End-to-end learning for Variational data assimilation : application to river networks and assessment of extremes

Nicolas Lafon[*1]

[1]LSCE – Laboratoire des Sciences du Climat et de l'Environnement, UMR CEA-CNRS-UVSQ, Gif sur Yvette, France – Orme des merisiers, batiment 714, Saint-Aubin, France

Résumé

Joint work with R.Fablet and P.Naveau The purpose of data assimilation is to reconstruct the time evolution of some state given a series of observations. This field is crucial in a wide range of applications, and aim to make the most of available observation datasets. For instance, it can include both in situ and satellite-derived data, combine with state-of-the-art numerical models. A vast family of data assimilation methods comes to the minimization of some energy or functional which involves two terms, a dynamical prior and an observation term. Variational data assimilation relies on a continuous state-space formulation and results in a gradient-based minimization of a defined variational cost. The variational cost we used is known as the weak-constraint 4D-Var (see Trémolet). Working with river flows, an explicit dynamical model is not known. To learn the model as well as to reconstruct the dynamics of river flows, we use the learning framework propose by *Fablet et al (2021)*. It relies on an end-to-end neural network architectures for data assimilation which comprises a variational model and a gradient-based solver both implemented as neural networks. Concerning the design of Neural Network architectures to find the state-space operator, we exploit a two-scale CNN architecture whose input and output has same dimension as the hidden state for more details. The spatial interpolation scheme and temporal dynamics are modelled by intregrating a spatial-temporal mask in the cost function. The global end-to-end learning architecture enables us to learn both model operator and solver with respect to some criterion which can differ from the variational cost. A supervised criterion is used. Once the model operator is learned, we can get insight in the features learned, and analyze the dependencies of the measurements in the river flow between different stations. We pay specific attention on the performance algorithms for extreme values, especially if the underlying dynamics of extremes is well restituted. We might introduce some methods to improve results on extreme values.
References :

- R.Fablet, B.Chapron, L. Drumetz, E. Memin, O. Pannekoucke and F. Rousseau. End-to-end learning for variational data assimilation models and solvers, 2021
- Trémolet, Y., 2007. Model-error estimation in 4D-Var. Q. J. R. Meteorol. Soc., 133, 1267–1280.

---

[*]Intervenant