∂'Alembert
Institut Jean le Rond d'Alembert

Institut des sciences du calcul et des données
ALLIANCE SORBONNE UNIVERSITÉ

ISIR
INSTITUT DES SYSTÈMES INTELLIGENTS ET DE ROBOTIQUE

Postdoc Position

Model-consistent Bayesian learning of turbulence models from sparse data

Contacts

Paola CINNELLA (paola.cinnella@sorbonne-universite.fr)

Location : Sorbonne Université, Pierre et Marie Curie Campus, 4 Place Jussieu, 75005, PARIS, France
Profile of the successful candidate : PhD in Fluid Mechanics, Applied Mathematics or Computer science, taste for multidisciplinary research, proved skills in scientific computing
How to apply: please send the following information to P. Cinnella: CV, motivation letter, references.
Duration: two years
Starting date : flexible, preferably september 2023.
Salary : fully funded positions, partial refunding of local mobility fees

Context

Numerical simulation of fluids plays an essential role in modeling complex physical phenomena in domains ranging from climate to aerodynamics. Fluid flows are well described by Navier-Stokes equations, but solving these equations at all scales remains extremely complex in many situations and only an averaged solution supplemented by a turbulence model is simulated in practice. Unfortunately turbulence models present important weaknesses (Xiao and Cinnella, 2019). The increased availability of large amounts of high fidelity data and the recent development and deployment of powerful machine learning methods has motivated a surge of recent work for using machine learning in the context of computational fluid dynamics (CFD), and specifically turbulence modelling (Durasaimy et al., 2019). Combining powerful statistical techniques and model-based methods leads to an entirely new perspective for CFD. From the machine learning (ML) side, modeling complex dynamical systems and combining model-based and data-based approaches is the topic of active new research directions. With the aim of fostering progress in the understanding, modeling and design of turbulent flows the **Sorbonne Institute of Computating and Data Science** (ISCD) has funded the **LearnFluidS** (Machine-LEARNing for FLUID Simulations) projet team: a well-balanced team of researchers well-known in the fields of CFD, Deep Learning, numerical analysis and turbulence modeling.
Our aim is to develop the interplay between Deep Learning (DL) and CFD in order to improve turbulence modeling and to challenge state of the art ML techniques.

Participants

The project team **LearnFluidS** promotes **the development of recent machine learning advances in the field of computational fluid dynamics**. Until very recently these two domains were completely separated and this is only during the last few years, thanks to the considerable advances of Deep Learning and the increased availability of simulation data, that researchers from both fields started to cooperate. The project gathers specialists from the two disciplines involved in the thesis topic: fluid dynamics at **Institut Jean Le Rond ∂'Alembert** (Institute of theoretical, computational and experimental mechanics) and the **machine learning team MLIA** (Prof. P. Gallinari) at **ISIR** (Institute of intelligent systems and robotics). **∂'Alembert** has a recognized expertise in CFD, turbulence modelling and in the development of machine-learned RANS models using sparse formal identification techniques. The **MLIA** team at ISIR is well known for its expertise in Deep Learning, and has pioneered work in physics-constrained deep learning. LearnFLuids also benefits of the contribution of researchers from the **Arts et Métiers Institute of Technology** and from **Safran Tech**, experienced in data-driven models and surrogate models for computational fluid dynamics.

Objective: Learning turbulence models from sparse and noisy data

A major problem in ML is the need for significant amounts of training data and the difficulty at dealing with scarce and noisy data. However, in fluid dynamics, abundant and reliable databases are available only from so-called high-fidelity simulations such as Direct and Large Eddy simulations (DNS, LES), which are extremely costly and limited to low/moderate Reynolds numbers and relatively simple geometries. For high-Reynolds-number, complex configurations of industrial interest, increasing masses of data are also available from experiments, but which are incomplete (not all flow variables are available, within restricted observation domains) and possibly noisy. A natural way for tackling the learning problem from sparse and noisy data is adopting a Bayesian viewpoint, which endowes the model parameters and structures with measures of probability (e.g., Barber 2012). Bayesian machine learning (BML) is better suited for learning from scarce or noisy data, thanks to the regularizing effect of prior distributions. The present P.I. has approximately ten years of experience in Bayesian methods, including recent attemps of symbolic sparse Bayesian learning of turbulent closures (Cherroud et al. 2022).

In the proposed work, we wish to further explore **Bayesian learning algorithms** in the context of data-driven turbulence modeling, with special focus on CFD-in-the-loop training procedures. Such approaches allow learning from any observed data corresponding to an output of the CFD models supplemented by the BML turbulent closure. Additionally, we expect the Bayesian formulation provide robust solutions with quantified uncertainty for the model parameters, which can be propagated back through the solver to achieve improved CFD solutions with uncertainty estimates.

A critical aspect of the procedure is represented by model training. Markov Chain Monte Carlo techniques often used for the Bayesian inference of complex models are hardly in CFD due to the cost of performing a large number of CFD solves. However, significant speed up is expected by using gradient information or by replacing the costly CFD model by a ML surrogate. The present research will adress the development of efficient training techniques.

Afterwards, model mixture approaches will be considered to favor model applicability to a large variety of flow (e.g. Cherroud et al., 2022b, Ling et al., 2022). ML techniques for finding optimal partitions of the covariate space (clustering), dimensionality reduction, and the construction of suitable gating functions will be adressed, with application again to flow problems from the **NASA Collaborative Testing challenge for data-driven turbulence models** (https://turbmodels.larc.nasa.gov/turb-prs2022.html).

References

- Arjovsky M. et al. 2020. Invariant Risk Minimization. *arxiv.org/abs/2002.04692* (2020), 1–31.
- Barber
- Bhattacharya K. et al. 2020, Model reduction and neural networks for para- metric pdes. Preprint arXiv:2005.03180 (2020).
- Chen, R.T.Q. et al., 2018. Neural Ordinary Differential Equations. *NIPS* (2018), 31–60.
- Cherroud S., Merle X., Cinnella P., Gloerfelt X., 2022. Sparse Bayesian Learning of Explicit Algebraic Reynolds-Stress models for turbulent separated flows. International Journal of Heat and Fluid Flow 98(December):109047
- Cherroud S., Merle X., Cinnella P., Gloerfelt X., 2022b. Spatial Model Aggregation (X-MA) of stochastic Explicit Algebraic Reynolds Stress models. 2022 Symposium on Turbulence Modeling: Roadblocks, and the Potential for Machine Learning. Suffolk, VI, July 27-29 2022.
- Duraisamy K., Iaccarino G., and Xiao H., Turbulence modeling in the age of data, Annual Review of Fluid Mechanics 51, 357 (2019).
- Gautier N. et al., 2015 Closed-loop separation control using machine learning. Journal of Fluid Mechanics, 770 (2015), 442–457
- Kochkov, D. et al. 2021. Machine learning accelerated computational fluid dynamics. (2021), 1–13.
- Le Guen V., Yin Y., Dona J., Ayed I., de Bézenac E., Thome N., Gallinari, P. 2021. Augmenting Physical Models with Deep Networks for Complex Dynamics Forecasting. *ICLR* (2021), 1–22.
- Ling Y., Arranz G., Williams E., Lozano-Duran A., 2022. Wall-modeled large-eddy simulation based on building-block flows. Preprint, December 2022. 10.48550/arXiv.2212.05120
- Schmelzer M., Dwight R.P., Cinnella P. 2020, Discovery of Algebraic Reynolds-Stress Models Using Sparse Symbolic Regression. Flow, Turbulence and Combustion 104 (2020), 579–603.
- Sirignano J., MacArt J.F., Freund J.B. 2020. DPM: A deep learning PDE augmentation method with application to large-eddy simulation. *Journal of Computational Physics*. 423, (2020), 109811.
- Um K. et al. 2020. Solver-in-the-Loop: Learning from Differentiable Physics to Interact with Iterative PDE-Solvers. *Neurips* (2020), 1–12.
- Wang R. et al. 2020, Towards physics-informed deep learning for tur- bulent flow prediction, in Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Dis- covery & Data Mining (2020), 1457–1466.
- Willard J.D. et al. 2020. Integrating physics-based modeling with machine learning: A survey. *arXiv* (2020), 1–34.
- Xiao H., Cinnella P. 2019, Quantification of model uncertainty in RANS simulations: A review. Progress in Aerospace Sciences, 108 (2019), 1-31.