Geophysical Research Abstracts
Vol. 19, EGU2017-4827, 2017
EGU General Assembly 2017

Scalable and balanced dynamic hybrid data assimilation

Tuomo Kauranne (1), Idrissa Amour (3), Martin Gunia (1), Kari Kallio (2), Ahti Lepistö (2), and Sampsa Koponen (2)

(1) Lappeenranta University of Technology, LENS, Lappeenranta, Finland (Tuomo.Kauranne@lut.fi), (2) Finnish Environment Institute, Helsinki, Finland, (3) University of Dar es Salaam, Dar es Salaam, Tanzania

Scalability of complex weather forecasting suites is dependent on the technical tools available for implementing highly parallel computational kernels, but to an equally large extent also on the dependence patterns between various components of the suite, such as observation processing, data assimilation and the forecast model. Scalability is a particular challenge for 4D variational assimilation methods that necessarily couple the forecast model into the assimilation process and subject this combination to an inherently serial quasi-Newton minimization process.

Ensemble based assimilation methods are naturally more parallel, but large models force ensemble sizes to be small and that results in poor assimilation accuracy, somewhat akin to shooting with a shotgun in a million-dimensional space. The Variational Ensemble Kalman Filter (VEnKF) is an ensemble method that can attain the accuracy of 4D variational data assimilation with a small ensemble size. It achieves this by processing a Gaussian approximation of the current error covariance distribution, instead of a set of ensemble members, analogously to the Extended Kalman Filter EKF. Ensemble members are re-sampled every time a new set of observations is processed from a new approximation of that Gaussian distribution which makes VEnKF a dynamic assimilation method. After this a smoothing step is applied that turns VEnKF into a dynamic Variational Ensemble Kalman Smoother VEnKS. In this smoothing step, the same process is iterated with frequent re-sampling of the ensemble but now using past iterations as surrogate observations until the end result is a smooth and balanced model trajectory.

In principle, VEnKF could suffer from similar scalability issues as 4D-Var. However, this can be avoided by isolating the forecast model completely from the minimization process by implementing the latter as a wrapper code whose only link to the model is calling for many parallel and totally independent model runs, all of them implemented as parallel model runs themselves. The only bottleneck in the process is the gathering and scattering of initial and final model state snapshots before and after the parallel runs which requires a very efficient and low-latency communication network. However, the volume of data communicated is small and the intervening minimization steps are only 3D-Var, which means their computational load is negligible compared with the fully parallel model runs.

We present example results of scalable VEnKF with the 4D lake and shallow sea model COHERENS, assimilating simultaneously continuous in situ measurements in a single point and infrequent satellite images that cover a whole lake, with the fully scalable VEnKF.