SORBONNE UNIVERSITÉ

PROGRAMME INTITUTS ET INITIATIVES

Appel à projet – campagne 2021

Proposition de projet de recherche doctoral (PRD)

SCAI - Sorbonne Center of Artificial Intelligence

Intitulé du projet de recherche doctoral (PRD):
Machine Learning of turbulence models in Computational Fluid Dynamics

Directeur.rice de thèse porteur.euse du projet (titulaire d'une HDR) :

NOM : **Cinnella** Prénom : **Paola**
Titre : Choisissez un élément : ou
e-mail :
Adresse professionnelle : Campus Pierre et Marie Curie, 5 Place Jussieu, Tour 55-65, bureau 516
(site, adresse, bât., bureau)

Unité de Recherche :
Intitulé : Institut Jean Le Rond D'Alembert
Code *(ex. UMR xxxx)* : UMR 7290

École Doctorale de rattachement de l'équipe (future école doctorale du.de la doctorant.e) : **ED391-SMAER**

Doctorant.e.s actuellement encadré.e.s par la.e directeur.rice de thèse (préciser le nombre de doctorant.e.s, leur année de 1e inscription et la quotité d'encadrement) : 0 dans l'ED SMAER; 2 dans l'ED 432 SMI de l'ENSAM (ancien établissement de rattachement); 1 doctorante de troisième année (cotutelle internationale, encadrement à 25%); 1 doctorant de troisième année (thèse CIFRE, encadrement à 33%)

--

Co-encadrant.e :

NOM : **Gallinari** Prénom : **Patrick**
Titre : Professeur des Universités ou HDR ☒
e-mail : patrick.gallinari@sorbonne-universite.fr

Unité de Recherche :
Intitulé : LIP6
Code *(ex. UMR xxxx)* : UMR 7606

École Doctorale de rattachement : **ED130-EDITE**
Ou si ED non Alliance SU : **1 en 1ere année**

Doctorant.e.s actuellement encadré.e.s par la.e co-directeur.rice de thèse (préciser le nombre de

doctorant.e.s, leur année de 1e inscription et la quotité d'encadrement) : 1 en 1ere année (50%), ; 2 en 2e année (50°% chacun); 2 en 3e année (50% chacun)

Co-encadrant.e :

NOM : Prénom :
Titre : Choisissez un élément : ou HDR ☐
e-mail :

Unité de Recherche :
Intitulé :
Code *(ex. UMR xxxx)* :

École Doctorale de rattachement : **Choisissez un élément :**
Ou si ED non Alliance SU :

Doctorant.e.s actuellement encadré.e.s par la.e co-directeur.rice de thèse (préciser le nombre de doctorant.e.s, leur année de 1e inscription et la quotité d'encadrement) :

Cotutelle internationale : ☒ Non ☐ Oui, précisez Pays et Université :

Selon vous, ce projet est-il susceptible d'intéresser une autre Initiative ou un autre Institut ?
☒ Non ☐ Oui, précisez Choisissez l'institut ou l'initiative :

Description du projet de recherche doctoral ***(en français ou en anglais) :***

Ce texte sera diffusé en ligne : il ne doit pas excéder 3 pages et est écrit en interligne simple.

Détailler le contexte, l'objectif scientifique, la justification de l'approche scientifique ainsi que l'adéquation à l'initiative/l'Institut.

Le cas échéant, préciser le rôle de chaque encadrant ainsi que les compétences scientifiques apportées. Indiquer les publications/productions des encadrants en lien avec le projet.
Préciser le profil d'étudiant(e) recherché.

VOIR ANNEXE

Machine Learning of turbulence models in Computational Fluid Dynamics

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. A typical example is turbulent fluid flows characterized by a a wide range of spatial and temporal scales. Direct Numerical Simulation (DNS) is usually prohibitive and one has to resort to smoothed versions of Navier-Stokes. Unfortunately these methods present important weaknesses. The increased availabilty 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), with the objectives of reducing computation costs and of solving complex problems like turbulence modeling. 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. All the lead conferences in the ML domain now feature topics such as physics and ML and dynamical systems and ML. This is then the context of this project, and 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.

Motivation

CFD models are the golden standard for preliminary and advanced aerodynamic design and optimization. Due to the high characteristic Reynolds numbers (representative of the ratio of inertial to viscous forces in the flow), most CFD applications deal with flows in the turbulent regime. High-fidelity CFD models like DNS or LES (Large Eddy Simulation) are very costly for complex high-Reynolds flows, and lower fidelity approaches, relying on the Reynolds-Averaged Navier-Stokes (RANS) equations represent the workhorse for engineering flow simulations. Due to the non-linearity of Navier-Stoles equations, unclosed terms appear in the averaging process and a turbulence model must be introduced to account for the unresolved flow scales. Despite much theoretical effort done for developing turbulence models on a physically sound basis, turbulence models yet rely on a simplified description of turbulence and they need a significant amount of empiricism and expert judgement both for defining the model mathematical structure and for calibrating the associated closure coefficients. A review of uncertainties and limitations of RANS models can be found in (Xiao and Cinnella, 2019). Building a universal RANS model valid for a wide range of flow applications is hard, and likely impossible, already for statistically steady turbulent flows. Inaccuracies are even more critical for flow simulations characterized by mean flow unsteadiness, where one needs to predict the average dynamics of the flow field.

The availability of a larger and larger amounts of high-fidelity numerical and experimental data for turbulent flow configurations has motivated interest in using modern ML tools to guide and automatize the development of turbulence models by learning from data (Durasaimy et al., 2019), while incorporating in the learning process physical constrains ensuring a physically acceptable behavior of the solution at any condition (see SOTA below). This gives promising results for RANS simulations statistically steady flows, leading to machine-learned, augmented turbulence models (e.g. the SpaRTA models of Schmeltzer et al., 2020). On the other hand, much work has been done in using ML to build reduced-order models of the turbulent dynamics of a given flow, namely, for the purpose of close-loop flow control (e.g., Gautier et al., 2015). Such models are limited to relatively low Reynolds number flows, and they can be hardly generalized to different flow cases. Finally, to our best knowledge, no attemps have been done yet to develop ML augmented RANS models for unsteady flow simulations.

Deploying ML for CFD is then a challenging and timely research area which has only started to be explored. We will address in the thesis project two main challenges as described below.

Combining CFD models and Deep Learning

Our objective is to improve traditional CFD models, both in terms of complexity and of accuracy of the predictions, with the addition of ML components. Recent progresses, and the generalized use of automatic differentiation both for differentiable solvers and DL algorithms have paved the road to the integration of DL techniques and ODE/PDE solvers. In the ML community, a starting point for such investigations was the Neural ODE paper (Chen 2018) that promoted the use of ODE solvers for ML problems. This however remains limited to simple temporal dynamics and the problems of solving complex PDE requiring efficient space and time discretization remains fully open. Combining statistical and numerical models raises many open questions such as characterizing the solutions of such systems and their coherence, and deriving efficient combination frameworks. We advocate for this research the use of DL modules for complementing CFD solvers, in the spirit of (Le Guen 2021) who introduced a principled approach however still limited to basic PDEs. In our new context, we will analyze how to model unclosed terms in the RANS equations. This approach can be seen as a generalization of classical closure models. In order to make easier this theoretical analysis, the approach will be first developed for a scalar surrogate of the Navier-Stokes equations, namely, the nonlinear Burgers' equation, which has been widely used in the literature as a simplified ansatz for Navier-Stokes turbulence. The framework will then be deployed and adapted to the specificity of unsteady RANS simulations. Turbulence model agmentation will be achieved by supplementing classical closure models for which we have prior knowledge with data-driven corrections. The whole system will be trained end to end with the DL modules and the numerical solvers using high-fidelity data. Note that the latter may be incomplete (selected flow properties are observed on a restricted portion of the flow domain) or noisy (this is often the case for experimental data), which must be properly accounted for in the learning procedure.

Learning in multiple environments

In order to be useful for CFD applications a learned model must accurately simulate flows outside of the training distribution: operational conditions and environment may vary according to different physical factors thus requiring models to extrapolate to these new conditions. DL could in principle be extremely efficient for learning complex dynamics. However since they do not include the underlying physics, they often do not enforce physical constraints such as incompressibility or conservation laws. Said otherwise, even when they perform well on training data, they struggle with generalization to out-of-distribution data. Recent ML research mostly considered this problem for static data and classification or regression problems, promoting either robust optimization methods or learning invariant representations (Arjovsky 2020). This is currently a very active research topic, but this is not adapted to our dynamical context which is much more complex. We will adopt a new perspective by considering learning dynamical models from multiple environments and propose a new framework leveraging the commonalities and discrepancies among environments. This could be achieved by capturing common dynamic characteristics in shared modules while additional terms capture environment specific dynamics. We expect this new setting to be more robust to new distributions than classical empirical risk minimization or robust optimization schemes. This setting will be analyzed both theoretically and for practical situations. We will in particular consider generalizing to different geometries sharing common flow features and to different flow conditions for a given geometry. Both issues are largely open.

Position w.r.t. State of the Art

The general problem of solving PDEs with neural networks is not new but it is only very recently, with the large scale deploiement of DL methods, that the topic gained momentum and started to motivate the interest of several communities. The first dedicated workshop was held at ICLR 2020, a major ML conference. In the field of CFD, a family of approaches uses ML to fit closures to classical turbulence models based on agreement with DNS (Duraisamy 2019). These models are mostly restricted to a narrow class of simple, steady, flows. Pure ML approaches advocate replacing the CFD pipeline with ML models trained from simulations (Bhattacharya 2020, Wang 2020). However, pure ML methods do not generalize well when trained on complex dynamics and they lack physical plausibility. Our approach is more in the line of combining model based and data based ideas in a hybrid framework (Willard 2020, Sirignano 2020, Le Guen 2021, Kochlov 2021, Um 2020). The problems adressed in this project build on this series of emerging ideas by adressing new problems.

Adequacy to SCAI and Role of the participants

The thesis project 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 2 years, thanks to the considerable advances of Deep Learning and the increased availability of simulation data, that tresearchers from both fields started to cooperate. This is then a typical emerging intedisciplinary domain which. As for the participants, the project gathers specialists from the two disciplines involved in the thesis topic: fluid dynamics @ ∂'Alembert and machine learning @LIP6. ∂'Alembert has a recognized expertise in CFD, turbulence modelling and in the development of open-box machine-learned RANS models using sparse formal identification techniques. The Machine Learning team at LIP6 is well known for its expertise in Deep Learning. One of the topic of the team has been for some years an interdisplinary research on dynamical systems involving cooperation with maths and climate specialists. This cooperation offers a new opportunity for enlarging this research domain to a new discipline (mechanics) and a new lab. (∂'Alembert) We consider that this has the potential to open the door to a new series of cooperations between mechanics and computer science at Sorbonne University.

Profil recherché: master ou école d'ingénieur, bonne bases en mathématiques appliquées à la mécanique des fluides, bonne connaissance du machine learning.

References

Xiao H., Cinnella P. 2019, Quantification of model uncertainty in RANS simulations: A review. Progress in Aerospace Sciences, 108 (2019), 1-31.

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.

Gautier N. et al., 2015 Closed-loop separation control using machine learning. Journal of Fluid Mechanics, 770 (2015), 442–457

Arjovsky M. et al. 2020. Invariant Risk Minimization. *arxiv.org/abs/2002.04692* (2020), 1–31.

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.

Duraisamy K., Iaccarino G., and Xiao H., Turbulence modeling in the age of data, Annual Review of Fluid Mechanics 51, 357 (2019).

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.

Um K. et al. 2020. Solver-in-the-Loop: Learning from Differentiable Physics to Interact with Iterative PDE-Solvers. *Neurips* (2020), 1–12.

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.

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.