Geophysical Research Abstracts
Vol. 12, EGU2010-9144, 2010
EGU General Assembly 2010

Beyond weak constraint 4DVAR: a bridge to Monte Carlo methods?

Dan Cornford (1), Yuan Shen (1), Michael Vrettas (1), and Manfred Opper (2)

(1) NCRG, Aston University, Computer Science, Birmingham, United Kingdom (d.cornford@aston.ac.uk), (2) Department of Artificial Intelligence, Technical University of Berlin, Berlin, Germany

Data assimilation is often motivated from a Bayesian perspective, however most implementations introduce approximations based on a very small number of samples (ensemble Kalman filter / smoother) to perform a statistical linearisation of the system model, or seek an approximate mode of the posterior distribution (4DVAR). In statistics the alternative approaches are based on Monte Carlo sampling optimally using particle filters / smoothers or Langevin path sampling, neither of which are likely to scale well enough to be applied to realistic models in the near future.

In this work we explain a new approach to data assimilation based on a variational treatment of the posterior distribution over paths. The method can be understood to be similar to a weak constraint 4DVAR where we seek the best approximating posterior distribution over paths rather than simply the most likely path. The method which we call Bayesian 4DVAR is based on the minimisation of the Kullback-Leibler divergence between distributions, and is suited to applications where simple additive model error in present as a random forcing in the system equations. The approximating distribution used is a Gaussian process, described by a time varying linear dynamical system, whose parameters form the control variables for the problem.

We will outline how this approach can be seen as an extension to weak constraint 4DVAR, where additionally the posterior covariance is approximated. We illustrate the method in operation on a range of toy examples including Lorenz 40D and Kuramoto-Shivashinsky PDE examples. We compare the approach to ensemble and traditional 4DVAR approaches to data assimilation and show its limitations. A principle limitation is that the method systematically underestimates the marginal (with respect to time) state covariance, although we show empirically that this effect is minor given sufficient observations. We discuss possible extensions based on a mean field approximation that will allow the application of the method to large systems. We also show how a local parametrisation of the time varying state between observations using an orthogonal polynomial basis allows further reduction in the number of parameters that need to be estimated.