Differentiable Programming for hybrid ocean data assimilation and machine learning

Despite the revolution in autonomous in-situ and satellite remote sensing, much of the ocean remains poorly sampled observationally in terms of its spatio-temporal evolution. Data assimilation seeks to combine the incomplete knowledge reservoirs from sparse observations and numerical models that require the provision of uncertain inputs to produce optimal state and parameter estimates of the system considered. In its various renderings, methods of optimal estimation and control are being used to learn uncertain parameters in models (calibration problem) or to produce optimal initial states for prediction (extrapolation problem). Scientific machine learning methods rely on similar mathematical concepts that underpin optimal estimation and may be successful when the system is heavily sampled. We explore some of the concepts in the context of ocean modeling, in particular the concept of differentiable programming. We provide a perspective on opportunities and challenges with a view toward hybrid data assimilation and machine learning to enable online or full-model learning. These ideas are being implemented in the context of general-purpose automatic differentiation applied to a new ocean model written in the Julia programming language.

Patrick Heimbach, University of Texas at Austin, USA