CENTRE EUROPÉEN DE RECHERCHE ET DE FORMATION AVANCÉE EN CALCUL SCIENTIFIQUE

Scientific Foundation Models for Computational Fluid Dynamics: threats and opportunities

Fernando González, Luciano Drozda

Sep 30, 2024

Foundation model

Definition

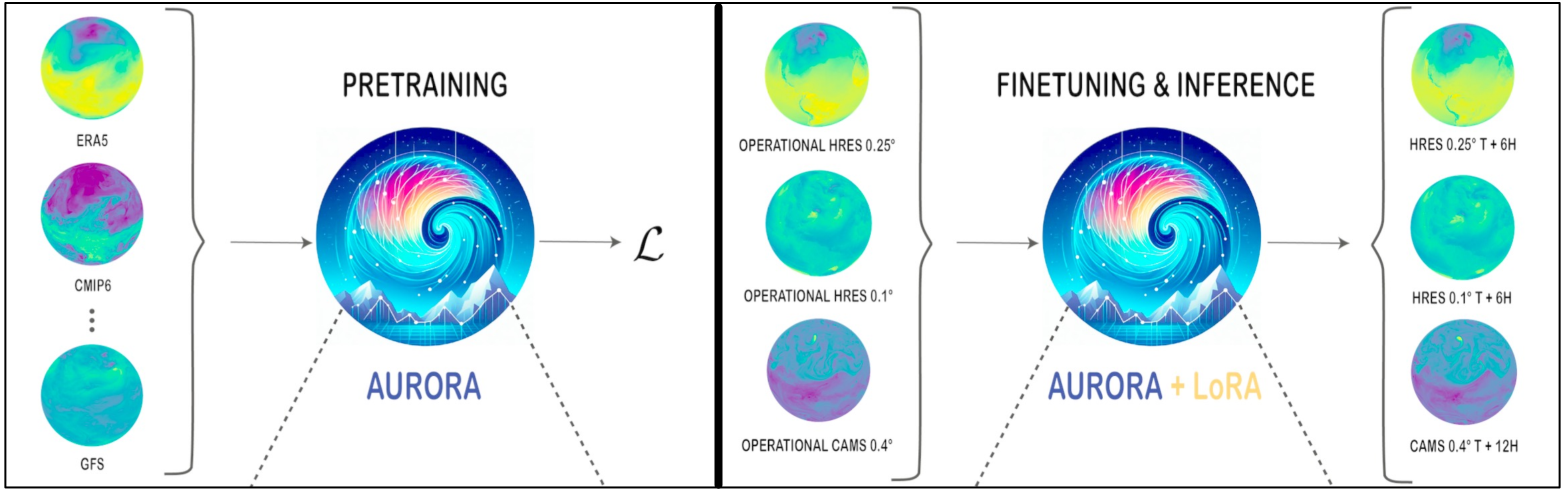
A large-scale machine learning model pretrained on a large body of data that can subsequently be adapted to solve multiple downstream tasks using some form of finetuning on a few task-specific samples

Bishop, C. M., & Bishop, H. (2024). *Deep learning: Foundations and concepts*. Cham: Springer International Publishing.

Foundation model

Example

A large-scale machine learning model <u>pretrained</u> on a large body of data that can subsequently be adapted to solve multiple downstream tasks using some form of <u>finetuning</u> on a few task-specific samples

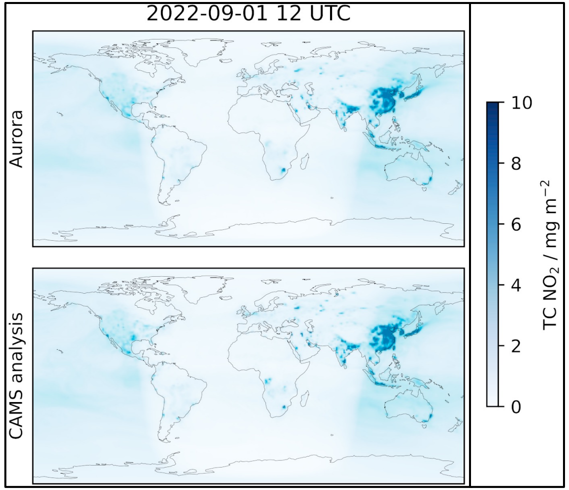


Bodnar, C., Bruinsma, W. P., Lucic, A., Stanley, M., Brandstetter, J., Garvan, P., … Perdikaris, P. (2024, May 28). *Aurora: A foundation model of the atmosphere*. arXiv.

Foundation model

Example

A large-scale machine learning model <u>pretrained</u> on a large body of data that can subsequently be adapted to solve multiple downstream tasks using some form of <u>finetuning</u> on a few task-specific samples



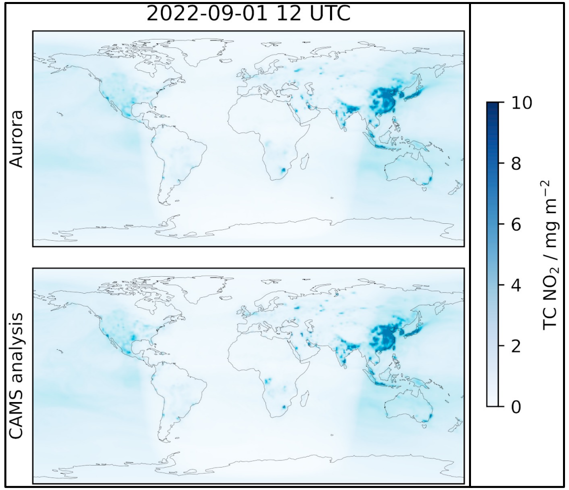
CERFACS

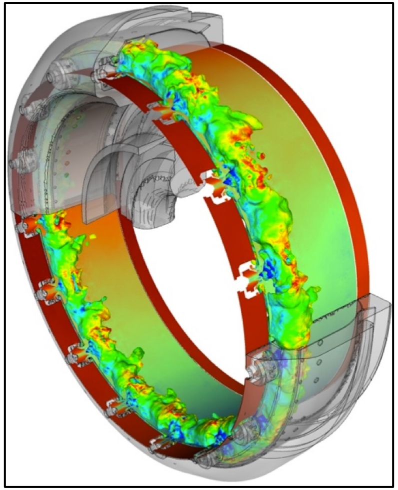
Bodnar, C., Bruinsma, W. P., Lucic, A., Stanley, M., Brandstetter, J., Garvan, P., … Perdikaris, P. (2024, May 28). *Aurora: A foundation model of the atmosphere*. arXiv.

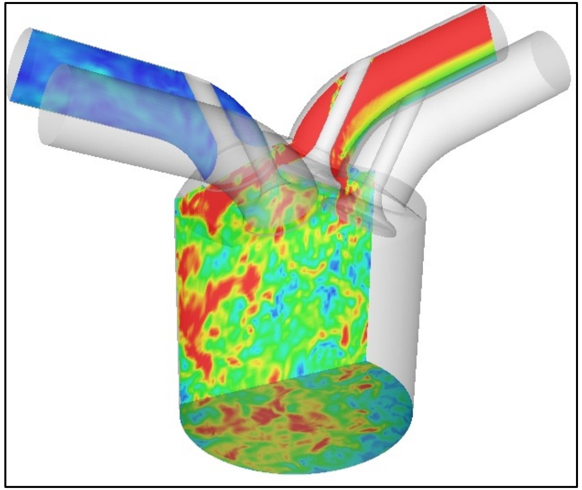
Threats

Data scaling

A large-scale machine learning model pretrained on a large body of data that can subsequently be adapted to solve multiple downstream tasks using some form of finetuning on a few task-specific samples







SCHÖNFELD, T., AND RUDGYARD, M. Steady and unsteady flows simulations using the hybrid flow solver AVBP.

Threats

Data secrecy

netCDF

NASA

Met Office

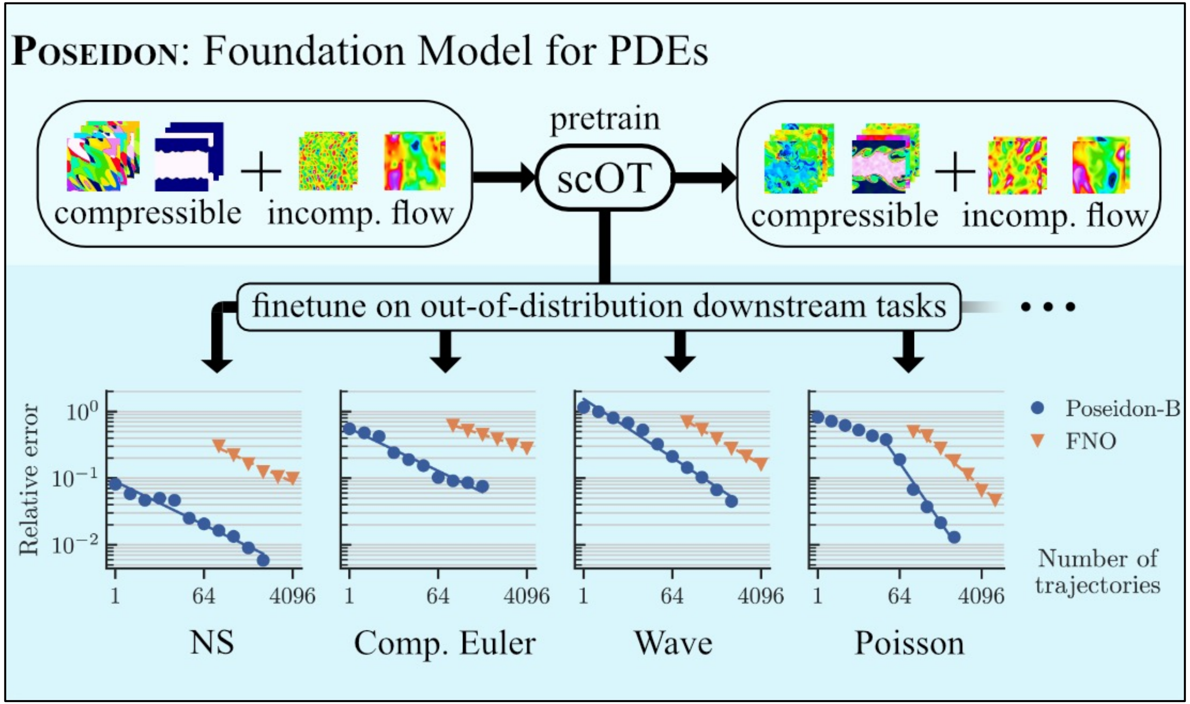
ECMWF

SAFRAN

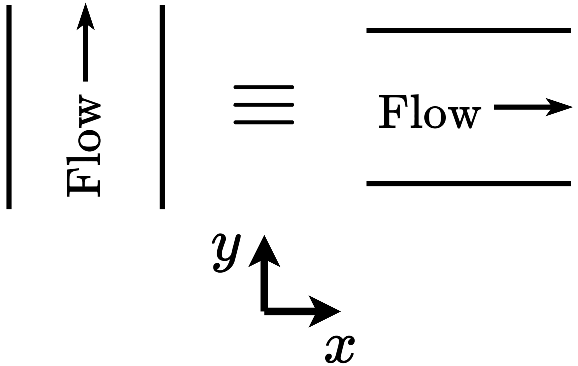
AIRBUS

Threats

Modeling assumptions



Lack of hard constraints on physical laws and Invariance / Equivariance properties

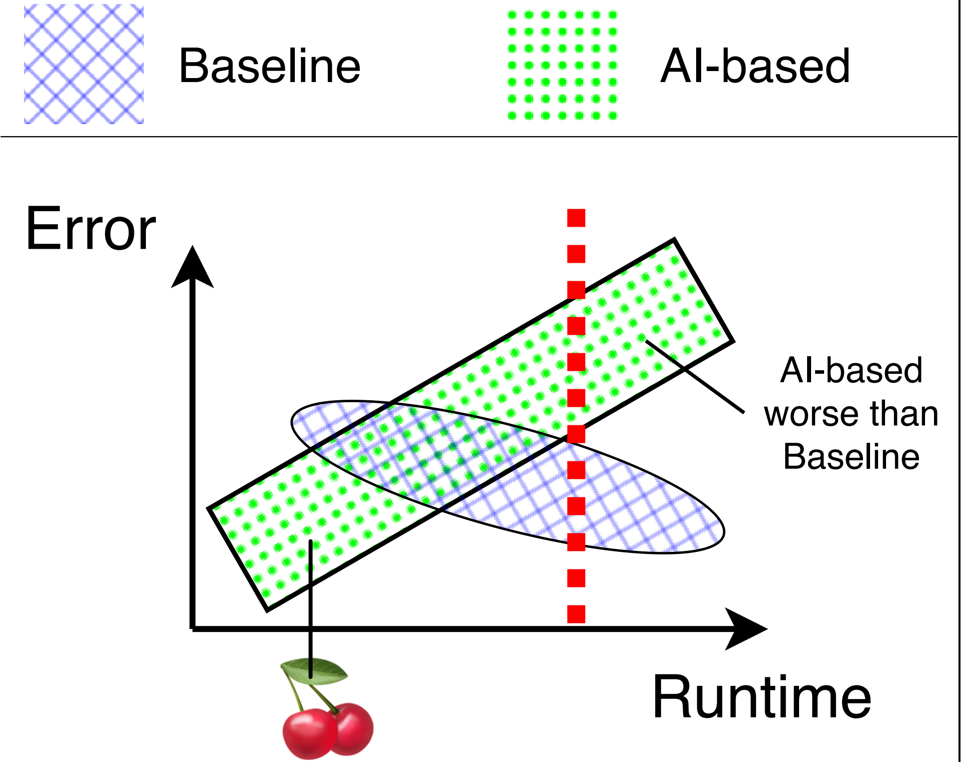


CERFACS

Herde, M., Raonić, B., Rohner, T., Käppeli, R., Molinaro, R., de Bézenac, E., & Mishra, S. (2024). *Poseidon: Efficient foundation models for pdes*. arXiv.

Threats

Weak baselines, outputs reliability, regulations

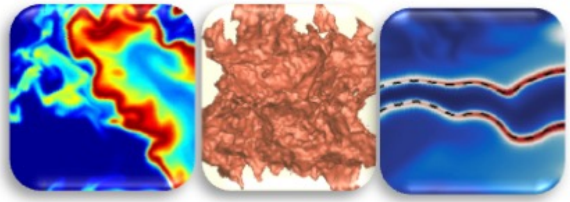


McGreivy, N., & Hakim, A. (2024). Weak baselines and reporting biases lead to overoptimism in machine learning for fluid-related partial differential equations. *Nature Machine Intelligence*. https://doi.org/10.1038/s42256-024-00897-5

Opportunities

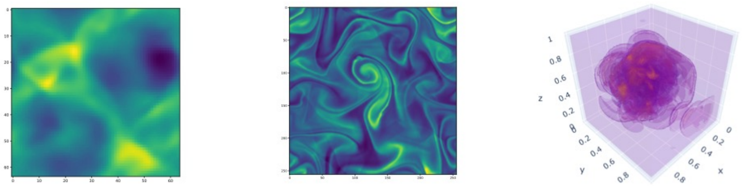
Emergence of Open-data Initiatives

1. BLASTNet

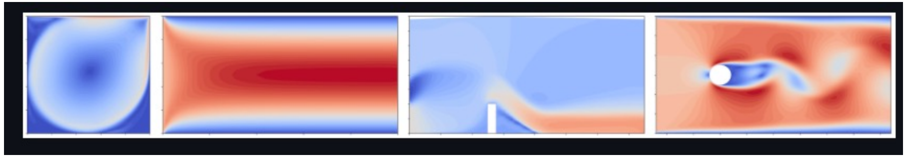


3. AirfRANS

2. PDE-BENCH



4. CFDBench



1. Chung, W.T., Jung, K.S., Chen, J.H., Ihme, M., 2022.

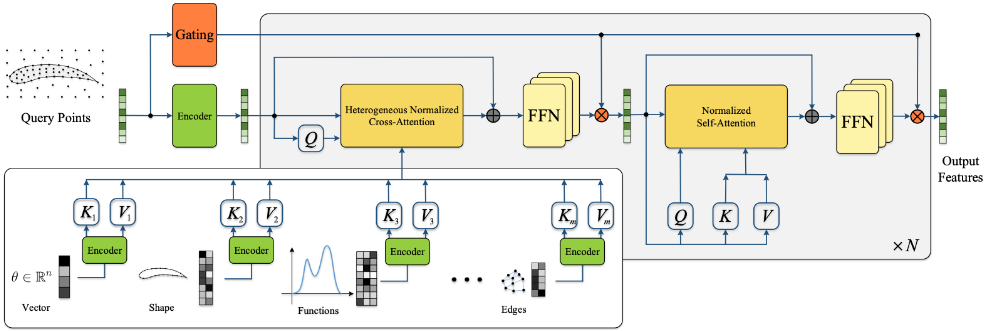
2. Takamoto, M., Praditia, T., Leiteritz, R., MacKinlay, D., Alesiani, F., Pflüger, D., Niepert, M., 2023.

3. Bonnet, F., Mazari, J.A., Cinnella, P., Gallinari, P., 2023.

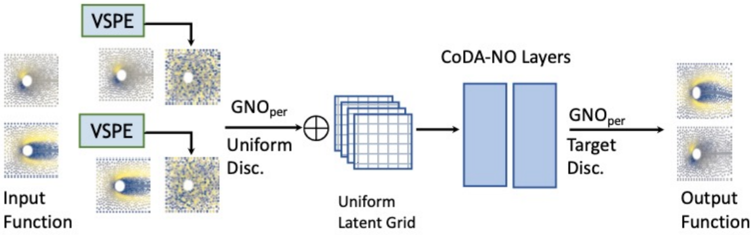
4. Luo, Y., Chen, Y., Zhang, Z., 2024.

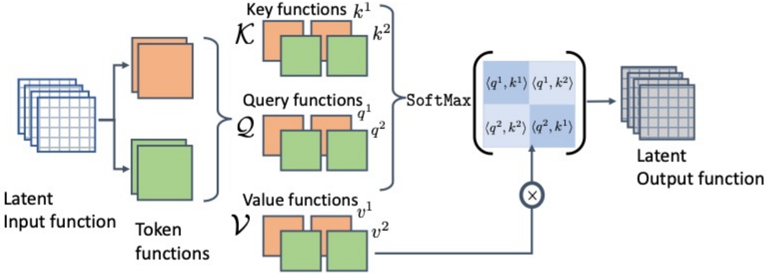
Opportunities

Scalable architectures for SciFMs



- Transformers can be scaled to many parameters and trained on parallel computer architecture.
- Neural Operators can adapt to different discretizations (Mesh independence/Convergence).



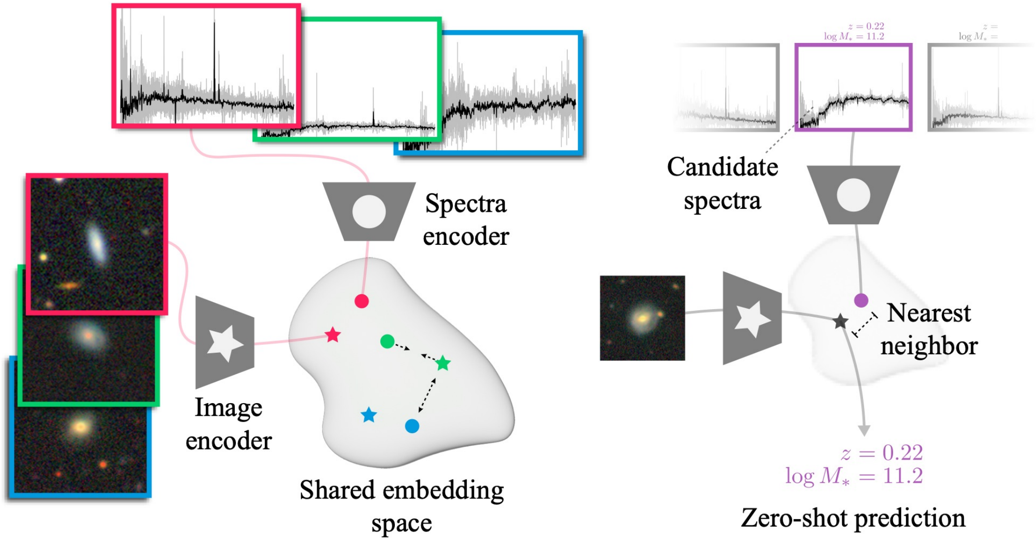


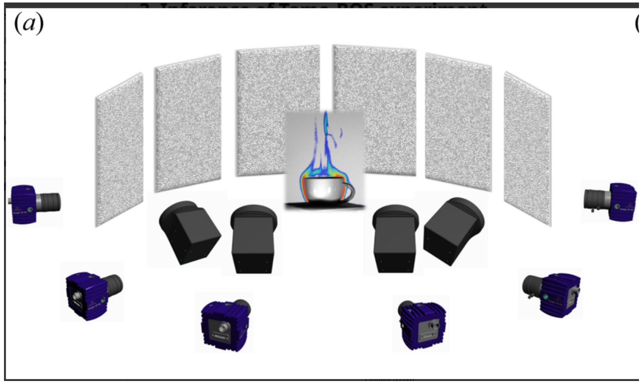
Hao, Z., Wang, Z., Su, H., Ying, C., Dong, Y., Liu, S., Cheng, Z., Song, J., Zhu, J., 2023. GNOT: A General Neural Operator Transformer for Operator Learning, in: Proceedings of the 40th International Conference on Machine Learning. Presented at the International Conference on Machine Learning, PMLR, pp. 12556–12569.

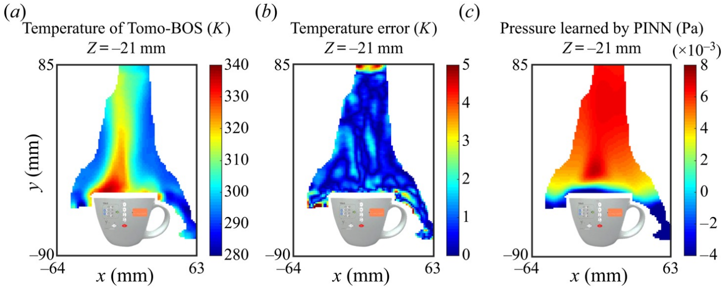
Rahman, M.A., George, R.J., Elleithy, M., Leibovici, D., Li, Z., Bonev, B., White, C., Berner, J., Yeh, R.A., Kossaifi, J., Azizzadenesheli, K., Anandkumar, A., 2024. Pretraining Codomain Attention Neural Operators for Solving Multiphysics PDEs.

Opportunities

Multimodality: Experiments & Simulations





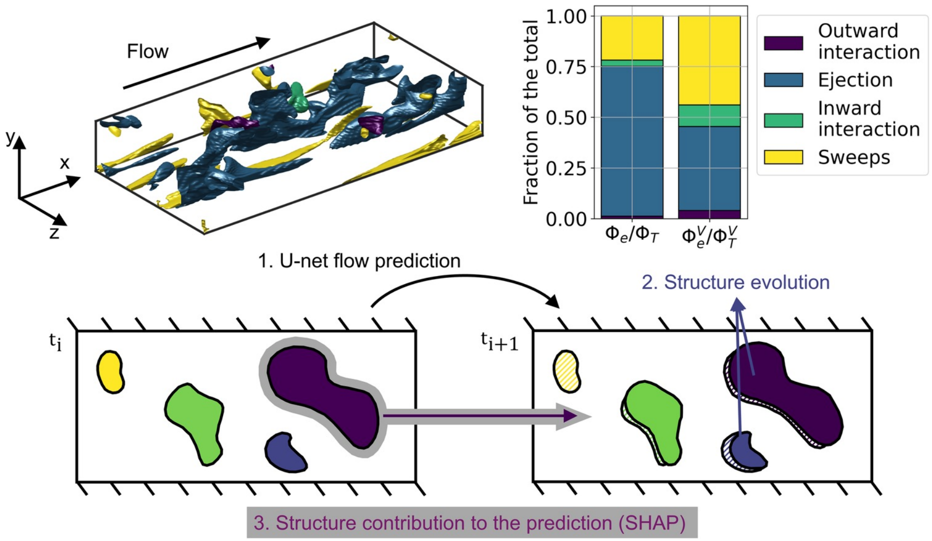


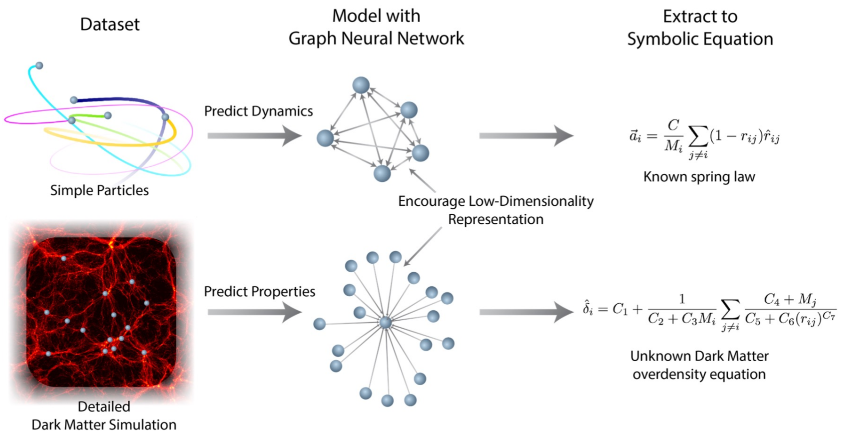
Parker, L., Lanusse, F., Golkar, S., Sarra, L., Cranmer, M., Bietti, A., Eickenberg, M., Krawezik, G., McCabe, M., Ohana, R., Pettee, M., Blancard, B.R.-S., Tesileanu, T., Cho, K., Ho, S., 2024. AstroCLIP: A Cross-Modal Foundation Model for Galaxies. Monthly Notices of the Royal Astronomical Society 531, 4990–5011. https://doi.org/10.1093/mnras/stae1450

Cai, S., Wang, Z., Fuest, F., Jeon, Y.J., Gray, C., Karniadakis, G.E., 2021. Flow over an espresso cup: inferring 3-D velocity and pressure fields from tomographic background oriented Schlieren via physics-informed neural networks. Journal of Fluid Mechanics 915, A102. https://doi.org/10.1017/jfm.2021.135ç

Opportunities

Explainability (xAI) & Scientific Discovery





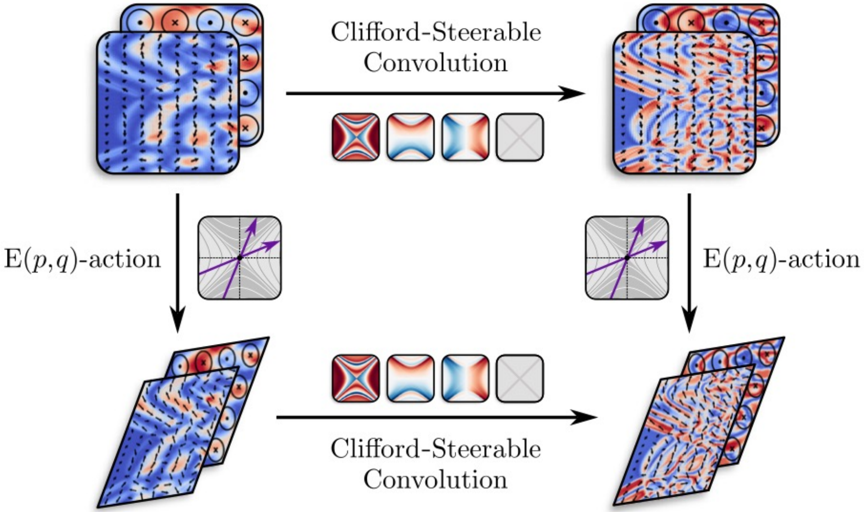
Cremades, A., Hoyas, S., Deshpande, R., Quintero, P., Lellep, M., Lee, W.J., Monty, J.P., Hutchins, N., Linkmann, M., Marusic, I., Vinuesa, R., 2024. Identifying regions of importance in wall-bounded turbulence through explainable deep learning. Nat Commun 15, 3864. https://doi.org/10.1038/s41467-024-47954-6

Cranmer, M., Sanchez-Gonzalez, A., Battaglia, P., Xu, R., Cranmer, K., Spergel, D., Ho, S., 2020. Discovering Symbolic Models from Deep Learning with Inductive Biases. arXiv:2006.11287 [astro-ph, physics:physics, stat].

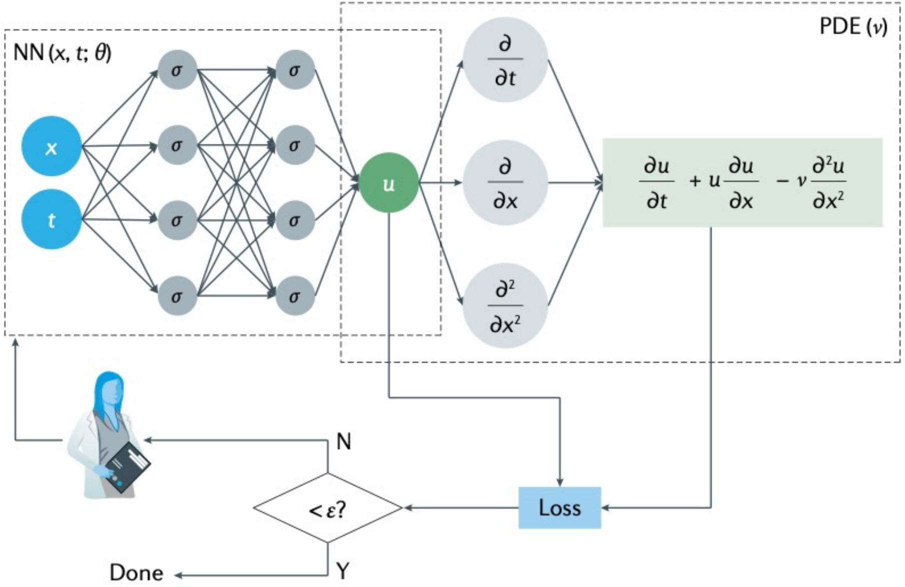
Opportunities

Leveraging Domain Knowledge

Symmetries and Equivariance



Conservation Laws

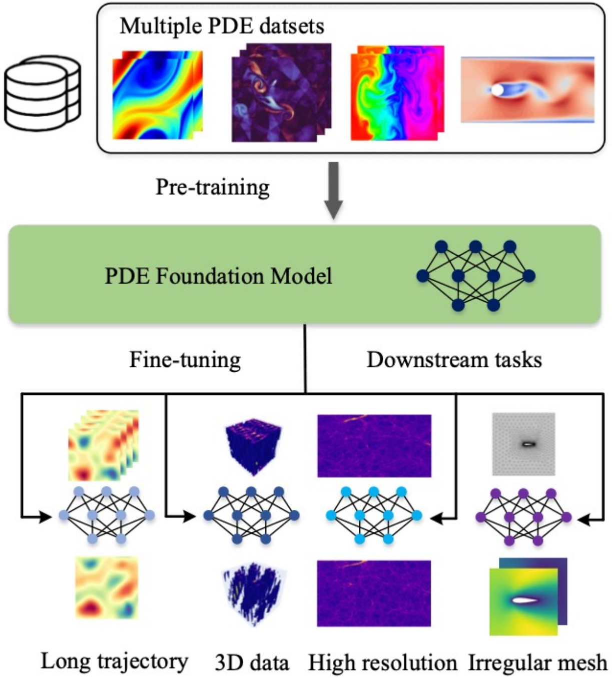


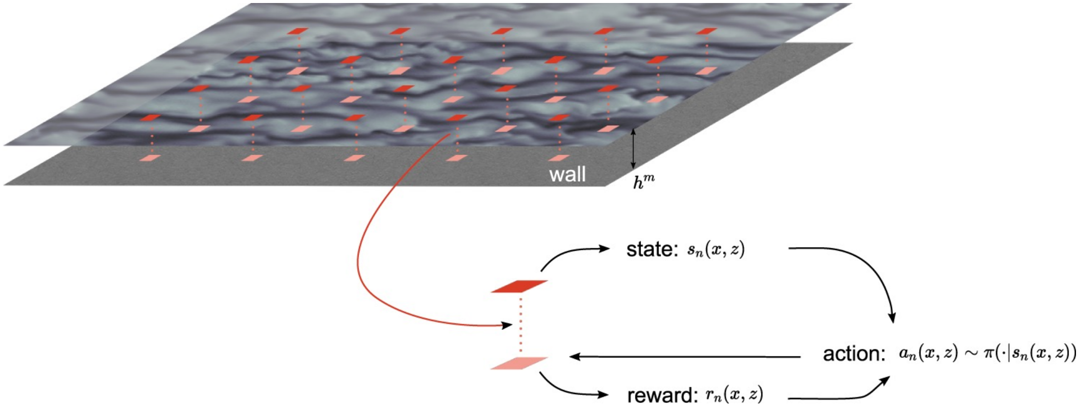
Zhdanov, M., Ruhe, D., Weiler, M., Lucic, A., Brandstetter, J., Forré, P., 2024. Clifford-Steerable Convolutional Neural Networks. https://doi.org/10.48550/arXiv.2402.14730

Karniadakis, G.E., Kevrekidis, I.G., Lu, L., Perdikaris, P., Wang, S., Yang, L., 2021. Physics-informed machine learning. Nat Rev Phys 3, 422–440. https://doi.org/10.1038/s42254-021-00314-5

Opportunities

Few-shot learning/Fine-tuning



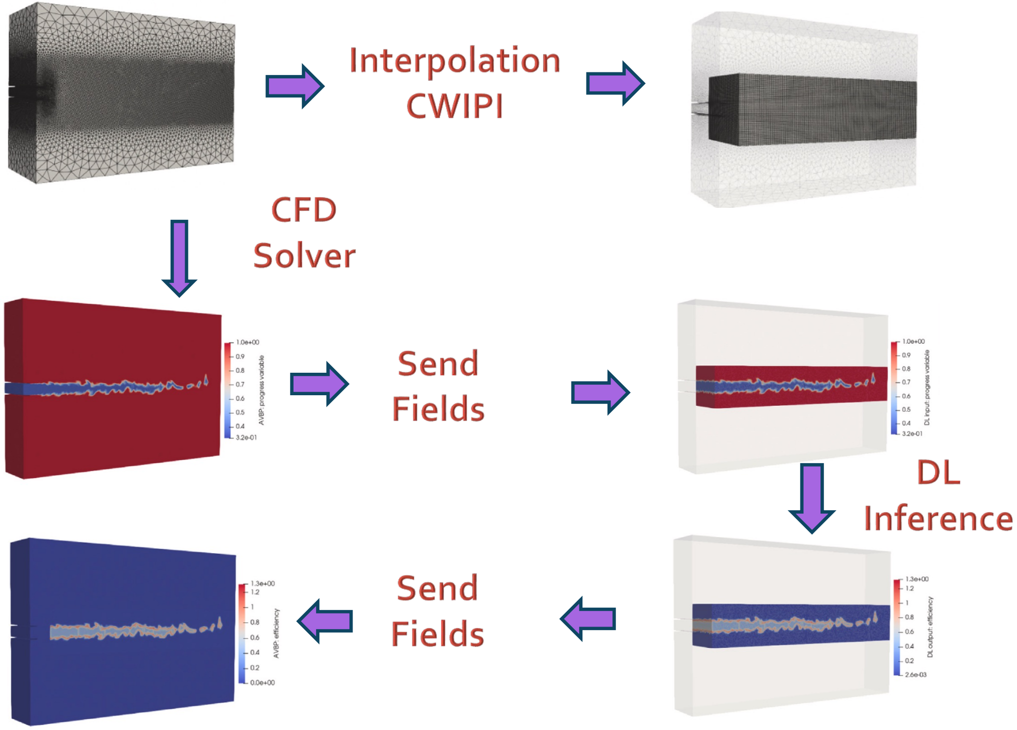


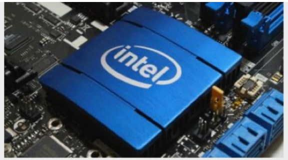
Hao, Z., Su, C., Liu, S., Berner, J., Ying, C., Su, H., Anandkumar, A., Song, J., Zhu, J., 2024. DPOT: Auto-Regressive Denoising Operator Transformer for Large-Scale PDE Pre-Training.

Bae, H.J., Koumoutsakos, P., 2022. Scientific multi-agent reinforcement learning for wall-models of turbulent flows. Nat Commun 13, 1443. https://doi.org/10.1038/s41467-022-28957-7

Opportunities

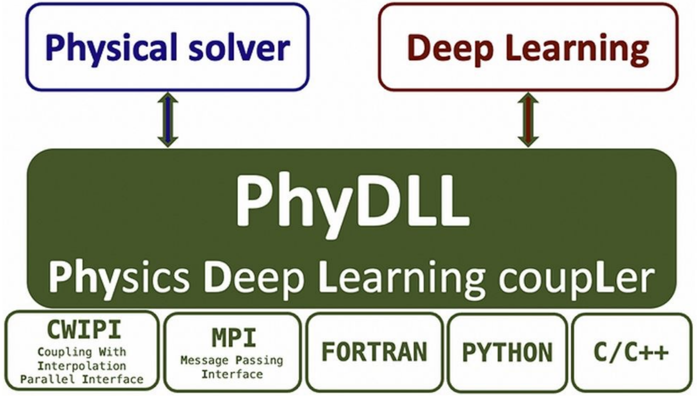
Hybrid DL/CFD models





x32 Intel Skylake

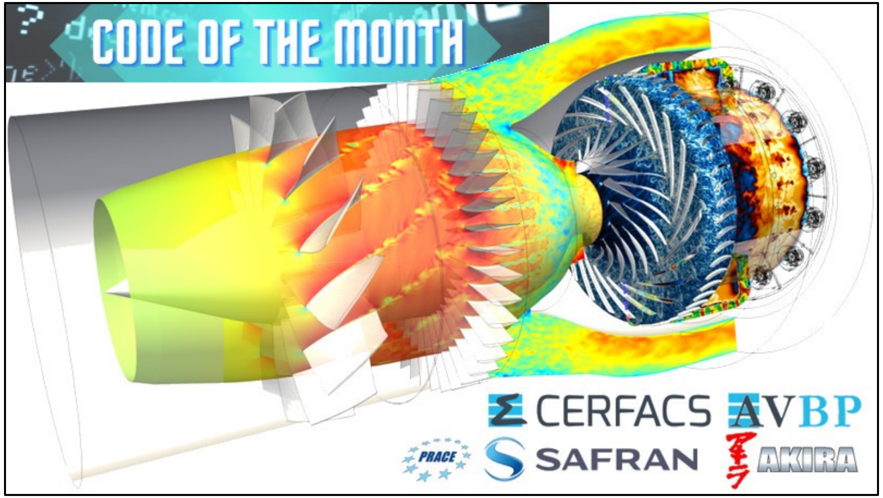
x4 Nvidia V100



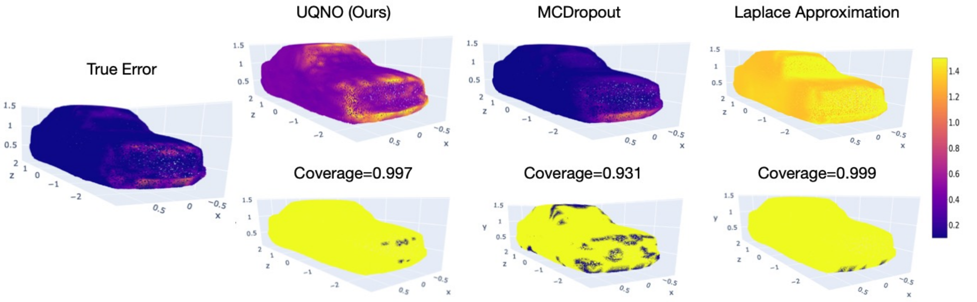
Serhani, A., Xing, V., Dupuy, D., Lapeyre, C., Staffelbach, G., 2024. Graph and convolutional neural network coupling with a high-performance large-eddy simulation solver. Computers & Fluids 278, 106306.
https://doi.org/10.1016/j.compfluid.2024.106306

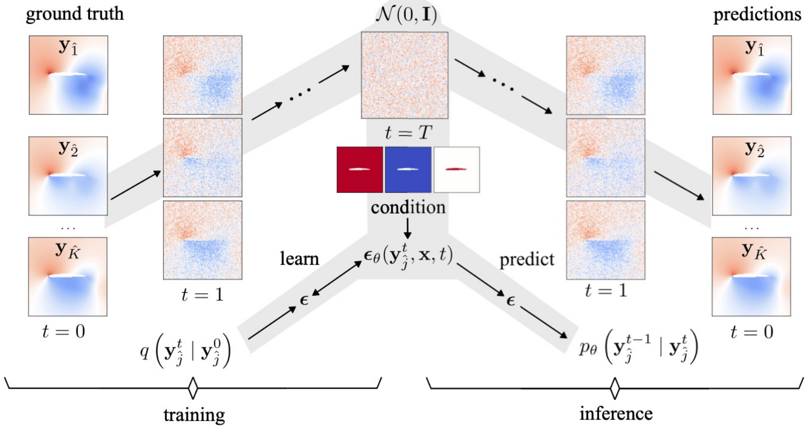
Opportunities

Fast and Reliable simulations for Engineering



SCHÖNFELD, T., AND RUDGYARD, M. Steady and unsteady flows simulations using the hybrid flow solver AVBP.





Ma, Z., Azizzadenesheli, K., Anandkumar, A., 2024. Calibrated Uncertainty Quantification for Operator Learning via Conformal Prediction. https://doi.org/10.48550/arXiv.2402.01960

Liu, Q., Thuerey, N., 2024. Uncertainty-aware Surrogate Models for Airfoil Flow Simulations with Denoising Diffusion Probabilistic Models. AIAA Journal 1–22. https://doi.org/10.2514/1.J063440

Conclusions & Perspectives

- Challenges around SciFMs for CFD are due to
  - Data heterogeneity, high-fidelity data scarcity, data secrecy, etc.
- Opportunities arise from community's commitment to open data
- CERFACS aims to build SciFMs for CFD using PhyDLL + AVBP
- Feedbacks are welcome !
  - Fernando González (gonzalez@cerfacs.fr)
  - Luciano Drozda (drozda@cerfacs.fr)