ParCFD'2020
32nd International Conference on Parallel Computational Fluid Dynamics
May 11-13 2020, Nice FRANCE

COUPLING CFD AND DEEP LEARNING TO BUILD DATA-DRIVEN MODELS: TWO CASE-STUDIES IN PROCESS ENGINEERING

A. MARCATO*, M. FERRARI, G. BOCCARDO*, A. BUFFO, M. VANNI, D. MARCHISIO

* Politecnico di Torino
Department of Applied Science and Technology
10129 Torino, ITALY
e-mail: gianluca.boccardo@polito.it, web page: http://www.disat.polito.it/it/

Key words: CFD methodology, Machine Learning, Packed-bed Reactors, Filtration, Deep Learning, Food emulsions.

Summary. In this work we have studied the coupling of CFD with machine learning (neural networks) models on two different case studies. The first case study is the mixing step of the mayonnaise production for which the performance of the data-driven model is very good, with errors below 1% between predictions of the surrogate model and CFD results: the network can also be used for extrapolation in ranges of operating conditions slightly outside the training range. The second case is fluid flow and transport of a colloid through porous media, where the surrogate model performance is slightly worse than for the previous case with average errors lower than 8% for the prediction of both permeability and average volume concentration.

1 INTRODUCTION

Machine learning techniques such as neural networks and deep learning spread in everyday applications and their impact is considered a revolution. In the scientific research these algorithms started to be employed in the last decade for example for new materials development [1] or biological molecules structure prediction [2]. Although chemical engineering is not properly considered a "big data" field, if simulations can be executed on the studied system it is possible to obtain a dataset large enough for the implementation of these techniques [3]. In this work computational fluid dynamics simulations are inserted in a workflow for the implementation of neural networks algorithms aimed at the construction of data-driven models that could be useful to improve multiscale modelling and industrial in-line control, as they are able to instantaneously predict a certain output after be trained with the CFD results. In this context two case studies are approached: the production of food emulsions and the flow and transport through porous media.

The production of food emulsions, mayonnaise in particular, is a process engineering example of application of neural networks. It is carried out in two stages: initially the ingredients are mixed in batch, then they are conveyed into a rotating equipment, the cone mill, which imparts the proper shear rate for the emulsification [4]. This last step of the proceeding is simulated through CFD and the data-driven model trained with those results is able to predict macroscopic properties of the system which can be useful to implement more

sophisticated models, such as population balance models.

Instead the study of particle transport in porous media touches a wide variety of different fields: from the study of contaminant transport in aquifers to the design of effective packed bed reactors in chemical engineering [5]. Porous media are characterized by many parameters, which are generally geometric, and whose impact and synergy of action may be impossible to analytically predict; these particularities make it a prime candidate for machine learning approaches, which are particularly suited to extract essential features hidden in data.

2 METHOD AND NUMERICAL DETAILS

The first step of the workflow is to build the computational models: the CFD simulations are performed on the open source toolbox OpenFOAM v6.

For the food emulsions case, the geometry is axisymmetric: the cone mill is modeled as two coaxial truncated cones, the internal one rotates and the external one is fixed. The simulations are performed under the hypothesis of an incompressible, pseudo-single-phase non-Newtonian fluid, where an appropriate rheological law is provided by Dubbelboer et.al. [4]. The velocity and pressure fields are determined by solving the continuity equation and the Navier-Stokes equations through the SRFSimpleFoam solver imposing a no-slip condition on the rotor and null velocity on the stator as boundary conditions.

For the porous media case, we implemented a bidimensional geometry constituted by a random periodic arrangement of grains having a Gaussian diameter distribution. Initially a representative elementary volume (REV) study has been conducted in order to evaluate the minimum number of grains (or the minimum dimension of the box) necessary to have a constant porosity. After the solver simpleFoam is used to solve the Navier-Stokes equation in laminar conditions, as boundary conditions we imposed periodic conditions on the upper and lower face of the domain, a non-slip condition on the surface of the grains and a constant inlet pressure. For the transport simulations the advection-diffusion equation is solved through the solver scalarTransportFoam imposing a constant unitary inlet concentration and a null concentration on the surface of the grains, representing a perfectly efficient filtration process.

After having designed the computational models, the simulations have been run on an HPC cluster in order to create the dataset for the training of the neural networks. The simulations differ in some input operating conditions and geometrical parameters which are varied in an established range: these will be the features for the prediction made by the neural network. For the food emulsion case the inlet velocity of the fluid, the rotating velocity of the cone mill, and the distance between the cones are varied and the volume average strain rate is the output result of the simulations; we chose this quantity as the output as it is the most impactful on droplet coalescence and breakage rates. For the case of flow in porous media, the porosity of the system, the mean diameter, and the standard deviation of the distribution of the grains diameter are the features chosen to predict the permeability of the medium, evaluated through the Darcy equation. The inlet pressure and the colloid diameter are added to the previous predictors to evaluate the volume average concentration of the colloid in the medium.

The dataset created is then fed to a fully connected neural network characterized by one hidden layer containing 20 neurons for the food emulsion case and 40 neurons for the porous

media case. The machine learning models are created using the Python library Keras.

3 RESULTS

Figure 1a shows a contour plot of the axial velocity of the fluid in the cone mill where, as Reynolds number is greater than 75, the formation of a Taylor vortex is apparent. Figure 1b shows a detail of the contour plot of the colloid concentration in the porous medium.

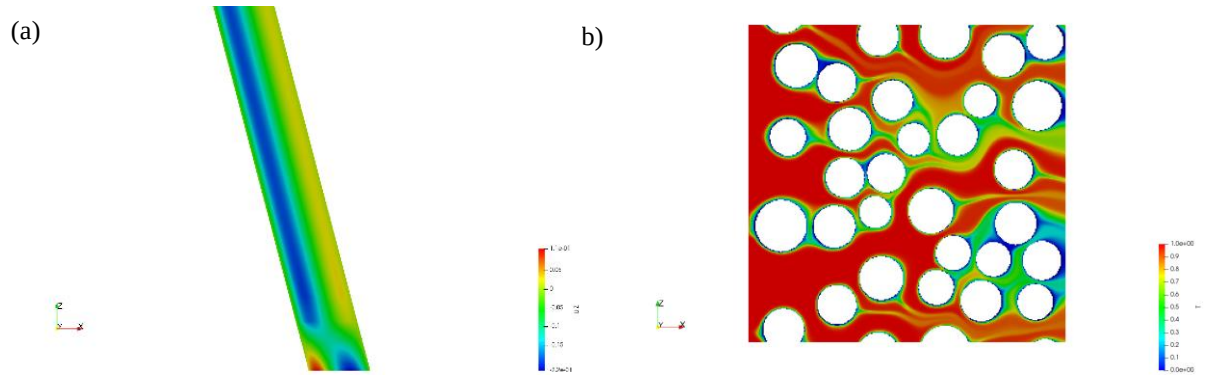

Figure 1: Contour plot for the axial velocity (Re = 138.5) (a) and for the concentration (Pe = 1629) (b)

As it has been mentioned, the CFD results were analysed and post-processed with the purpose of obtaining, for each geometry and operating condition considered, a set connecting the input features of the simulation with the result(s) obtained. The numerosity of this set is equal to the number of simulations performed to build the training set for the neural network (which is numbering in the thousands). After the neural network training process, the end result is a surrogate black-box model capable of predicting the output values when given a new set of input features, with varying accuracy depending on the model considered and on the distance of the new input feature from the original set of features used in the training process.

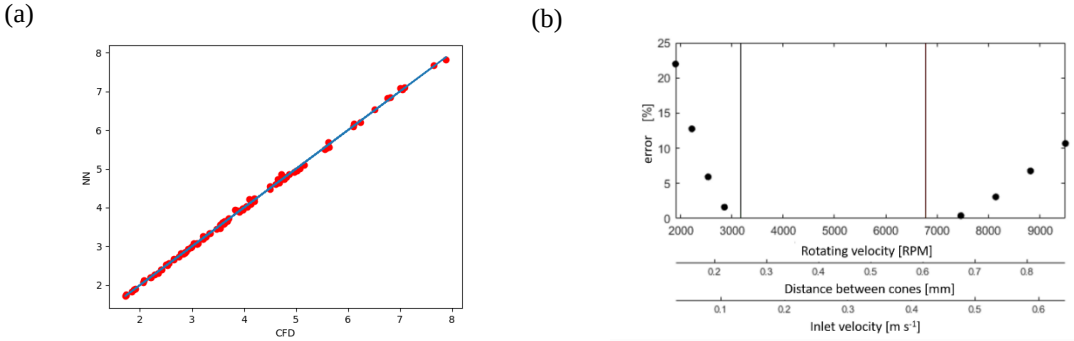

Figure 2: Neural network performance for the food emulsion case: (a) parity diagram, (b) evaluation on the model outside the range explored during training

This accuracy is defined as the error between results of CFD with new inputs, and predictions of the neural networks with these new inputs.

For the food emulsions case the accuracy of the data-driven model inside the range of

feature explored during the training is very good, in fact, the error is less than 1% (Figure 2a); while the error increases the more the feature moves away from the training range (Figure 2b).

For the porous media case the average error associated to the prediction of the permeability is 5.8% (Figure 3a), and the one of the volume average concentration is 5.6% (Figure 3b).

(a) (b)

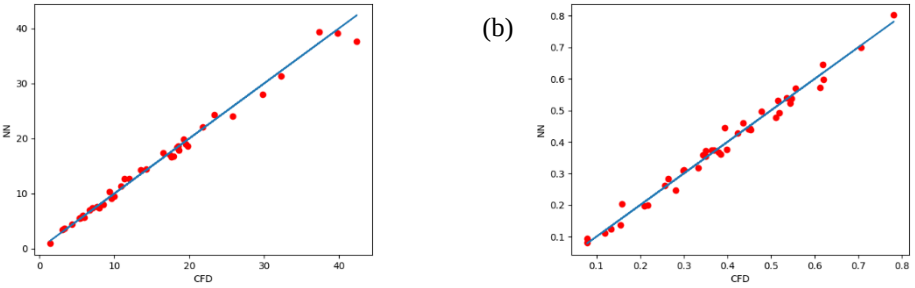

Figure 3: Neural network performance for the flow and transport in porous media case: (a) permeability prediction, (b) volume-averaged concentration prediction

4 CONCLUSIONS

In this work we proved the possibility of using CFD simulations results as training data for machine learning algorithms on two relevant case studies in chemical engineering: the production of food emulsions and the flow and transport through porous media. The time required for the resolution of the simulations varies from few minutes for the food emulsion case to about thirty minutes for the transport ones; the neural network training needs less than five minutes in both cases, while the calls for new outputs, given new input samples, are instantaneous: this means that data-driven models can be easily integrated in other multi-scale models. The future perspectives (and current work) are the integration of an upscaling procedure for the characterization of colloid transport in the porous medium, and the exploration of new deep learning techniques, such as convolutional neural networks, for the prediction/generation of the entire flow field (and not just averaged integral values) starting from a simple geometric description of the case.

REFERENCES

[1] F. Oviedo, Z. Ren, S. Sun, C. Settens, Z. Liu, N.T.P. Hartono, S. Ramasamy et al., "Fast and interpretable classification of small x-ray diffraction datasets using data augmentation and deep neural networks", *npj Computational Materials*, **5**, 1-9 (2019).

[2] F. Noé, G. De Fabritiis and C. Clementi, "Machine learning for protein folding and dynamics", *Current Opinion in Structural Biology*, **60**, 77-84 (2020).

[3] V. Venkatasubramanian, "The promise of artificial intelligence in chemical engineering: Is it here, finally?", *AIChE Journal*, **65**(2), 466–478 (2019).

[4] A. Dubbelboer et al., "Pilot-scale production process for high internal phase emulsions: Experimentation and modeling", *Chemical Engineering Science*, **148**, 32–43 (2016).

[5] G. Boccardo, D. L. Marchisio and R. Sethi, "Microscale simulation of particle deposition in porous media", *Journal of colloid and interface science*, **417,** 227–237 (2014).