Predicting extreme events using dynamics based machine learning

Dario Lucente[1], George Miloshevich[1], Corentin Herbert[1] & Freddy Bouchet[1]

Univ Lyon, ENS de Lyon, Univ Claude Bernard, CNRS, Laboratoire de Physique, F-69342 Lyon, France
dario.lucente@ens-lyon.fr

Many phenomena in the climate system lie in the gray zone between weather and climate : they are not amenable to deterministic forecast, but they still depend on the initial condition. A natural example is medium-range forecasting, which is inherently probabilistic because it lies beyond the predictability time of the atmosphere. Similarly, one may ask the probability of occurrence of an El-Niño event several months ahead of time or the probability of occurrence of a heat wave a few weeks in advance based on the observed atmospheric circulation. In this talk, we introduce a quantity which corresponds precisely to this type of prediction problem : the committor function is the probability for an event to occur in the future, as a function of the current state of the system. In the first part of this presentation, we explain the main mathematical properties of this probabilistic concept, and compute it in the case of a low-dimensional stochastic model for El-Niño, the Jin and Timmerman model. This example allows us to show that the ability to predict the probability of occurrence of the event of interest may differ strongly depending on the initial state : in some regions of phase space, the committor function is smooth (intrinsic probabilistic predictability) and in some other regions, it depends sensitively on the initial condition (intrinsic probabilistic unpredictability), see Fig. (1). We stress that this predictability concept is markedly different from the deterministic unpredictability arising because of chaotic dynamics and exponential sensivity to initial conditions. The second part of the talk is about how to efficiently compute the committor function from data through several data-driven approaches, such as direct estimates, kernel-based methods and neural networks. We discuss two examples : a) the computation of committor function for the Jin and Timmerman model, b) the computation of committor function for extreme heat waves. Both systems are highly nonlinear but, considering the dimensionality of the two, their level of complexity is profoundly different. This therefore allows us to explore and discuss the performance and limits of the different methods proposed. Finally, we propose a method for learning effective dynamics by introducing a Markov chain on the data. Using the Markov chain we are able to quickly and easily compute many interesting quantities of the original system, including the committor function. The goal is to overcome some of the limitations of the methods introduced previously and to develop a robust algorithm that can be useful even in the lack of data.

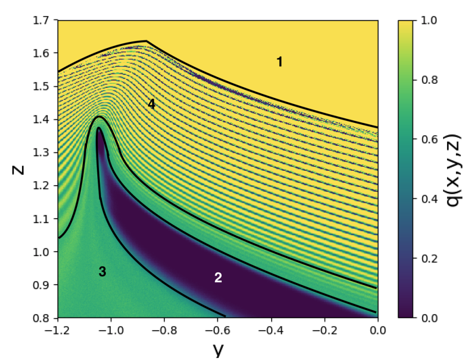

Figure 1. Colour plot of the committor function $q(x, y, z)$ in the plane $x = -2.8310$. Regions with uniform $q = 0$ or 1 values correspond to deterministic predictability, smooth regions with $0 < q < 1$ to probabilistic predictability, and regions with sensitive dependence on intial conditions to unpredictable parts of phase space.