Data-driven model discovery from fusion plasma turbulence simulations

I. Abramovic[1,3], E. P. Alves[2], M. Greenwald[3]

[1] *University of Technology Eindhoven, Eindhoven, The Netherlands*

[2] *University of California, Los Angeles, United States*

[3] *Massachusetts Institute of Technology, Cambridge, United States*

A machine learning method has shown the ability to extract the underlying physics equations using only the data from fusion plasma turbulence simulations. This can potentially be applied for derivations of reduced models and validation against experimental data. Turbulence, being a highly non-linear process, is challenging to theoretically describe and computationally model. High-fidelity computational models come at a high computational cost and cannot be applied for routine simulation of plasma turbulence. Reduced models based on artificial neural networks could fill the need for affordable simulations. However, the training requires very large volumes of data and the obtained models lack interpretability, making it difficult to understand the physics being learned and whether it can be extrapolated to scenarios not encountered previously during training. This is especially problematic for predictions in future machines which will operate in unexplored parameter ranges. In particular, we explore parameter ranges foreseen for MIT's high field tokamak SPARC [1]. SPARC will feature high temperature superconducting magnets enabling it to have a field of 12 T on axis and demonstrate a fusion net energy gain. To circumvent the problem of extrapolability and data-inefficiency, we explore a data-driven model discovery approach based on sparse regression and infer governing partial differential equations from the data [2, 3]. Our data is generated by simulations of drift-wave turbulence according to Hasegawa-Wakatani model [4]. Balancing model accuracy and complexity enables the reconstruction of the governing equations, accurately describing the dynamics from the input data sets. The methodology is further applied to more complex synthetic data produced by the Global Drift Ballooning code [5]. The findings show that the methodology is promising for development of reduced and computationally cheap models as well as for existing model validation.

References

[1] A. J. Creely, M. J. Greenwald et al., *Overview of the SPARC tokamak*, J. Plasma Phys (2020)

[2] S. Brunton, J. L. Proctor and J. N. Kutz, *Discovering governing equations from data by sparse identification of non-linear dynamical systems*, PNAS 113 15 (2016)

[3] E. P. Alves and F. Fiuza, *Data-driven discovery of reduced plasma physics models from fully-kinetic simulations*, arXiv:2011.01927 (2020)

[4] M. Wakatani and A. Hasegawa, *A collisional drift wave description of plasma edge turbulence*, The Physics of Fluids (1984)

[5] D. M. Fisher and B. N. Rogers, *Two-fluid biasing simulations of the large plasma device*, Phys. Plasmas 24 (2017)