EGU22-5692
https://doi.org/10.5194/egusphere-egu22-5692
EGU General Assembly 2022

Model error correction with data assimilation and machine learning

Alban Farchi[1], Marc Bocquet[1], Patrick Laloyaux[2], Massimo Bonavita[2], Marcin Chrust[2], and Quentin Malartic[1,3]
[1]CEREA, École des Ponts and EDF R&D, Île-de-France, France
[2]ECMWF, Shinfield Park, Reading, United Kingdom
[3]LMD/IPSL École Normale Supérieure and PSL University, École Polytechnique, Université Paris-Saclay, Sorbonne Université, CNRS, Paris, France

The idea of using machine learning (ML) methods to reconstruct the dynamics of a system is the topic of recent studies in the geosciences in which the key output is a surrogate model meant to emulate the dynamical model. In order to treat sparse and noisy observations in a rigorous way, ML can be combined with data assimilation (DA). This yields a class of iterative methods in which, at each iteration a DA step estimates the system's state, and alternates with a ML step to learn the system's dynamics from the DA analysis.

This framework can be used to correct the error of an existent, physical model. The resulting surrogate model is hybrid, with a physical and a statistical part. In practice, the correction can be added as an integrated term (i.e. in the model resolvent) or directly inside the tendencies of the physical model. The resolvent correction is easy to implement but is not suited for short-term predictions. The tendency correction is more technical since it requires the adjoint of the physical model, but also more flexible and can be used for any forecast lead time.

In this presentation, we start by a proof of concept for the use of joint DA and ML tools to correct model error. We use the resolvent correction with simple neural networks to correct the error of a two-dimensional, two layer quasi-geostrophic layer. The difference between the resolvent and the tendency correction is then illustrated with the two- scale Lorenz model. Finally, we show that the tendency correction opens the possibility to make online model error correction, i.e. improving the model progressively as new observations become available. We compare online and offline learning using the same twin experiment with the two-scale Lorenz model.