Learning missing part of physics-based models within a variational data assimilation scheme

Arthur Filoche1, Julien Brajard2,3, Anastase Charantonis2,4, and Dominique Béréziat1
1Sorbonne Université, LIP6, Paris, France (arthur.filoche@lip6.fr)
2Sorbonne Université, LOCEAN, Paris, France
3Nansen Environmental and Remote Sensing Center, Bergen, Norway
4École Nationale Supérieure d’Informatique pour l’Industrie et l’Entreprise, Evry, France

The analogy between data assimilation and machine learning has already been shown and is still being investigated to address the problem of improving physics-based models. Even though both techniques learn from data, machine learning focuses on inferring model parameters while data assimilation concentrates on hidden system state estimation with the help of a dynamical model.

Also, neural networks and more precisely ResNet-like architectures can be seen as dynamical systems and numerical schemes, respectively. They are now considered state of the art in a vast amount of tasks involving spatio-temporal forecasting. But to train such networks, one needs dense and representative data which is rarely the case in earth sciences. At the same time, data assimilation offers a proper Bayesian framework allowing to learn from partial, noisy and indirect observations. Thus, each of this field can profit from the other by providing either a learnable class of dynamical models or dense data sets.

In this work, we benefit from powerful and flexible tools provided by the deep learning community based on automatic differentiation that are clearly suitable for variational data assimilation, avoiding explicit adjoint modelling. We use a hybrid model divided into 2 terms. The first term is a numerical scheme that comes from the discretisation of physics-based equations, the second is a convolutional neural network that represents the unresolved part of the dynamics. From the Data Assimilation point of view, our network can be seen as a particular parametrisation of the model error. We then jointly learn this parameterisation and estimate hidden system states within a variational data assimilation scheme. Indirectly, the issue of incorporating physical knowledge into machine learning models is also addressed.

We show that the hybrid model improves forecast skill compared to traditional data assimilation techniques. The generalisation of the method on different models and data will also be discussed.
