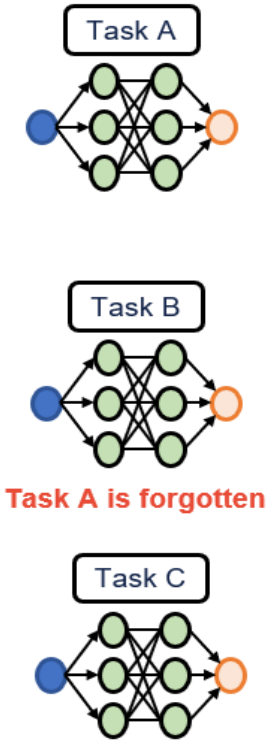
Bottlenecks of AI Models

Bottlenecks of AI Models—Why traditional AI/Digital twins not gonna work?



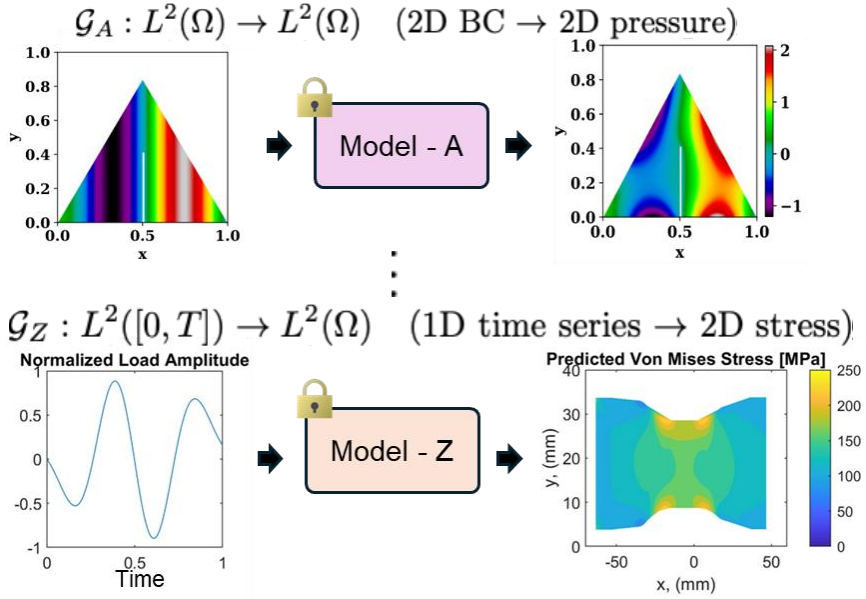
One-PDE-per-Model Paradigm

Each model is trained for a single PDE, requiring retraining from scratch with no shared learning and high time/data cost.



Continual Learning Bottleneck

Models forget previous tasks when trained on new ones, making continual learning difficult and leading to loss of prior knowledge.



Rigid Architectures (Cross-component/cross-geometry/cross conditions)

Architectures are inflexible, mapping specific inputs to outputs (e.g., 2D BC to 2D pressure), limiting adaptability across tasks.

Lee, Yoonpyo, S. Alam et al. "Agentic Physical AI toward a Domain-Specific Foundation Model for Nuclear Reactor Control." Nature Comm, review (2025).
Foundation Models for Scientific Discovery and Innovation, **National Academies of Sciences, Engineering, and Medicine (2025)-** Appointed as **one of Committee Members**

Towards Foundation Model Blueprint

Fundamental Reconstruction Problem

- Effective dose cannot be directly measured at scale→ The quantity of interest is unobservable.
- Only indirect secondary particles (neutrons) are observable→ Sparse, indirect proxies are measurable.
- Virtual sensing reconstructs radiation damage from sparse neutron measurements.

Neutron sensor located on the earth → **Effective Dose due to Cosmic-Ray Exposure**

*1 Retrieved from https://astronuclphysics.info/JadRadFyzika6.htm)

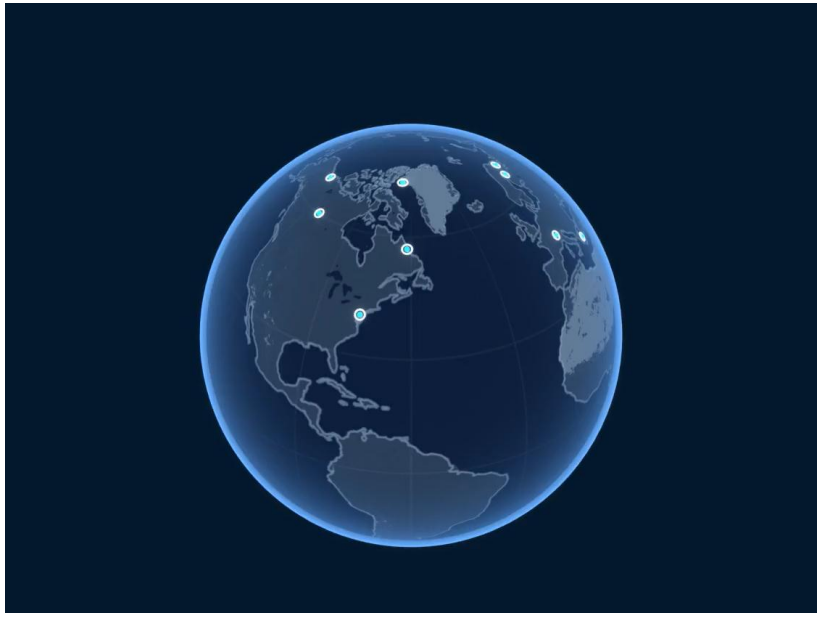
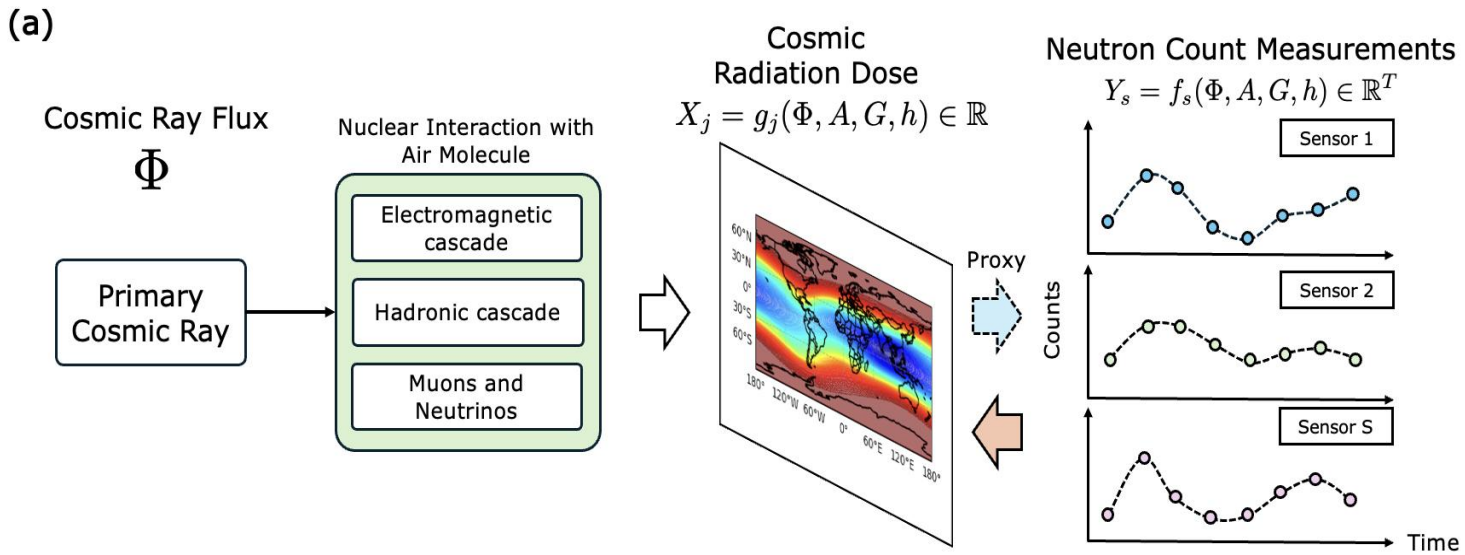


Fig. cosmic radiation*1



- What sensors see are sparse neutron time series at a limited locations.
- What we want is a **spatiotemporal field of radiation dose**, which is a latent physical quantity.

K. Kobayashi, S. Alam, et. al. From Proxies to Fields: Spatiotemporal Reconstruction of Global Radiation from Sparse Sensor Sequences, Under review at Nature Communications (2025)

Estimation of Effective Dose due to Cosmic-Ray Exposure

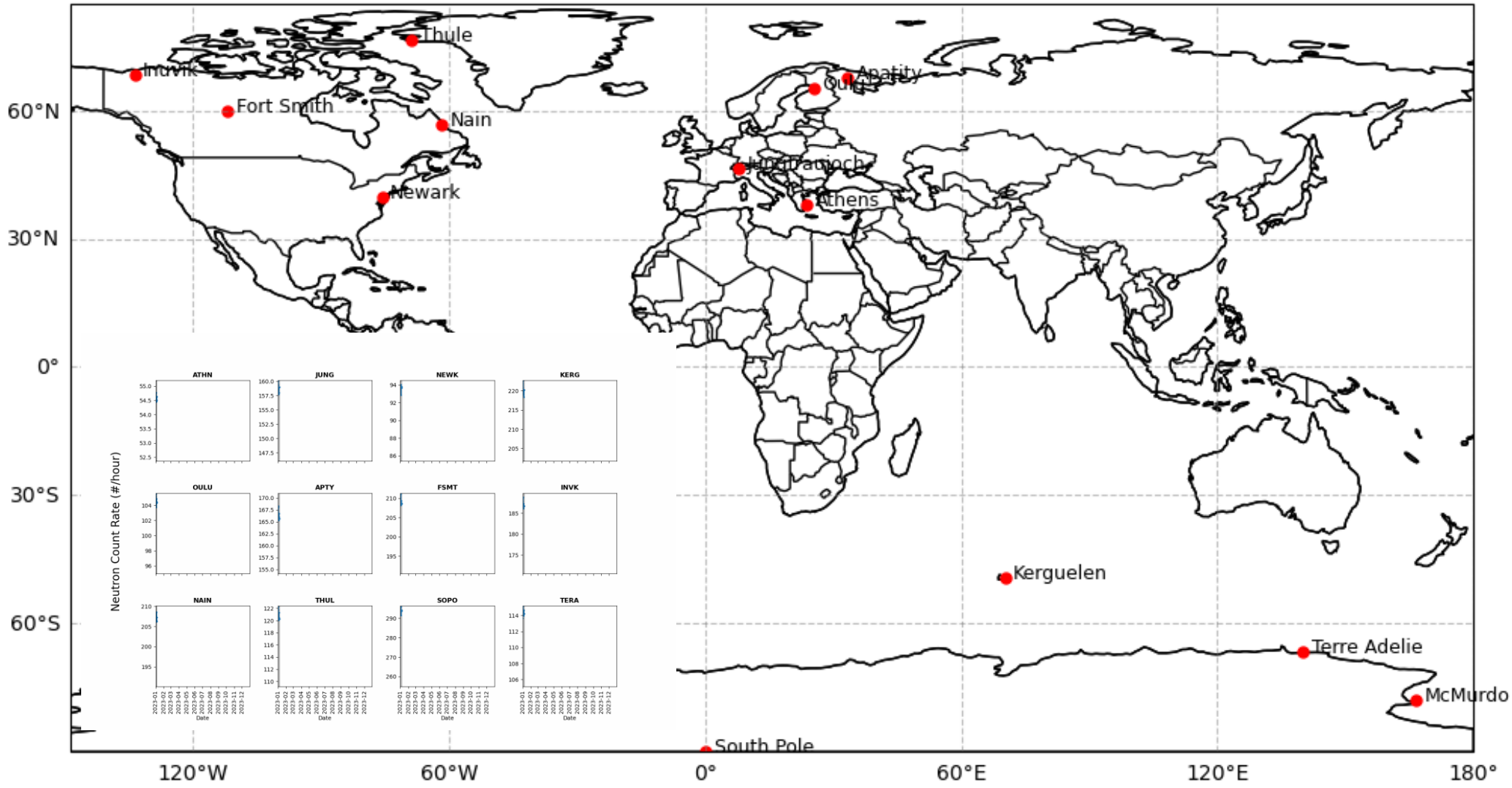
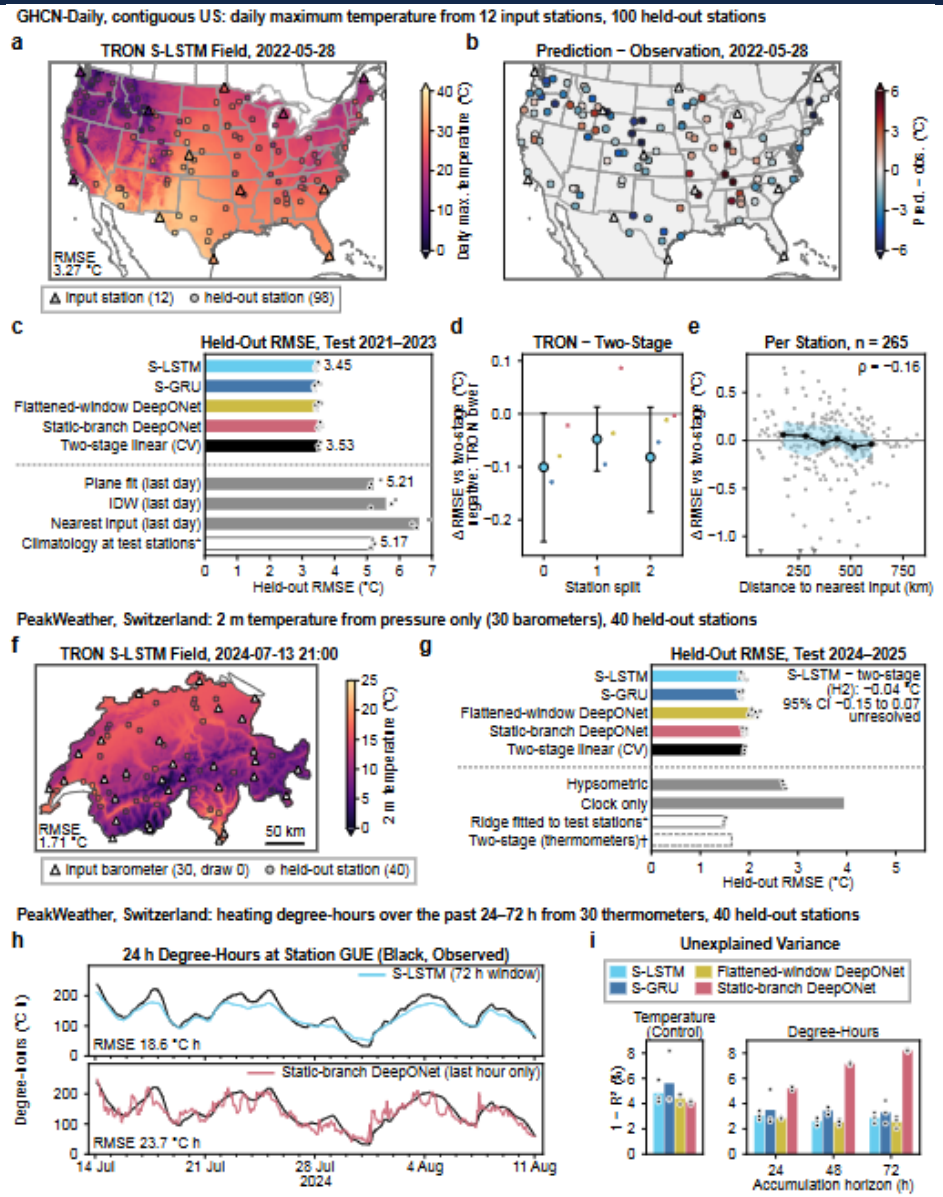
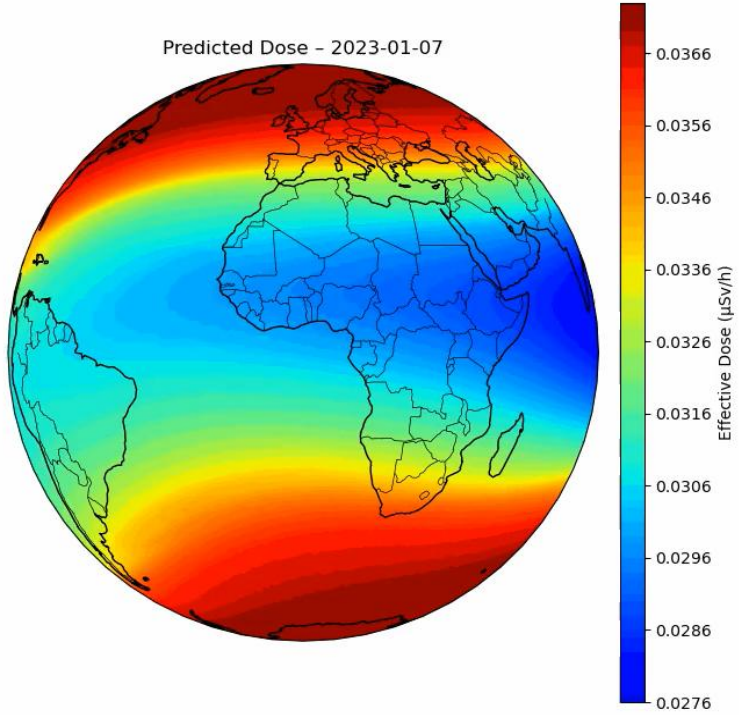


Fig. Neutron monitoring **12 stations** – sensor locations **(red markers)**

Example: Reconstruction of Radiation Dose Field





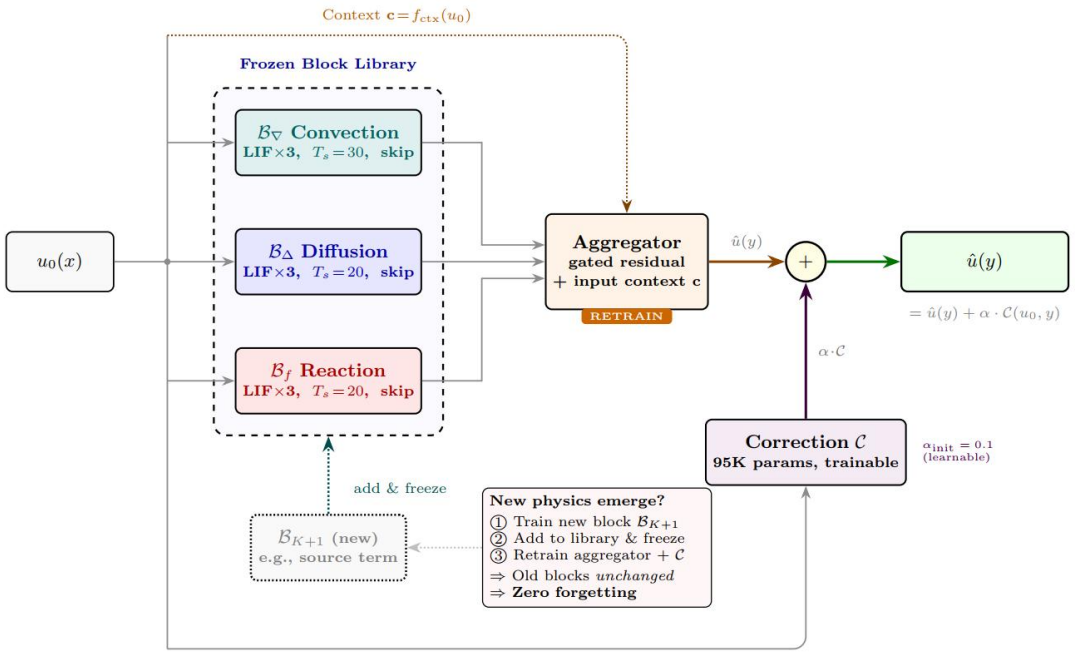
- (The model was trained with dataset from 2001-2021)
- Long & Lat resolutions: 1 degree

Table. Model performance

Model	Relative L_2 Error (%)	RMSE	MAE	R^2
FNN-DeepONet	1.64×10^{-1}	5.51×10^{-5}	4.75×10^{-5}	9.995×10^{-1}
S-GRU	1.34×10^{-1}	4.51×10^{-5}	3.67×10^{-5}	9.997×10^{-1}
S-LSTM	$\mathbf{9.96 \times 10^{-2}}$	$\mathbf{3.35 \times 10^{-5}}$	$\mathbf{2.75 \times 10^{-5}}$	$\mathbf{9.998 \times 10^{-1}}$
M-GRU	7.00×10^{-1}	2.36×10^{-4}	2.11×10^{-4}	9.905×10^{-1}
M-LSTM	9.00×10^{-1}	3.03×10^{-4}	2.55×10^{-4}	9.864×10^{-1}

K. Kobayashi, S. Alam, et. al. From Proxies to Fields: Spatiotemporal Reconstruction of Global Radiation from Sparse Sensor Sequences, Under review at Nature Communications (2025)

SCNO: Compose Simple Physics Blocks → Solve Reactor PDEs on Neuromorphic Hardware (Spiking Compositional Neural Operator)

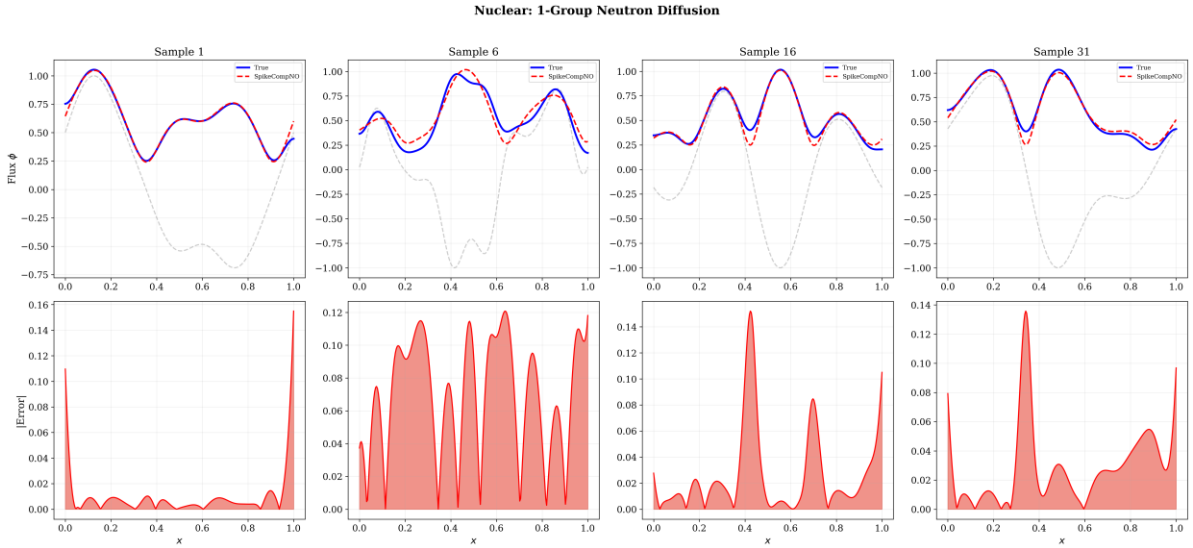


Modular — Train once on elementary operators (diffusion, convection, reaction), compose them to solve coupled PDEs never seen during training

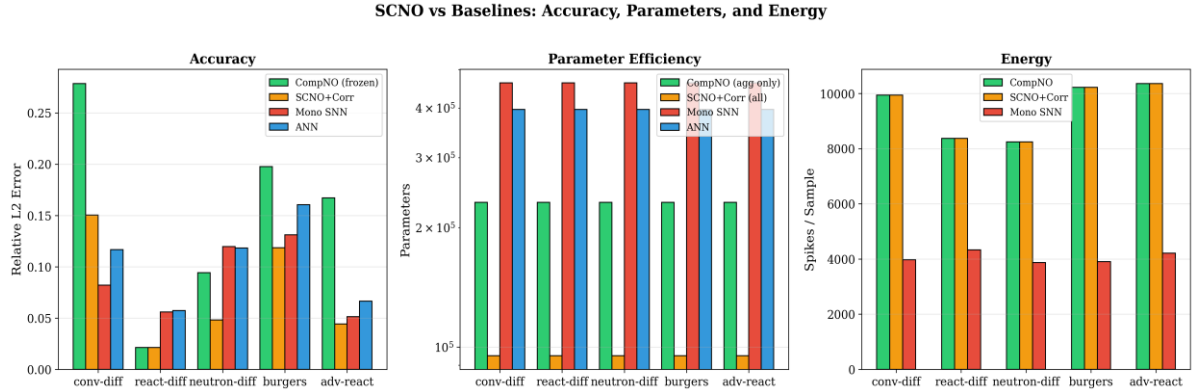
Neuromorphic — Spike-based inference on Loihi 2 chip (<1 W) instead of GPU digital twins (30–300 W), designed for remote reactor edge deployment

Expandable — New physics blocks added without retraining existing ones, zero catastrophic forgetting by architectural guarantee

Diffusion — Predictions vs Ground Truth



Accuracy, Parameters & Energy Across 5 Coupled PDEs



(a) Foundation Model

Capabilities

- Cross-domain generalization
- Multimodal Dataset
- Reusability

System Label (ℓ) — Which System?

Cooling System: {1,0,…,0}
Heat Pipe: {0,1,…,0}
⋮

Multiple Sensors (a) — Which Sensors?

Gauge, Function, Images, Text, Graph, History

Spatiotemporal Grid (x) — Where and When?

T_{i-2}, T_{i-1}, T_i; y, x

Context Embedding → MLP → Lifting Transformation

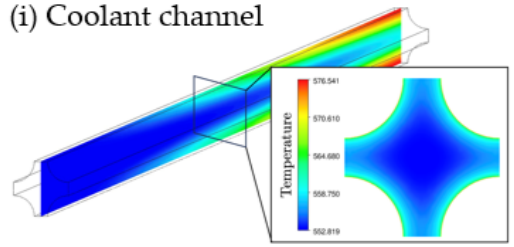
Neural Operators (PI's Prior Works)

v_0 → Expert Block 1 → v_1 → · · · → Expert Block h → v_h

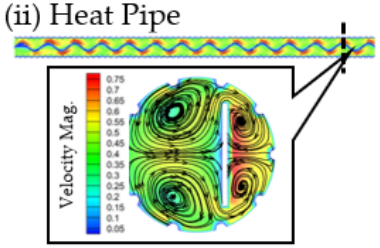
Down Lifting Transformation → Solution (u)

Example of Solutions from the Compact FM

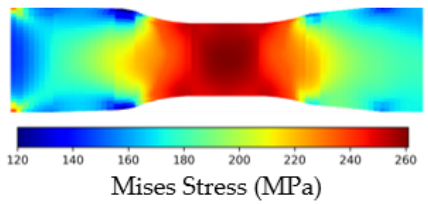
(i) Coolant channel



(ii) Heat Pipe



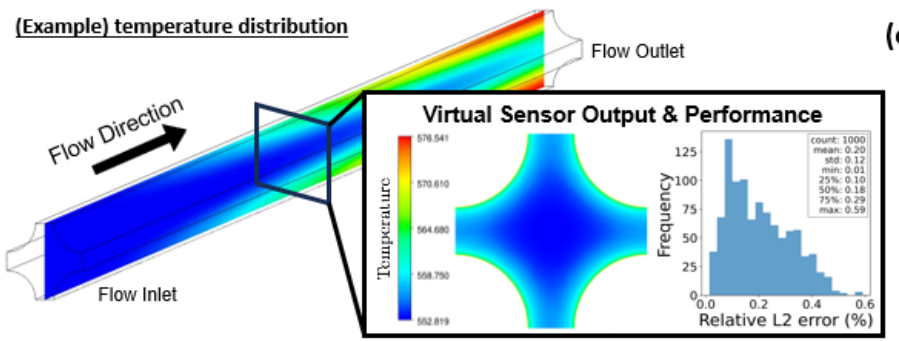
(iii) Structure & Material



(a) Virtual Sensor on a Cooling System

u_1 : inlet temperature
u_2 : inlet velocity
↦
f_1 : temperature distribution
f_2 : velocity distribution

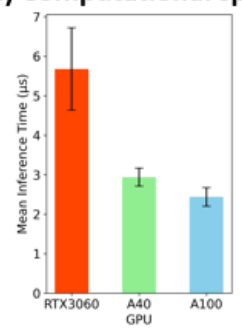
(Example) temperature distribution



(b) Model Performance on unseen dataset

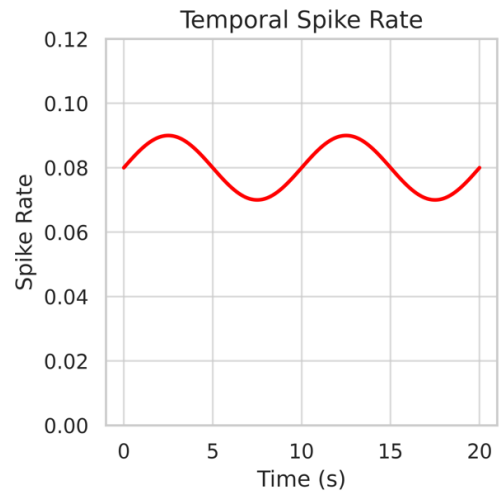
Metrics	5-fold validation MSE	Test dataset relative L_2 errors (%)	
		f_1	f_2
Value	$(4.32 \pm 0.81) \times 10^{-3}$	$(2.02 \pm 1.16) \times 10^{-1}$	$(8.43 \pm 5.53) \times 10^{-1}$

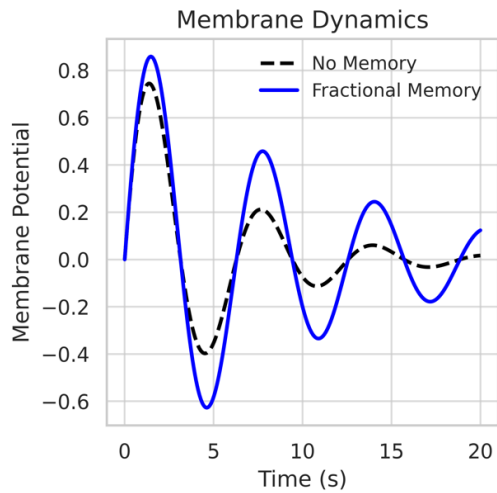
(c) Computational speed comparison

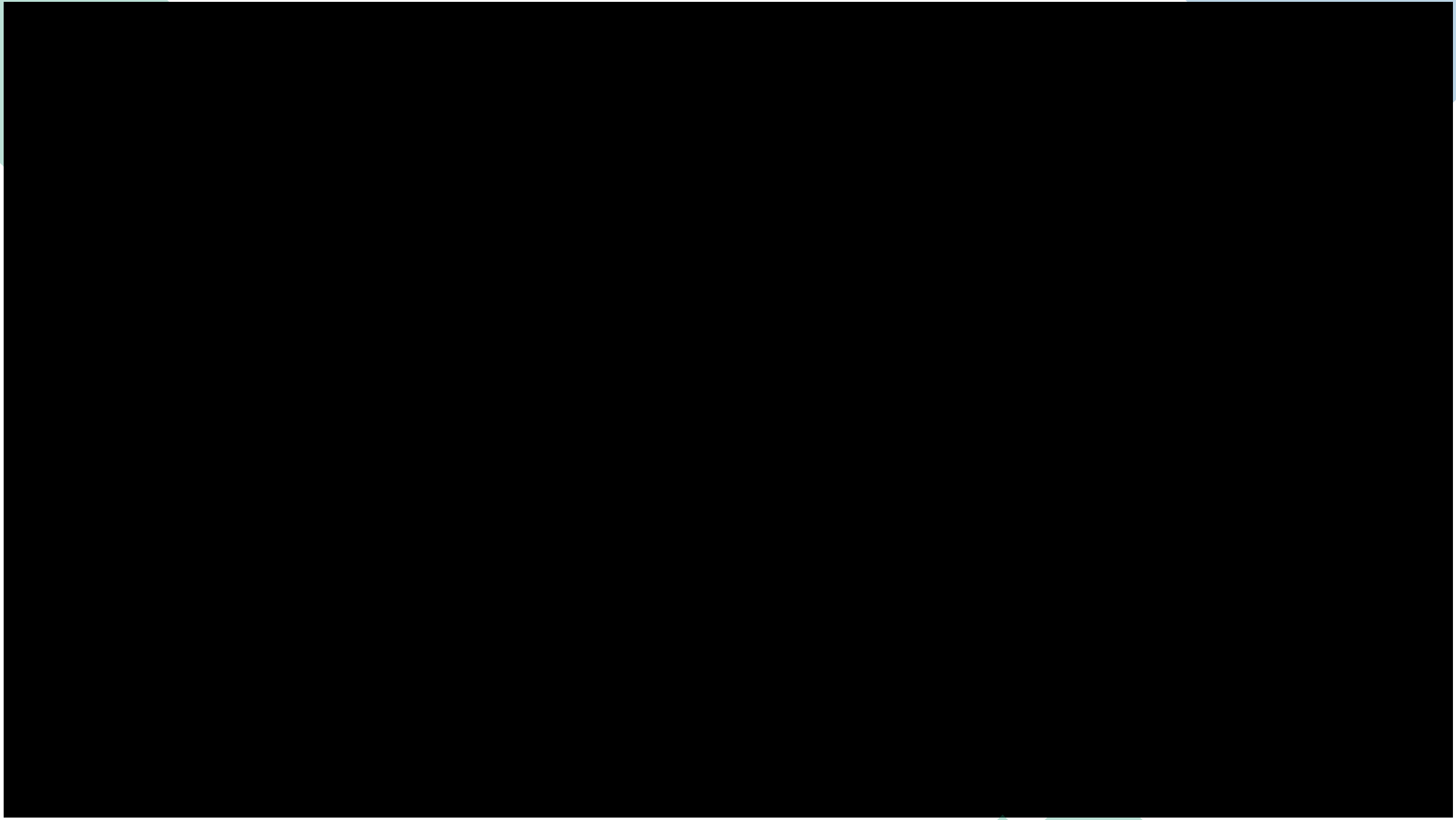


Simulation with CPU (12 core)
- ~ 580 seconds

AI Model inference with GPU
- ~ Microsecond order
- Does not need retraining for new scenario







Thank you

Syed Bahauddin Alam

alams@illinois.edu

MARTIANS LAB

University of Illinois Urbana-Champaign

SCAN TO VISIT OUR LAB

martianslab.web.illinois.edu