Building Ensemble-Based Data Assimilation Systems with Coupled Models

Lars Nerger, Dmitry Sidorenko

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

Overview

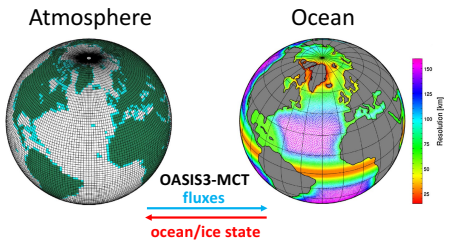
We discuss a strategy to modify a coupled model so that we can use it for efficient ensemble data assimilation. The method uses a direct connection between a coupled model and the ensemble data assimilation framework PDAF [1, http://pdaf.awi.de]. The strategy allows us to set up a data assimilation program with high flexibility and parallel scalability with only small changes to the model.

The direct connection is obtained by

1. adapting the source codes of the coupled model so that it is able to run an ensemble of model states
2. adding a filtering step to the source codes.

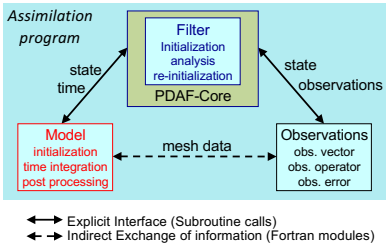
We discuss this connection for the coupled atmosphere-ocean model AWI-CM. For this coupled model, we have to augment the codes of both the ocean and atmosphere, adapt the parallelization, and add routines for the handling of observations and model fields specific for each model compartment.

Coupled Model: AWI-CM



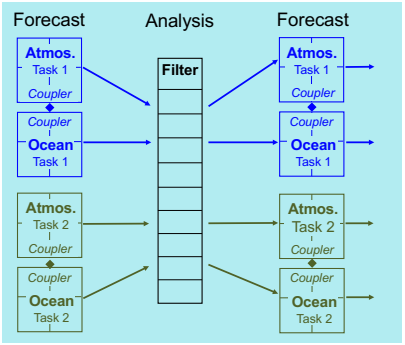
The coupled model AWI-CM [2] consists of ECHAM6 for the atmosphere including the land model JSBACH, and the finite-element sea ice-ocean model FESOM for the ocean compartment. The models are separate programs. They are coupled with OASIS3-MCT. Fluxes between the models are computed each 6 hours by OASIS3-MCT using the fields from FESOM.

Data Assimilation System



The data assimilation system can be separated into three components: Model, filter algorithm, and observations. The filter algorithms are model-independent, while the model and subroutines to handle observations are provided by the user. The routines are either directly called in the program code or share information, e.g., through Fortran modules.

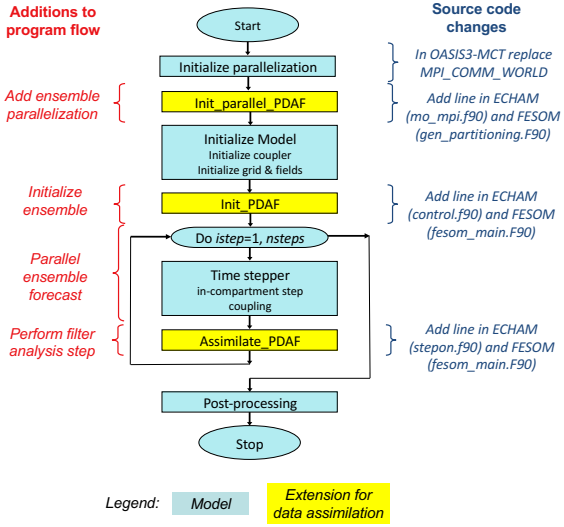
Coupled Ensemble Forecasts



Example of an ensemble integration with two ensemble members. Both models and the filter are parallelized. The ensemble adds one level of parallelization to integrate all members at once.

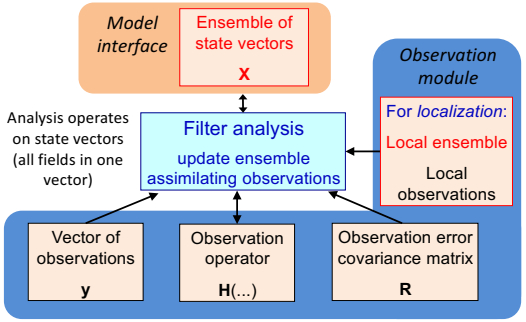
Augmenting the Model Codes

To augment the coupled model with data assimilation functionality, three subroutine calls for PDAF are inserted into the source codes of ECHAM6 and FESOM. Further, we need to replace a communicator in OASIS3-MCT [3] so that each coupled ensemble task is treated separately by the coupler.



Call-back Routines for Analysis Step

The filter analysis step needs information on the assimilated observations. Further, model fields need to be written into the state vectors. The functionality is provided by call-back routines. The programs of the atmosphere and ocean models use distinct user routines for handling observations and model fields.



References:

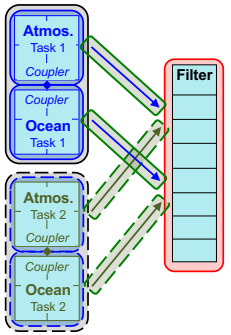
[1] Nerger, L., Hiller, W. Software for Ensemble-based Data Assimilation Systems - Implementation Strategies and Scalability. Comp. & Geosci., (2013) 55: 110-118

[2] Sidorenko, D. et al. Towards multi-resolution global climate modeling with ECHAM6–FESOM. Part I: model formulation and mean climate, Clim. Dyn. (2015) 44:757–780

[3] Kurtz, W. et al. TerrSysMP–PDAF (version 1.0): a modular high-performance data assimilation framework for an integrated land surface–subsurface model, Geosci. Model Dev., (2016) 9: 1341–1360

Configuring the Parallelization

The parallelization is adapted to enable the coupled ensemble integrations, field exchanges between model and filter, and the computation of the filter step. Usually, we use the processors of the model task 1 to compute the filter.



Decomposition into process groups using parallel (MPI) communicators:

- Coupled model task
- Compartment in each task (created by coupler)
- Filter (1 for strongly, 2 for weakly coupled assimilation)
- Connection for collecting ensembles for filtering (for each model sub-domain)

Summary

The discussed strategy to build the data assimilation system uses a combination of in-memory access and parallel communication to create a particularly efficient online-coupled ensemble assimilation program.

The analysis step is computed in between time steps. It is independent of the actual model coupler. There is no need to write the ensemble into files and no need to restart the model.

Care needs to be taken when implementing the model interface and observation handling routines, which are specific to the two programs for atmosphere and ocean.