

Performance-Carbon Trade-Offs across Architectural Biases in Shear Flow Forecasting

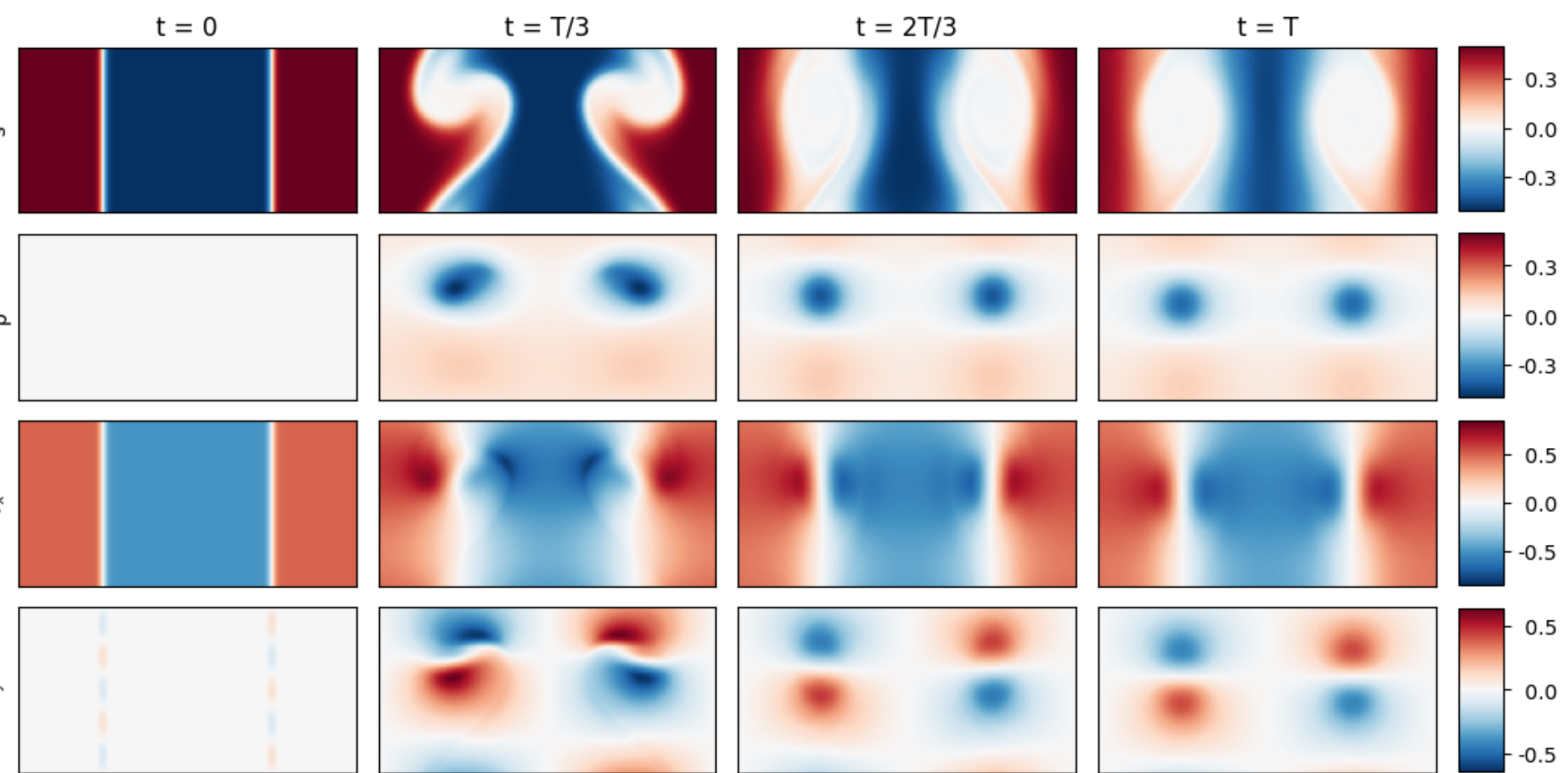
Sophia N. Wilson<sup>1</sup> · Jens Hesselbjerg Christensen<sup>2</sup> · Raghavendra Selvan<sup>1</sup>  
<sup>1</sup>Department of Computer Science, University of Copenhagen, Denmark · <sup>2</sup>Niels Bohr Institute, University of Copenhagen, Denmark

INTRODUCTION

**Can architectural biases improve the trade-off between predictive performance and carbon footprint?**

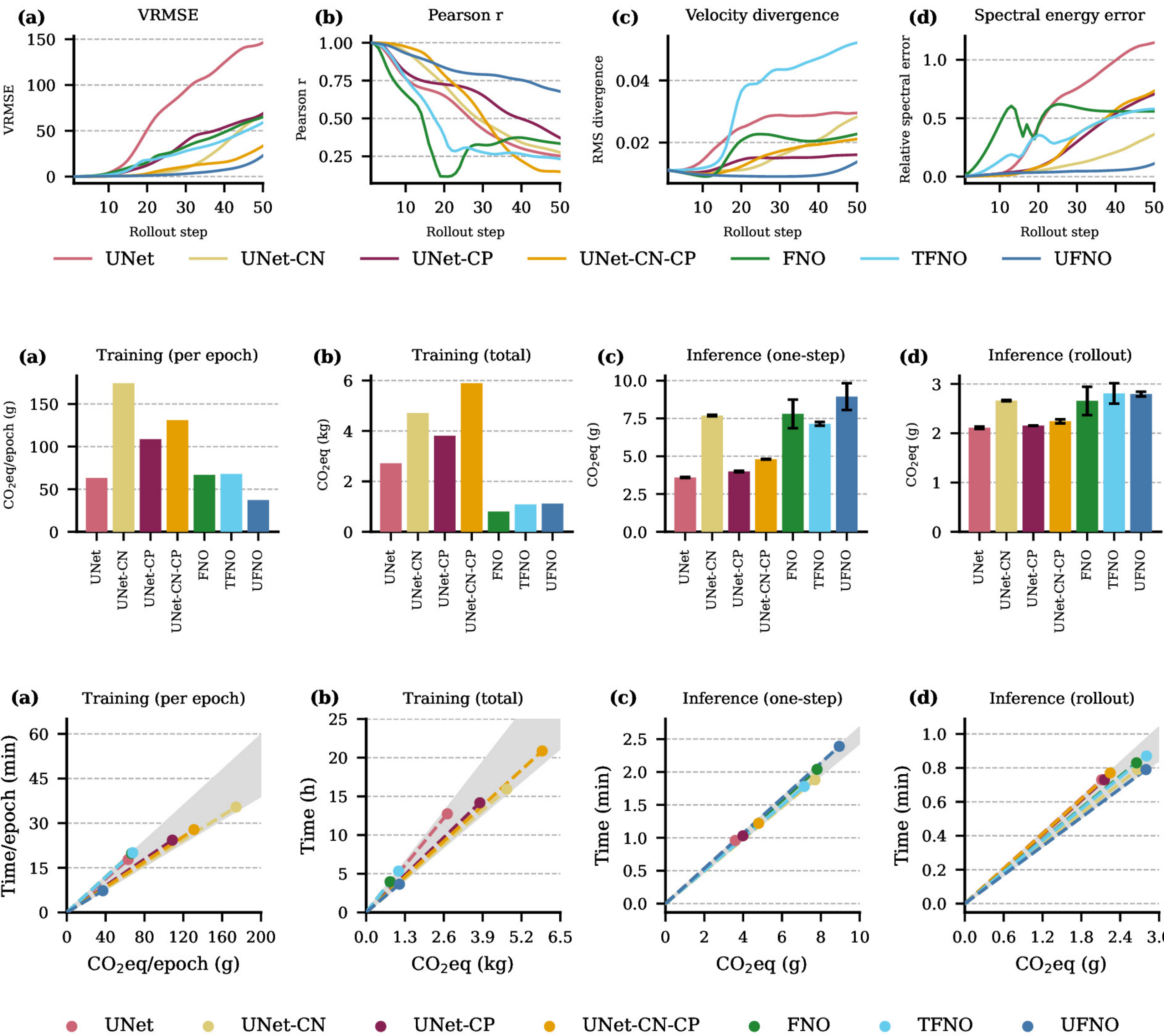
We study seven models varying in receptive field and periodicity, evaluating accuracy, physics fidelity, and carbon cost across training and inference. Our contributions are:

- 1. Carbon-aware evaluation:** we argue for evaluation practices that weigh training and inference carbon cost alongside predictive accuracy and physics fidelity.
- 2. Characterising the performance-carbon trade-off:** we empirically show how receptive field and periodicity biases relate to predictive performance and carbon footprint in spatio-temporal forecasting of incompressible shear flow.



The dataset [1] comprises 240 trajectories of 2D incompressible Navier-Stokes flow, spanning six PDE parameters and 40 initial conditions. Each trajectory has four fields: two velocity components, pressure, and a tracer.

PERFORMANCE-CARBON TRADE-OFFS



We evaluate seven models from two model families, UNet [2,3] and Fourier Neural Operator (FNO) [4-6], which differ in receptive field and periodicity assumptions, as summarised in the table below.

| Model               | Receptive field | Periodicity                 |
|---------------------|-----------------|-----------------------------|
| UNet, UNet-CN       | Local           | None (zero-padded)          |
| UNet-CP, UNet-CN-CP | Local           | Explicit (circular padding) |
| FNO, TFNO, UFNO     | Global          | Implicit (Fourier basis)    |

The UNets have a local receptive field due to their convolutional layers, whereas the FNOs have a global receptive field because they operate in Fourier space.

Two UNets make no periodicity assumption, while the other two impose an explicit periodic assumption via circular padding at the domain boundary. All FNOs carry an implicit periodicity assumption throughout their representation, a consequence of the discrete Fourier transform assuming a periodic signal.

All models are trained for one-step forecasting. We measure the energy consumption during training and inference using CarbonTracker and convert this into a carbon footprint, as seen in the table below.

| Model      | Training                     |                               |         |          | Inference (one-step)   |          | Inference (rollout)    |          |
|------------|------------------------------|-------------------------------|---------|----------|------------------------|----------|------------------------|----------|
|            | CO <sub>2</sub> eq/epoch (g) | CO <sub>2</sub> eq total (kg) | #Epochs | Time (h) | CO <sub>2</sub> eq (g) | Time (s) | CO <sub>2</sub> eq (g) | Time (s) |
| UNet       | 63                           | 2.72                          | 43      | 12.7     | 3.59 ± 0.03            | 58       | 2.11 ± 0.02            | 44       |
| UNet-CN    | 174                          | 4.71                          | 27      | 15.9     | 7.70 ± 0.05            | 113      | 2.66 ± 0.01            | 47       |
| UNet-CP    | 109                          | 3.81                          | 35      | 14.2     | 3.99 ± 0.05            | 62       | 2.15 ± 0.00            | 44       |
| UNet-CN-CP | 131                          | 5.89                          | 45      | 20.9     | 4.80 ± 0.02            | 73       | 2.24 ± 0.04            | 46       |
| FNO        | 67                           | 0.80                          | 12      | 4.0      | 7.80 ± 0.94            | 122      | 2.66 ± 0.29            | 50       |
| TFNO       | 68                           | 1.09                          | 16      | 5.3      | 7.14 ± 0.12            | 107      | 2.81 ± 0.21            | 52       |
| UFNO       | 37                           | 1.12                          | 30      | 3.6      | 8.94 ± 0.88            | 143      | 2.79 ± 0.05            | 48       |

CONCLUSION

Our study explores how architectural biases relate to the trade-off between predictive performance and carbon footprint. As the plots above show, the answer is not straightforward, as carbon trade-offs vary across training and inference. **This motivates the need to consider resource costs across the entire model lifecycle.**

The plots also show that runtime is not a reliable proxy for carbon footprint, which is why **resource costs must be measured explicitly.**

We focus on architectural biases, but the logic of **incorporating domain knowledge, where possible**, extends to other forms of prior knowledge.

**Ultimately, we argue for evaluating progress in machine learning in terms of value gained relative to cost incurred.**

LINKS



[1] Ohana, R., et al. (2024): The Well: A Large-scale Collection of Diverse Physics Simulations for Machine Learning. NeurIPS.  
 [2] Ronneberger, O., Fischer, P., Brox, T. (2025): U-Net: Convolutional Networks for Biomedical Image Segmentation. MICCAI.  
 [3] Liu, Z., et al. (2022): A ConvNet for the 2020s. CVPR.  
 [4] Li, Z., et al. (2021): Fourier Neural Operator for Parametric Partial Differential Equations. ICLR.  
 [5] Kossaifi, J., et al. (2023): Multi-grid Tensorized Fourier Neural Operator for High-resolution PDEs. TMLR.  
 [6] Rahman, M. A., et al.: U-NO: U-shaped Neural Operators. TMLR.

