EGU22-654
https://doi.org/10.5194/egusphere-egu22-654
EGU General Assembly 2022

On the derivation of data-driven models for partially observed systems

Said Ouala[1], Bertrand Chapron[2], Fabrice Collard[3], Lucile Gaultier[3], and Ronan Fablet[1]
[1]IMT Atlantique, Lab STICC, Signal et communications, France (said.ouala@imt-atlantique.fr)
[2]Ifremer, LOPS, Plouzané, France (bertrand.chapron@ifremer.fr)
[3]OceanDataLab, Locmaria-Plouzané, France (dr.fab@oceandatalab.com, lucile.gaultier@oceandatalab.com)

When considering the modeling of dynamical systems, the increasing interest in machine learning, artificial intelligence and more generally, data-driven representations, as well as the increasing availability of data, motivated the exploration and definition of new identification techniques. These new data-driven representations aim at solving modern questions regarding the modeling, the prediction and ultimately, the understanding of complex systems such as the ocean, the atmosphere and the climate.

In this work, we focus on one question regarding the ability to define a (deterministic) dynamical model from a sequence of observations. We focus on sea surface observations and show that these observations typically relate to some, but not all, components of the underlying state space, making the derivation of a deterministic model in the observation space impossible. In this context, we formulate the identification problem as the definition, from data, of an embedding of the observations, parameterized by a differential equation. When compared to state-of-the-art techniques based on delay embedding and linear decomposition of the underlying operators, the proposed approach benefits from all the advances in machine learning and dynamical systems theory in order to define, constrain and tune the reconstructed sate space and the approximate differential equation. Furthermore, the proposed embedding methodology naturally extends to cases in which a dynamical prior (derived for example using physical principals) is known, leading to relevant physics informed data-driven models.