Building Ensemble-Based Data Assimilation Systems

Lars Nerger and Paul Kirchgessner

Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research, Bremerhaven, Germany
Contact: Lars.Nerger@awi.de; Paul.Kirchgessner@awi.de  ·  http://www.awi.de

Introduction

We discuss different strategies for implementing ensemble-based data assimilation systems. Ensemble filters like ensemble Kalman filters and particle filters can be implemented so that they are nearly independent from the model into which they assimilate observations.

Offline coupling through disk files avoids changes to the numerical model, but is computationally not efficient. An online coupling strategy is computationally efficient. In this coupling strategy, subroutine calls for the data assimilation are directly inserted into the source code of an existing numerical model and augment the numerical model to become a data assimilative model.

Using the example of the parallel data assimilation framework (PDAF, http://pdaf.awi.de) and the ocean model NEMO, we demonstrate how the online coupling can be achieved with minimal changes to the numerical model.

Assimilation experiments are performed with a box configuration (SEABASS) of NEMO that simulates a double-gyre (see [1]). The configuration is one of the benchmarks of the SANGOMA project. To simulate a high-dimensional model, the resolution is increased to 1/12°. The grid has 361 x 241 grid points and 11 layers. The state vector has a size of about 3 million.

Synthetic observations of sea surface height at ENVISAT and Jason-1 satellite tracks and temperature profiles on a 3’ x 3’ grid are assimilated each 48 hours over 360 days. Observation errors are respectively set to 5 cm and 0.3°C. The assimilation uses the local ESTKF filter [4].

Parallel Performance of Online Coupling

The parallel compute performance of the assimilation system is described by the speedup (ratio of the computing time on n processes to the time on one process). The speedup of the assimilation system is dominated by the speedup of the NEMO model itself. The assimilation slightly increases the speedup due to a better scalability.

The online coupling shows a good computational scalability on supercomputers and is hence well suited for high-dimensional numerical models, including coupled earth system models.

Further, a clear separation of the model and data assimilation components allows to continue the development of both components separately.

Summary

Implementations using online coupling have been performed also for other models like FESOM, BSHmod, HBIM, NOBM, ADCIRC, and MITgcm. PDAF is coded in Fortran with MPI parallelization. It is available as free software. Further information and the source code of PDAF are available on the web site:

http://pdaf.awi.de

References