Building an Efficient Ensemble Data Assimilation System for Coupled Models with the Parallel Data Assimilation Framework

Lars Nerger, Qi Tang, Longjiang Mu, Dmitry Sidorenko

Alfred-Wegener Institute Helmholtz Center for Polar and Marine Research, Bremerhaven, Germany

Contact: Lars.Nerger@awi.de http://www.awi.de

PDAF is open source. The code, documentation, and this poster are available at **http://pdaf.awi.de**

Overview

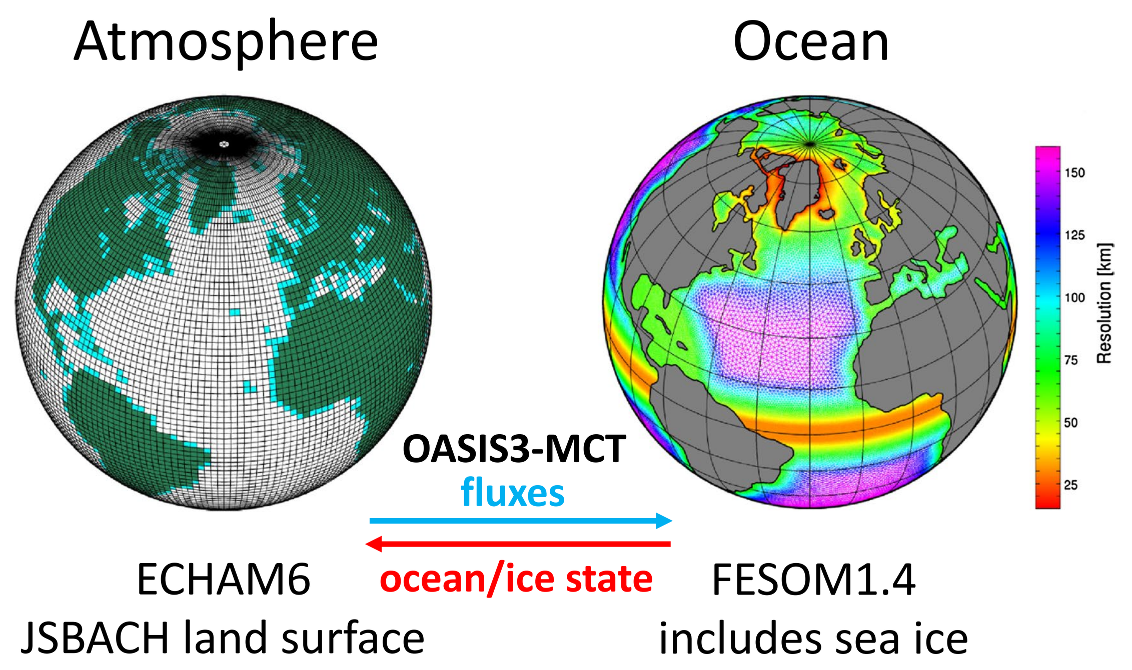
We show how to modify a coupled model so that we can use it for efficient ensemble data assimilation. We use a direct connection between the coupled model and the ensemble data assimilation framework PDAF [1]. Augmenting the model allows us to set up a data assimilation program with high flexibility and parallel scalability with only small changes to the model.

Data assimilation in the coupled model 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.**

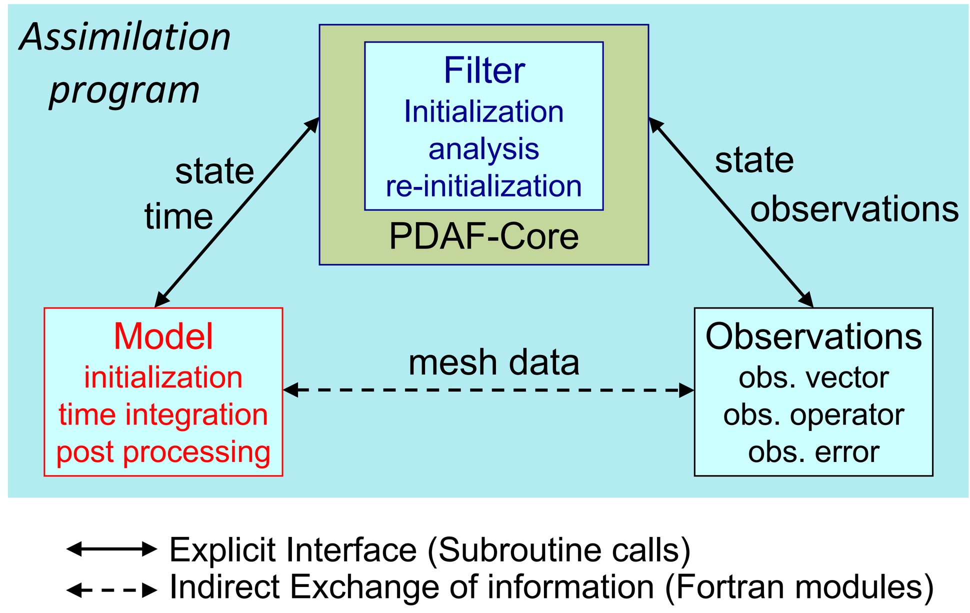
We discuss this connection for the coupled atmosphere-ocean model AWI-CM. We augment the model codes of both the ocean and atmosphere, adapt the parallelization, and add routines for the handling of observations and model fields specific for each model compartment.

Example Coupled Model: AWI-CM



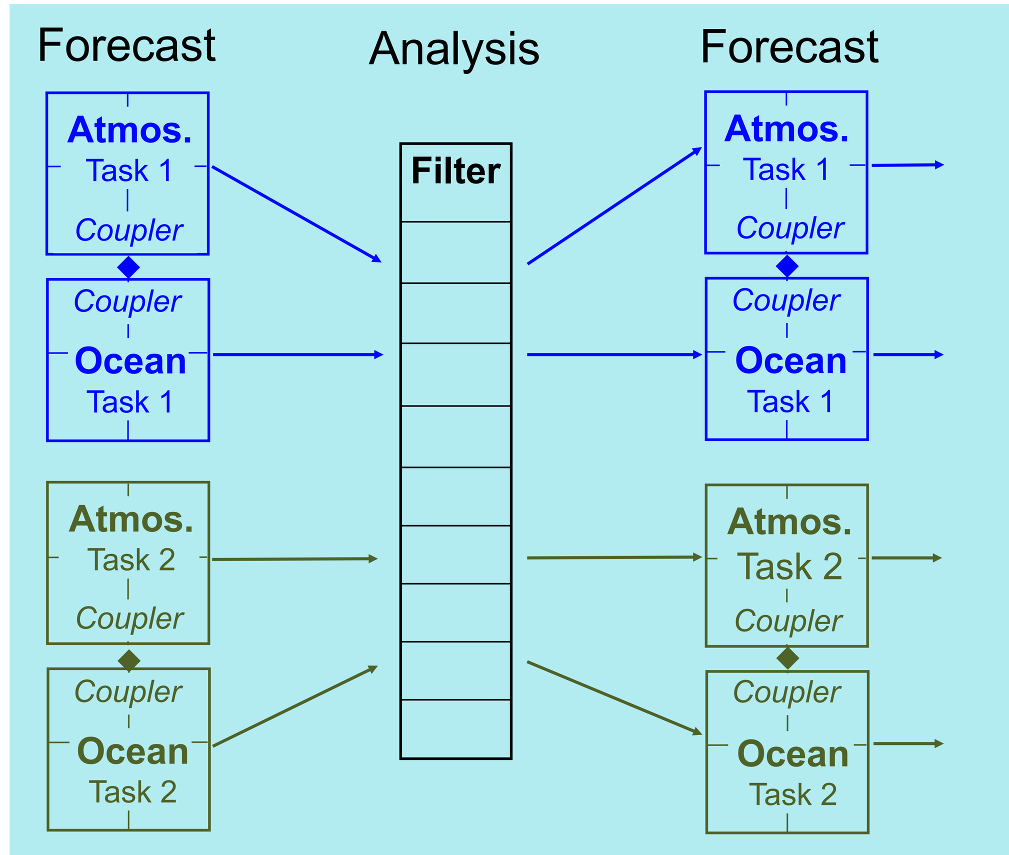
AWI-CM [2] consists of two separate programs: FESOM and ECHAM6. Both are coupled with OASIS3-MCT and run in parallel. Fluxes between the models are computed and exchanged each 6 hours by OASIS3-MCT using parallel communication.

Data Assimilation Program



The data assimilation system has three components: Model, filter algorithm, and observations. The filter algorithms are model-agnostic, while the model and subroutines to handle observations are provided by the user. The observation routines are called by PDAF as call-back routines.

Coupled Ensemble Forecasts

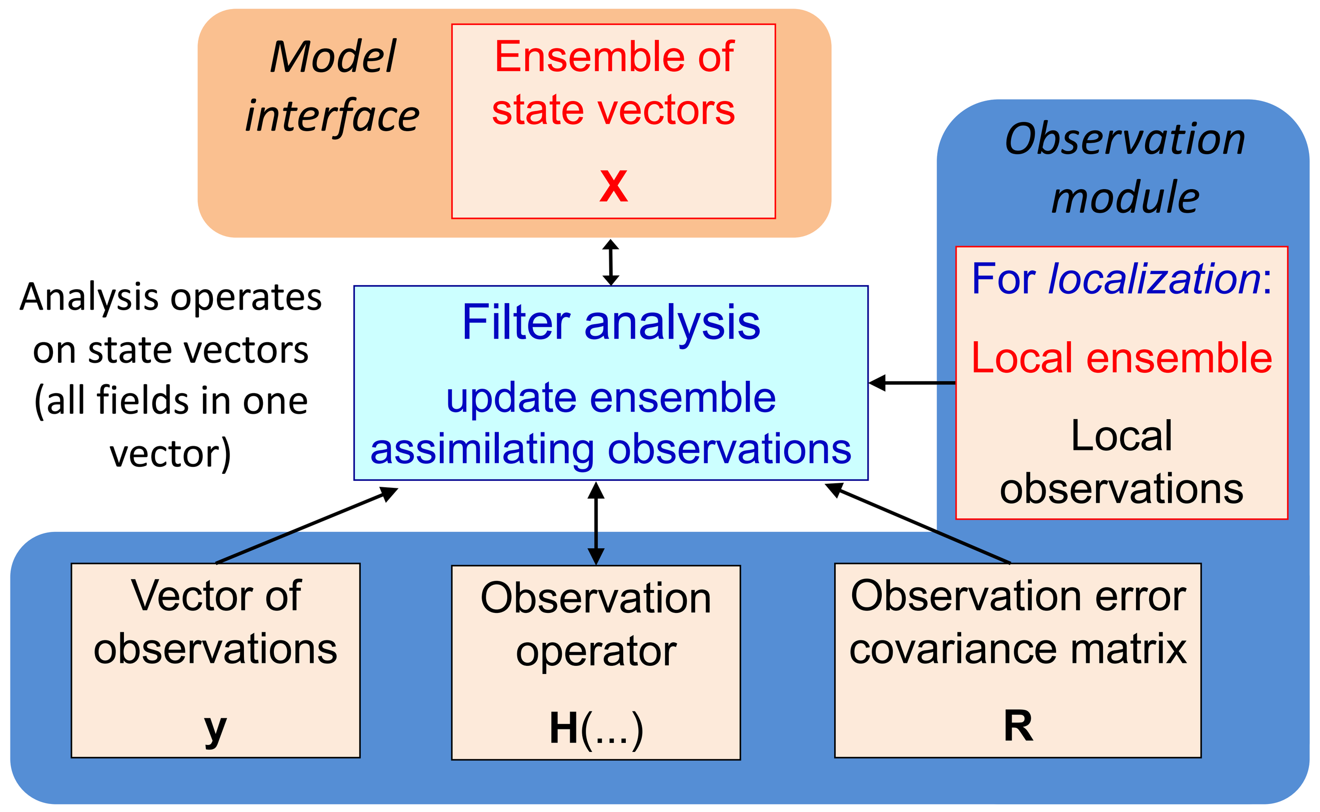


Example of an ensemble integration with two ensemble members. Both models and the filter are parallelized. The ensemble adds one level of parallelization to integrate all members at once.

PDAF lets you easily build a highly efficient program for ensemble data assimilation

Call-back Routines for Analysis Step

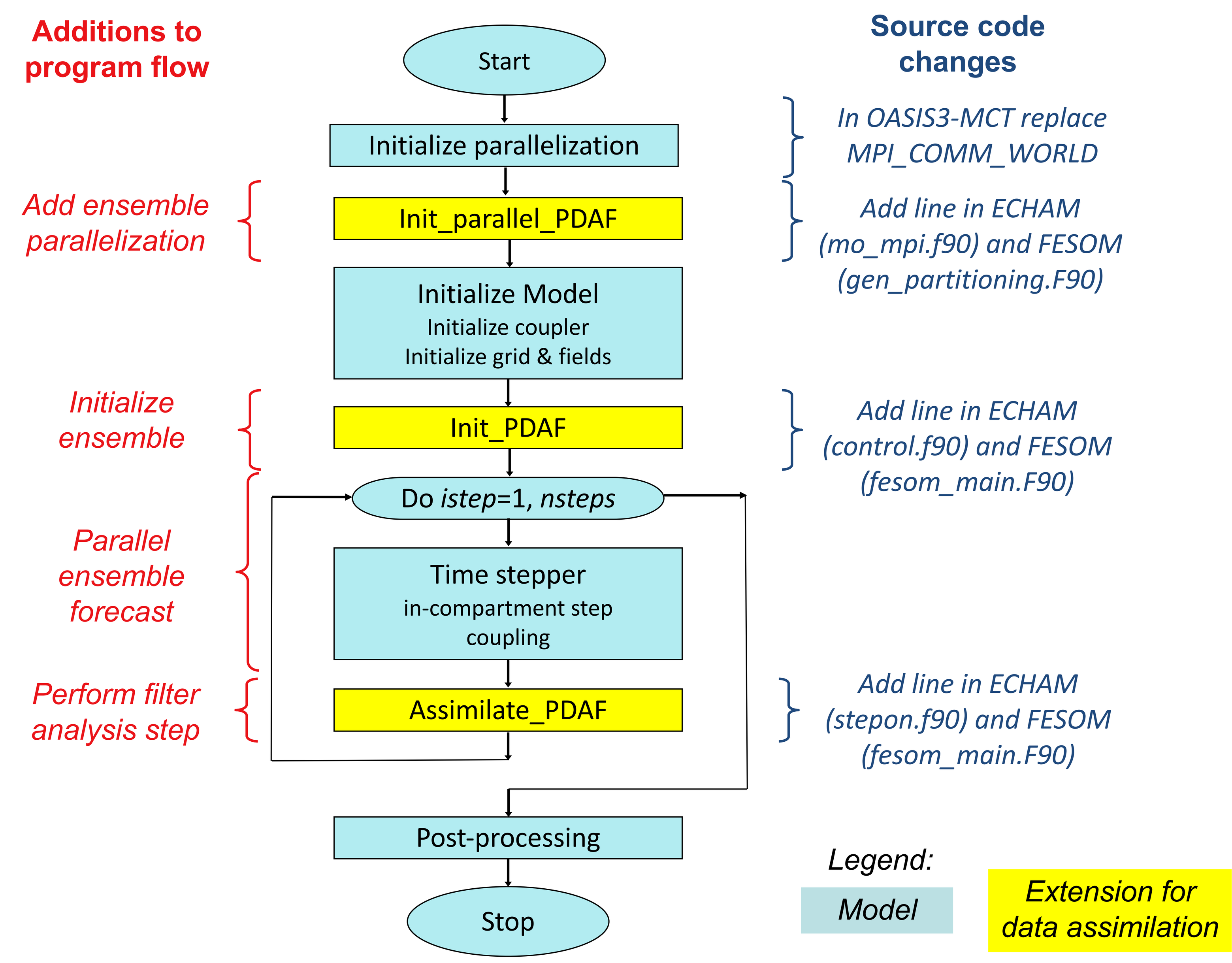
Model fields need to be written into the state vectors and back. The filter analysis step needs information on the assimilated observations. PDAF uses call-back routines for this. The programs of the atmosphere and ocean models use distinct user routines for handling observations and model fields.



Adapting the Model Codes

We insert three subroutine calls for PDAF into the source codes of ECHAM6 and FESOM to add data assimilation functionality to the coupled model.

Further, we need to replace a communicator in OASIS3-MCT so that it treats each coupled ensemble task separately.



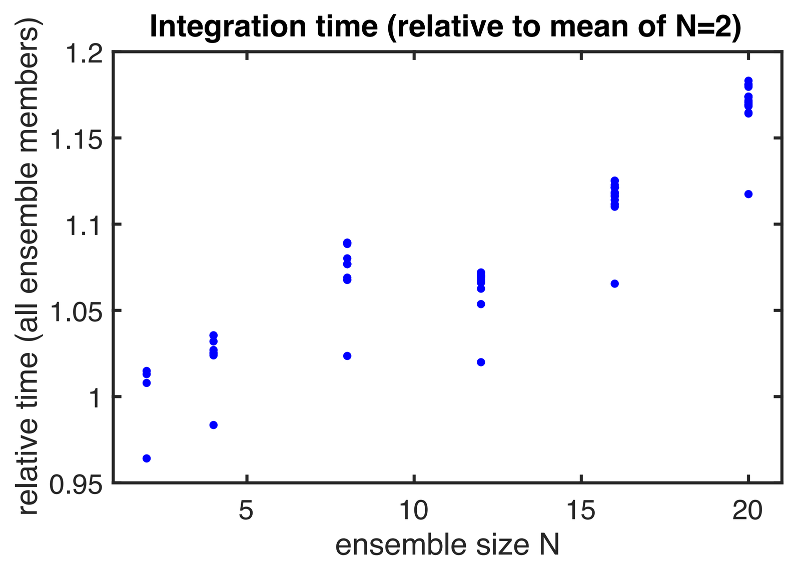
Compute Performance

The experiment

- Weakly-coupled assimilation into the ocean
- State vector: ocean surface height, temperature, salinity, velocities
- Ensemble size: up to 23 state realizations
- Assimilation method: Local Error-Subspace Transform Kalman Filter (LESTKF)
- Simulation period: full year 2016, daily assimilation update

Compute Performance

- Run time for ensemble size 23: 6.5 hours (fully parallelized on 12,144 processors)
- Scaling test: increase ensemble size and number of processors
    - Slightly different forecast duration for each ensemble member
    - Run time only increases by 17% for 10-fold ensemble size



Summary

- Using PDAF we add data-assimilation functionality to the model to build a data assimilation program.
- PDAF uses in-memory access and parallelization to ensure high efficiency.
- The addition is independent of the actual model coupler.
- The analysis step is computed in between time steps without stopping the program. There is no need to write the ensemble into files.
- Routines for the model interface and observation handling need to be implemented for each of the two programs for atmosphere and ocean.

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