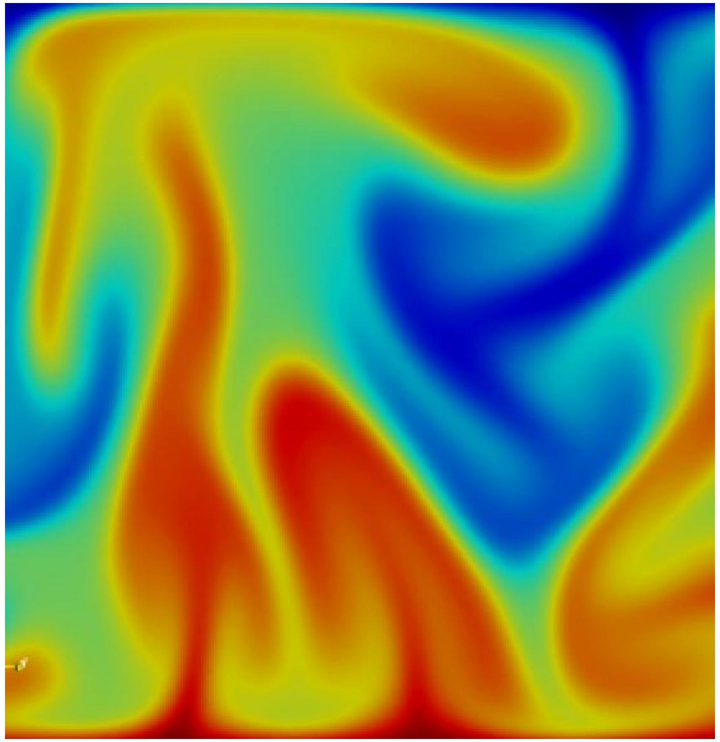
UNIVERSITÄT BIELEFELD

Faculty of Technology

SAIL

SUSTAINABLE LIFE-CYCLE OF INTELLIGENT SOCIO-TECHNICAL SYSTEMS

Surrogate modeling: Accelerating Scientific Discovery with Machine Learning



Michiel Straat

Machine Learning Group, Bielefeld University

Lamarr Lab Visits 2025

18.02.2025

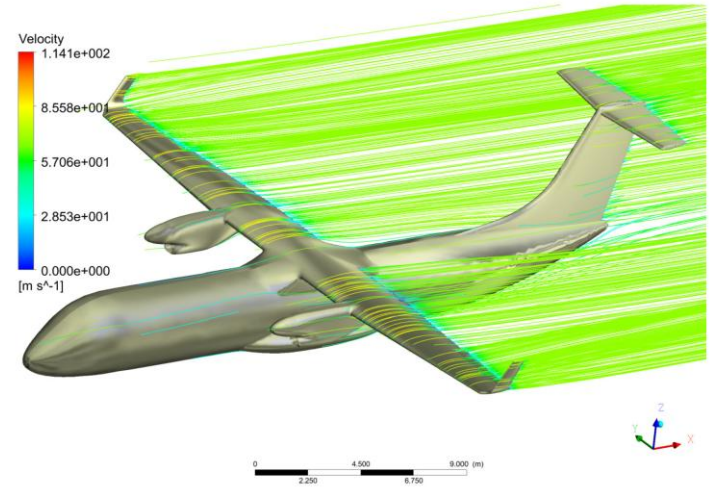
Outline of the session

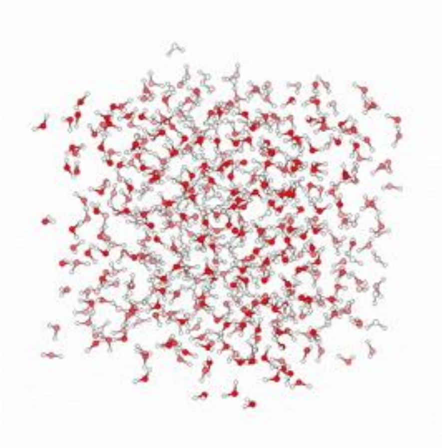
Discussing Machine Learning in physics applications and using physics to improve Machine Learning

- **Talks (14:30-15:30)**
  - Surrogate modeling using Machine Learning (Michiel Straat)
  - Physics-informed machine learning (Thorben Markmann)
  - Application of physics-informed ML in Water Distribution Networks (André Artelt)
  - Symmetry-Informed Reinforcement Learning (Sebastian Peitz)
- **Time for discussion (15:30-)**

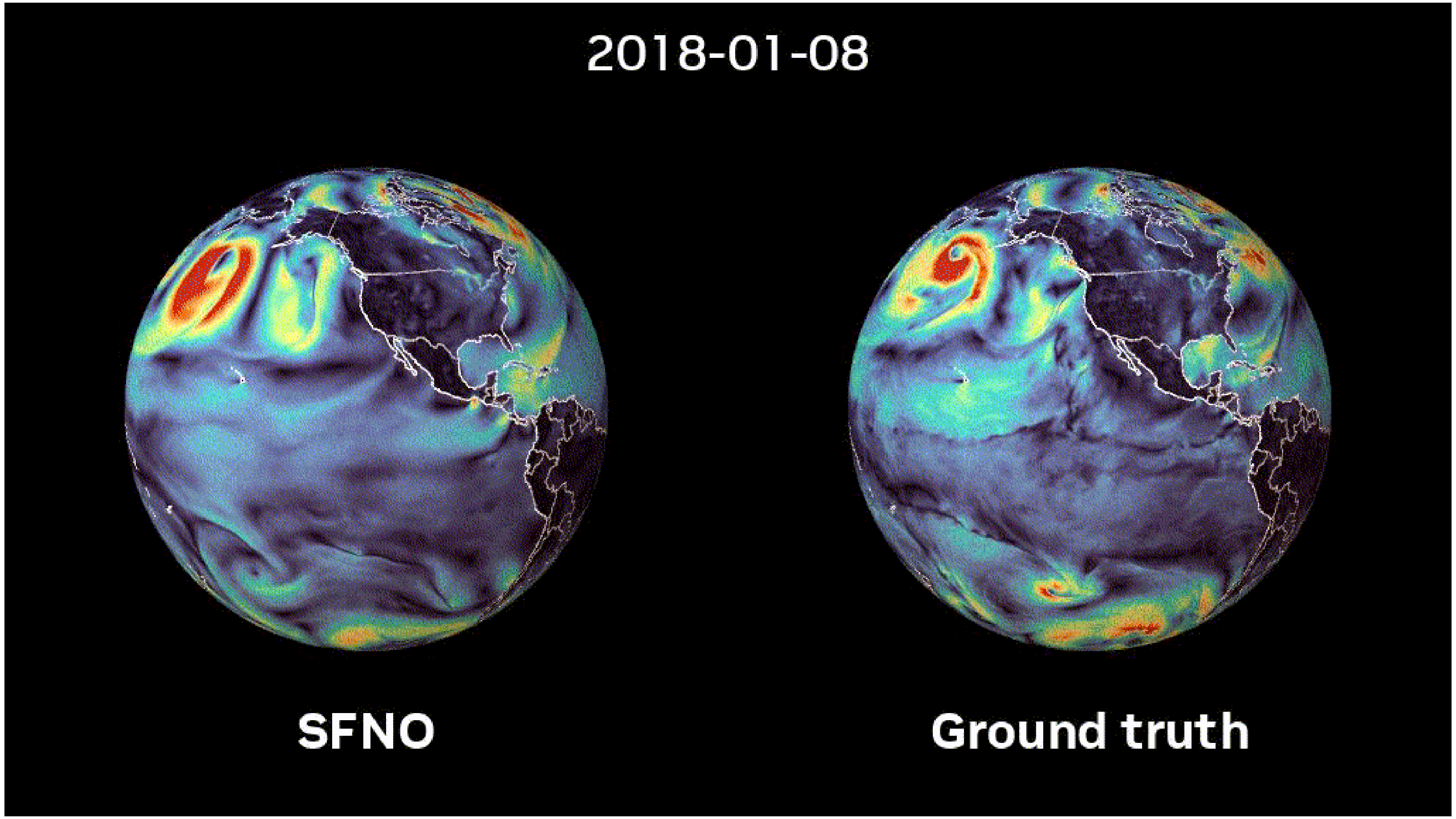
Introduction to Surrogate Modeling

- A surrogate model approximates a more complex or expensive model, and it is used to accelerate simulations, optimization, and decision-making.
- Often used when direct simulations or experiments are too costly
  - Reinforcement learning environments
  - Parameter studies and inverse problems.
- Examples and application areas:
  - Climate science: Estimating climate patterns based on computationally intensive weather simulations.
  - Engineering: Predicting the behavior of an aircraft wing under different conditions without running full CFD simulations.
  - Material science: Predicting material properties without running full molecular dynamics simulations.





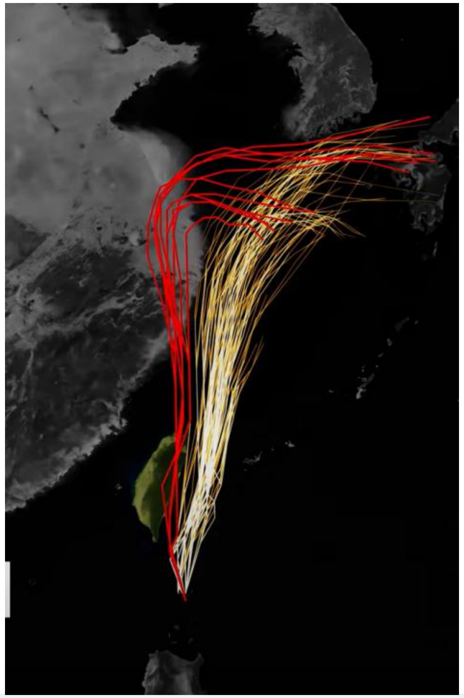
Weather prediction

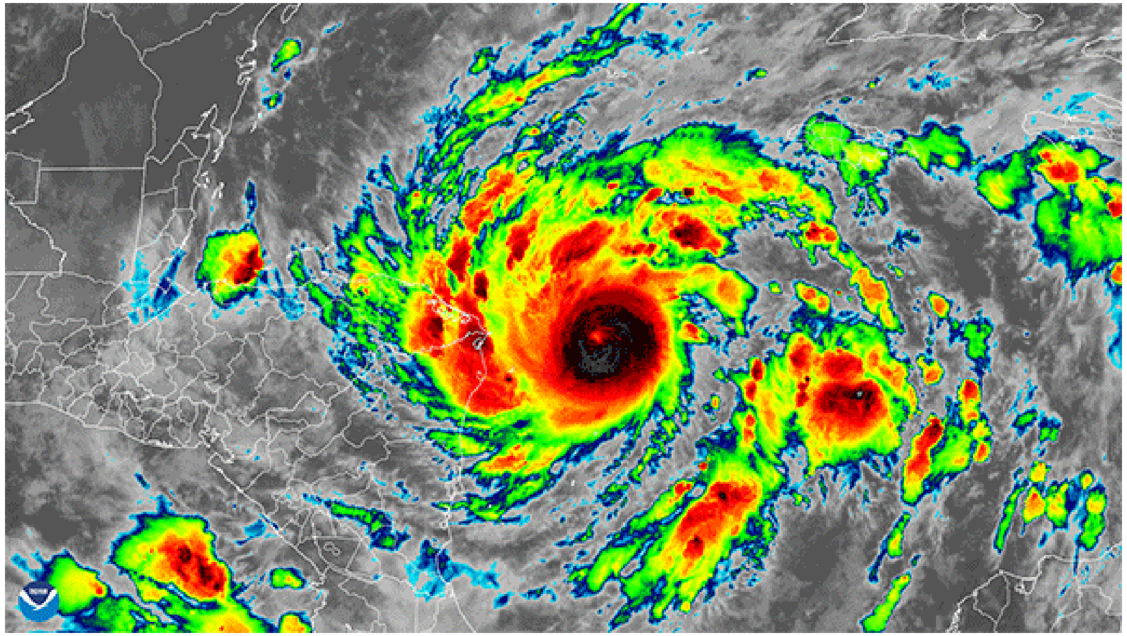


1-month-long rollout of SFNO. Surface windspeed predictions with SFNO and ground truth data are compared to each other, trained on the ERA5 dataset.
1 year rollout (1460 steps), takes about 13 minutes

Ensemble Weather Forecast

- simulations with many slightly different initial conditions and model details (*parameter studies*)
  - Gives possible developments (e.g. different tracks of a storm) and uncertainty quantification.





https://developer.nvidia.com/blog/ai-accurately-forecasts-extreme-weather-up-to-23-days-ahead/

Application area: Inverse problems

- Determine unknown inputs from observed outputs.
- Examples are in engineering: airfoil optimization, ocean current prediction from sparse temperature measurements.
- Surrogates help in solving the problem by:
  - Rapid forward evaluations: $y = f(\theta) \approx s(\theta)$
  - Approximating the inverse mapping directly: $\theta = f^{-1}(y) \approx s(\theta)$
  - (Combine both approaches).

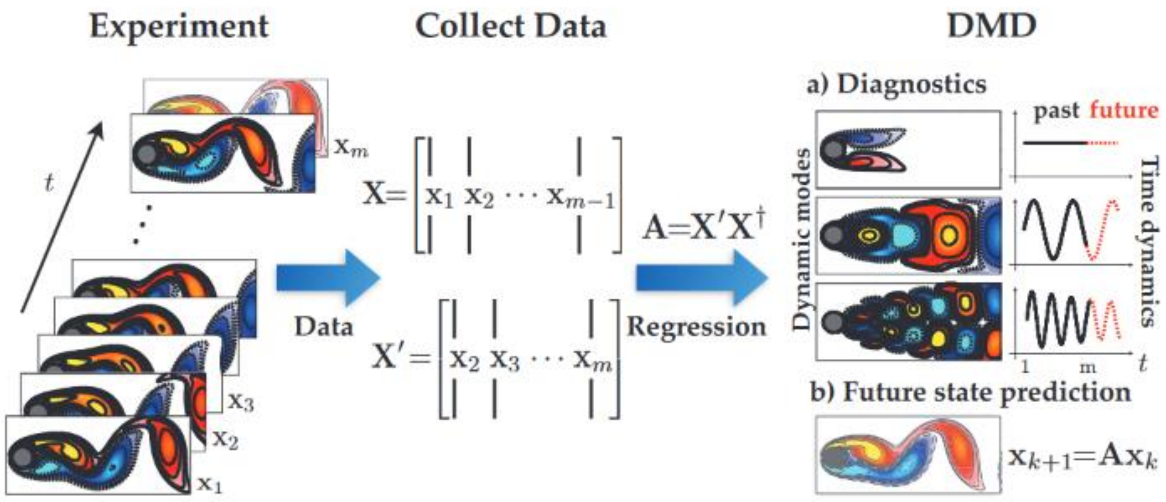
Machine Learning and Surrogate Modeling

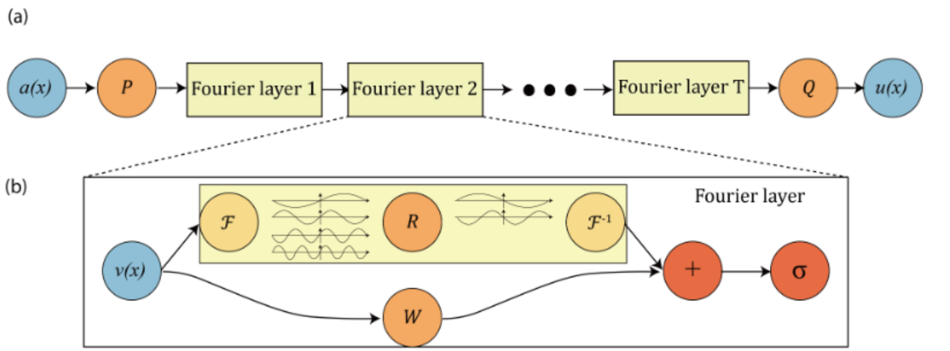
- Traditional models: polynomial approximations, kriging, Reduced Order Models (ROMs).
- Big-data and compute era: ML-based models, such as (Convolutional) Neural Networks, Gaussian Processes, Neural Operators.
- ML techniques can capture complex, high-dimensional dependencies that traditional methods struggle with.
- After the computational effort of training, the models provide very fast predictions.
- Surrogate models can be fully data-driven or include physics to accelerate training, improve generalization and alleviate data requirements.

Surrogate modeling in Reinforcement Learning

- Many RL environments rely on expensive simulations (e.g. robotics, autonomous systems, finance).
- Model-Based Reinforcement Learning
  - Surrogate models (a.k.a. World Models) make training more efficient by approximating system dynamics.

Case study: ML surrogate modeling in fluid dynamics

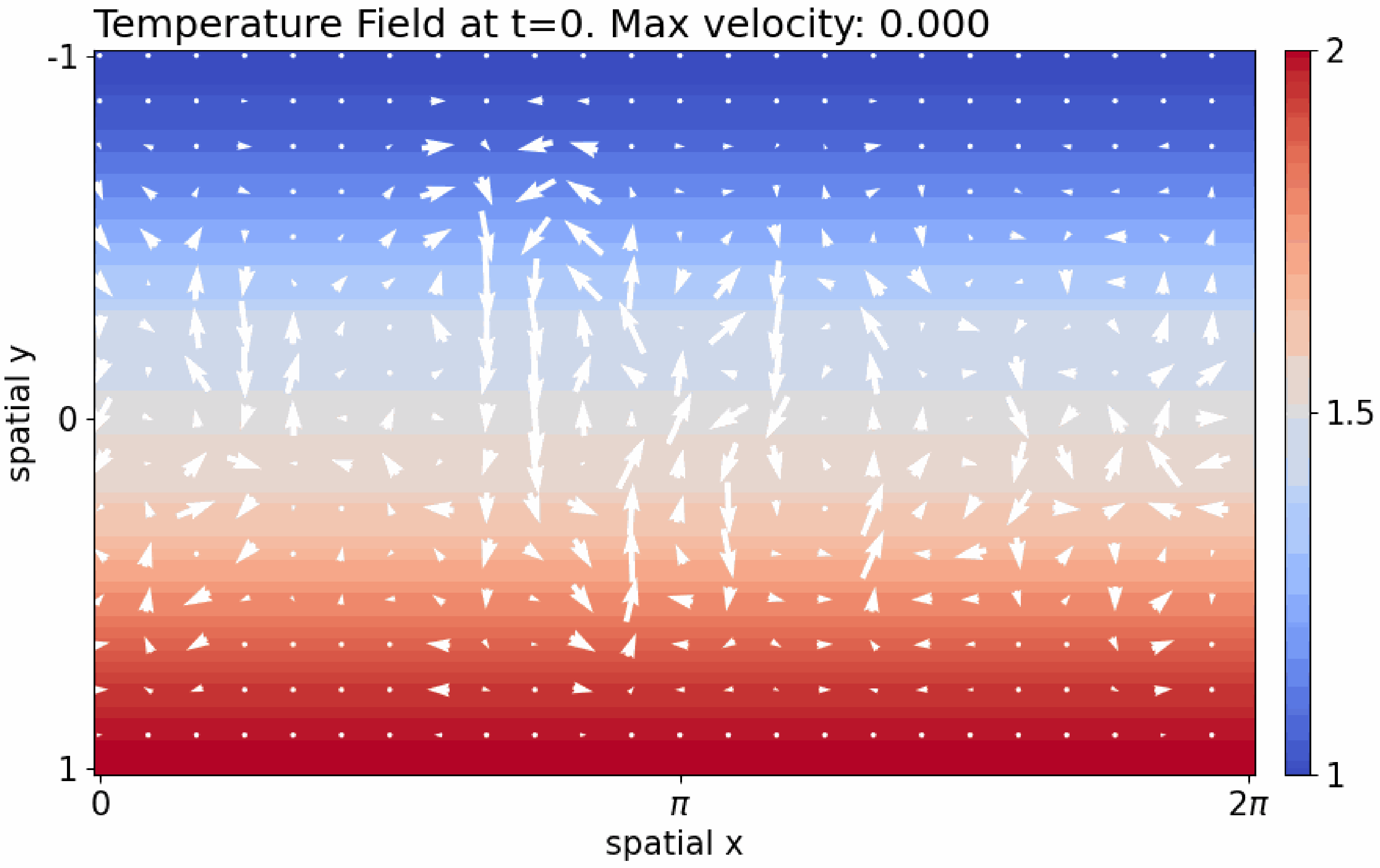




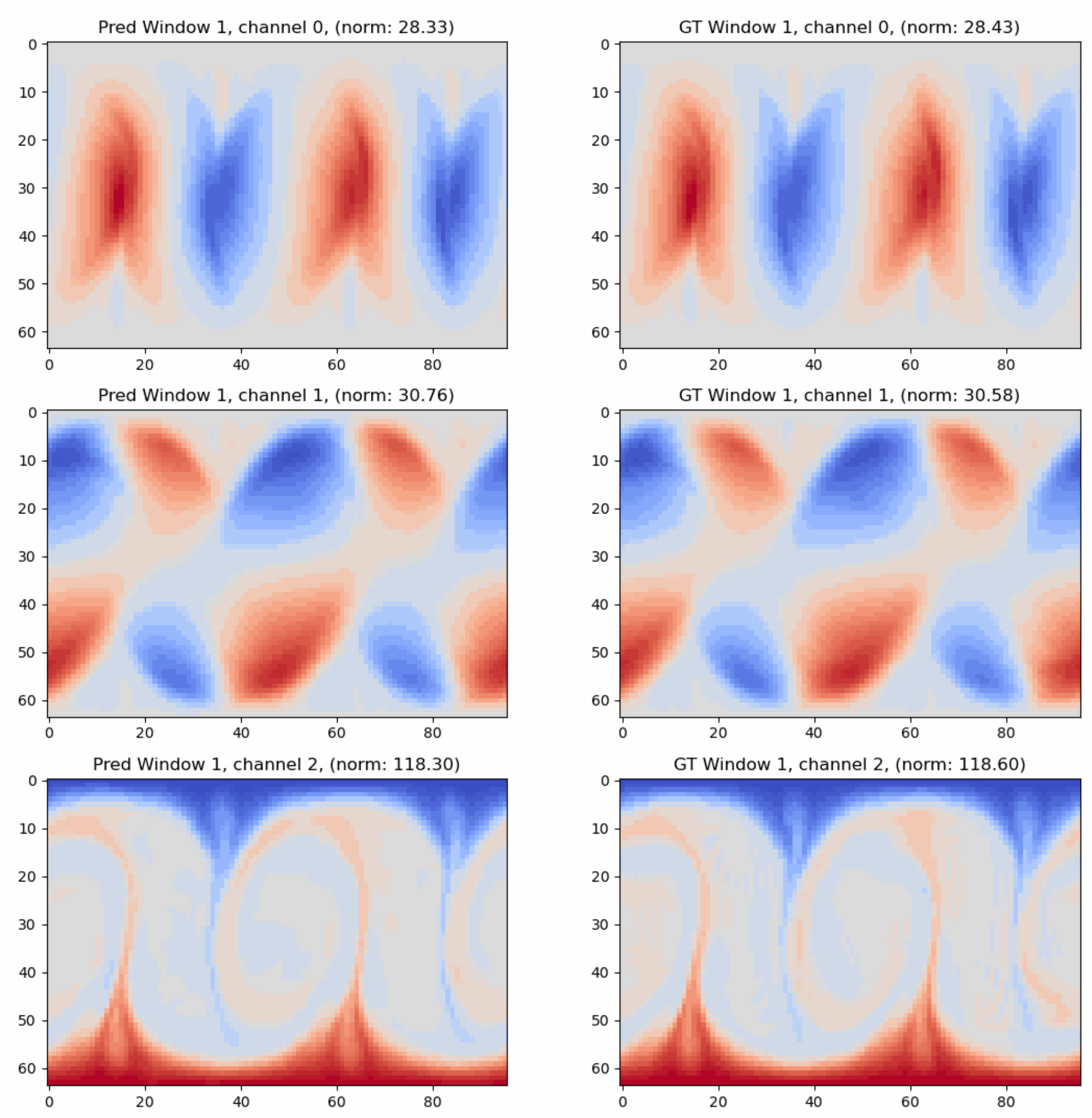
(a) The full architecture of neural operator: start from input a. 1. Lift to a higher dimension channel space by a neural network P. 2. Apply four layers of integral operators and activation functions. 3. Project back to the target dimension by a neural network Q. Output u. **(b) Fourier layers**: Start from input v. On top: apply the Fourier transform $\mathcal{F}$; a linear transform R on the lower Fourier modes and filters out the higher modes; then apply the inverse Fourier transform $\mathcal{F}^{-1}$. On the bottom: apply a local linear transform W.

Figure 2: **top:** The architecture of the neural operators; **bottom:** Fourier layer.

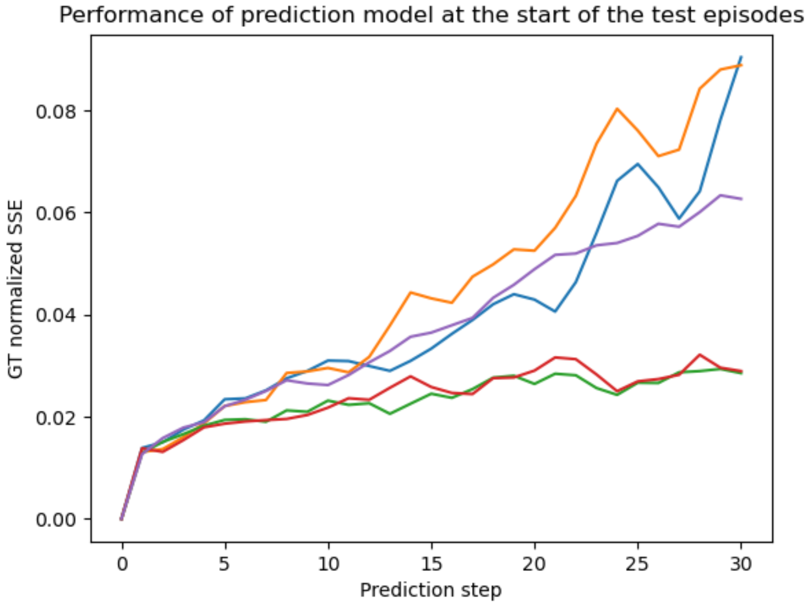
Predicting convection patterns



Results FNO $Ra = 2 * 10^6$



Parameter	Value	Parameter	Value
Hidden channels	32	Nr Episodes	20
Lifting channels	64	Projection Channels	64
Nr Layers	6	Seq. Delay	0.7
Nr Modes x	16	Seq. Length	2
Nr Modes y	16		



Challenges in Surrogate Modeling

- Generalization beyond the training data.
- Integrate physical constraints and domain knowledge into ML models (increase generalization and data efficiency)
  - Physics Informed Neural Networks
  - Hybrid AI-Simulation: e.g. in computation fluid dynamics, AI-based surrogate model predicts turbulence behavior while the rest of the simulation is handled by classical numerical solvers.
- Transfer learning: Encourage re-use of surrogate models for slight changes in tasks or domains, fine-tune to problem specifics using small datasets. Sim-to-Real Transfer.

Summary

- Surrogate models enable faster and more efficient simulations.
- Machine learning-based surrogates have shown impressive results for high-dimensional problems.
- Open question: How to best combine physics, machine learning and domain expertise?

UNIVERSITÄT BIELEFELD
Faculty of Technology

SAIL
SUSTAINABLE LIFE-CYCLE OF INTELLIGENT SOCIO-TECHNICAL SYSTEMS

Thank you for your attention!

Contact:

Michiel Straat

mstraat@techfak.uni-bielefeld.de

SAIL is funded by

Ministry of Culture and Science
of the State of
North Rhine-Westphalia

under the grant no NW21-059A