A PLATFORM FOR DECODING AND RECODING TURBULENCE

Understanding and modeling turbulent phenomena in Fluid Mechanics is pivotal in many areas of science, ranging from life and environmental sciences to engineering applications. As numerical simulations and experimental techniques have developed in recent years, leading to the generation of extensive databases, opportunities for testing and developing new theories and more complex, realistic, models have multiplied, but leave us with the challenge of developing adequate strategies for the management and exploitation of large databases. Moreover, independently of their size, the analysis of the databases always needs to be carried out in a context of missing information: for instance, numerical simulations are limited in time-span, laboratory experiments only provide sparse spatio-temporal measurements and the Fluid Mechanics parameters describing a real-life problem may be unknown or cannot be determined with precision. The challenge is to overcome these limitations in order to provide (*i*) efficient and relevant characterization of the turbulent phenomena, (*ii*) reconstruction of missing data and/or generation of virtual data in a realistic context.

The long-term aim of the project is to develop an interdisciplinary platform for the analysis and reconstruction of turbulent flows bridging the gap from academic configurations to real-life situations. Historically, the approach to turbulence has always been grounded on statistical methods, and we believe that the Center for Data Science represents a unique opportunity for the turbulence community of Université Paris-Saclay to interact in a new way with statisticians and data scientists. The goal of the project is not only to improve our understanding of the fundamental physics of turbulence using statistical methods, but also to provide models for applications of interest to other scientific communities on issues related to high dimensional, wide-spectrum, data from turbulence.

The methods will be tested on a collection of databases of different sizes, natures, and origins, with the aim to foster two-way transfers of knowledge from one application to the other as well as between real-life applications and academic problems.

The project is divided into two connected tasks:

1) Analysis of Fluid Mechanics databases using application-tailored data science analysis tools (*e.g.* clustering, goal-oriented dictionary learning, graph analysis in a network-theoretic framework),
2) Reconstruction/ Generation of data in a missing information context using statistical methods (*e.g.* compressed sensing, deep learning).

The current team consists of B. Podvin, L. Mathelin and C. Tenaud, who are researchers at LIMSI in the Mechanics Department, with an expertise in turbulence and/or statistical methods.