Extending NEMO for Ensemble Data Assimilation on Supercomputers with the Parallel Data Assimilation Framework PDAF

Alfred Wegener Institute
Helmholtz Center for Polar and Marine Research

Lars Nerger and Paul Kirchgessner

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

SANGOMA

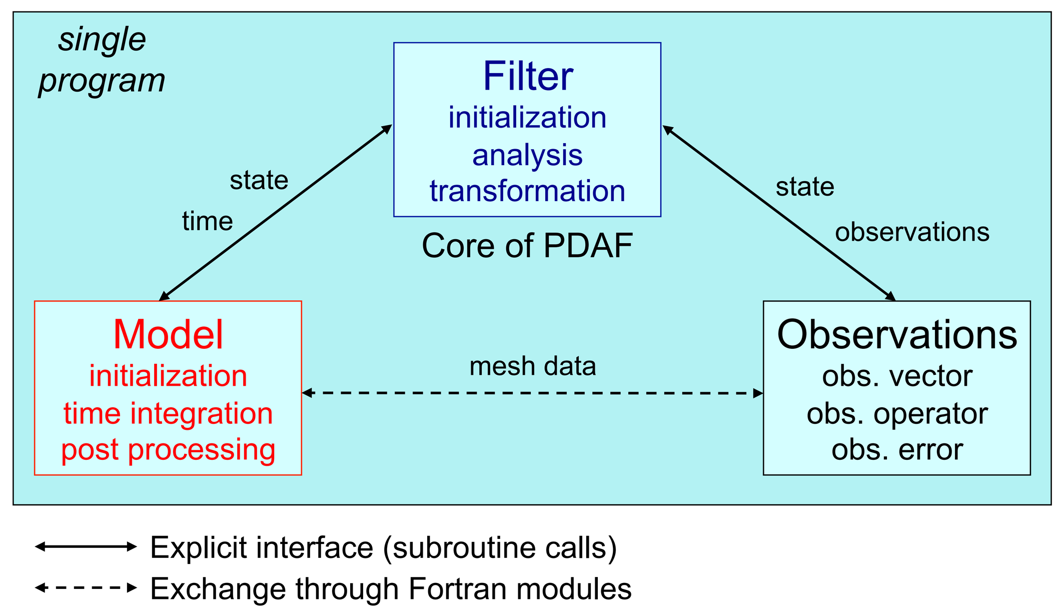
Introduction

We introduce a data assimilation system for the ocean circulation model NEMO that is built using the parallel data assimilation framework PDAF [http://pdaf.awi.de]. Inserting three subroutine calls to the source code of NEMO, one extends NEMO to a data assimilation system that consists of a single program. Utilizing the parallelization capacity of today's supercomputers, the system performs both the ensemble forecasts and the analysis step of the filter algorithm in a single execution of the program.

This system is in contrast to other assimilation systems that run NEMO separately from the assimilation algorithms. These exchange data through disk files holding the ensemble of model states. Using the online-coupled data assimilation system with PDAF, repeated storage of all ensemble states in disk files is avoided. In addition, re-starting the model is not required. These features lead to a computationally very efficient data assimilation program. A square-box configuration of NEMO [see 1] is used to test the assimilation system.

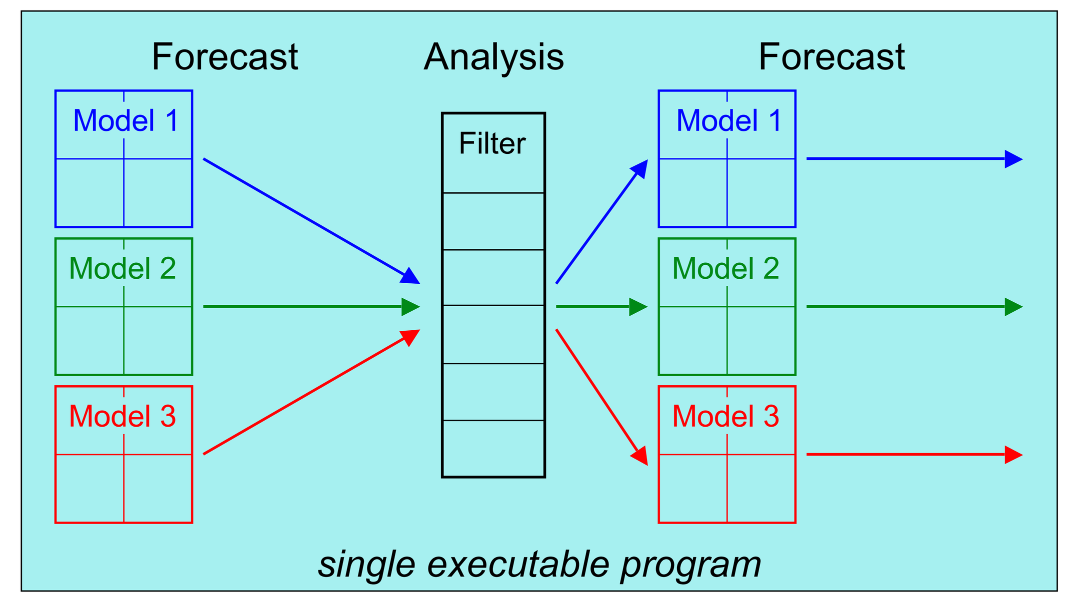
A Parallel Data Assimilation System

Logical separation of the assimilation system



PDAF separates the data assimilation system into three components: Model, filter algorithm, and observations. The filter algorithms are part of PDAF's core, while the model and subroutines to handle observations are provided by the user. A standard interface for all filter algorithms connects the three components. All user-supplied subroutines can be implemented like model routines.

2-level parallelization of the assimilation system

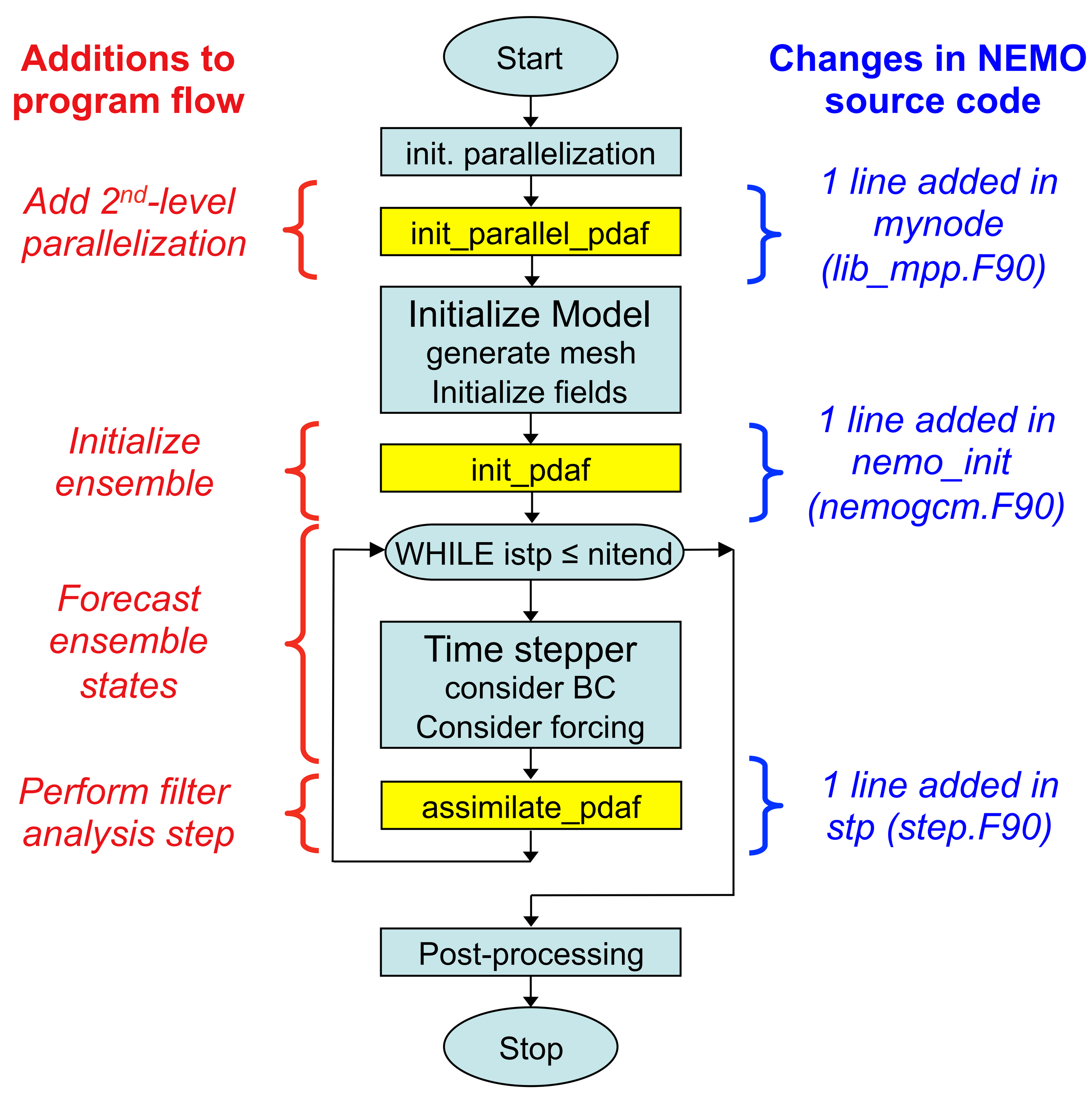


PDAF provides support for a 2-level parallelization of the assimilation system:

1. Each model task can be parallelized.
2. All model tasks are executed concurrently.

Thus, ensemble integrations can be performed fully parallel. In addition, the filter analysis step uses parallelization. All components are combined in a single program.

Coupling NEMO with PDAF



Legend: Model | *Extension for data assimilation*

NEMO is coupled with PDAF [2,3] by adding three subroutine calls the model source code and utilizing parallelization. In contrast to other frameworks, the model does not need to exist as a separate subroutine. Model- and observation-specific operations are performed in user-supplied call-back routines that are called through PDAF. The ensemble forecast is also controlled by user-supplied routines.

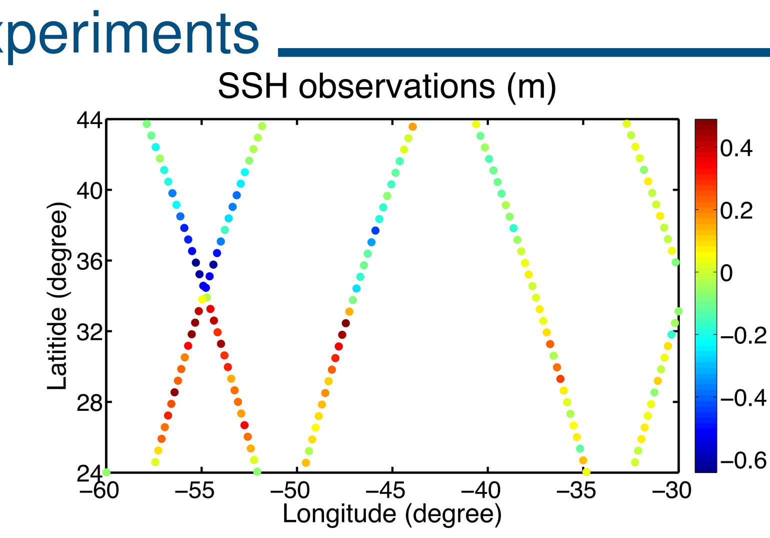
Implementations using this online coupling scheme have been performed also for other models like FESOM, BSHcmod, HBM, NOBM, ADCIRC, and PARODY.

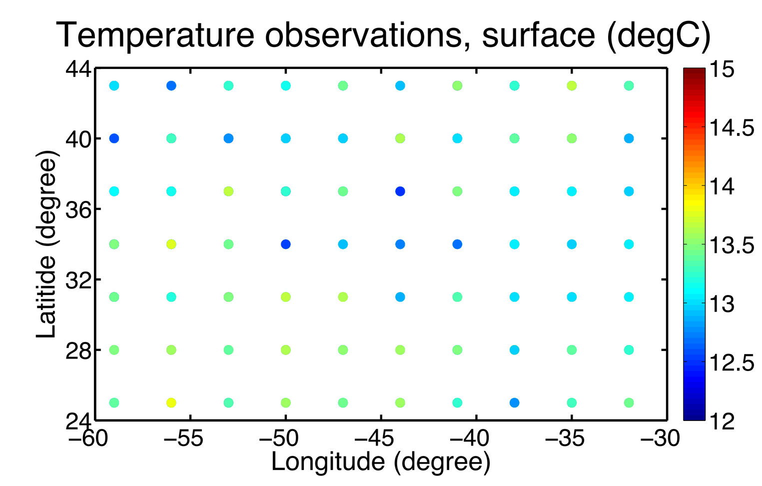
PDAF is coded in Fortran with MPI parallelization. It is available as free software. Further information and the source code of PDAF are available on the web site:

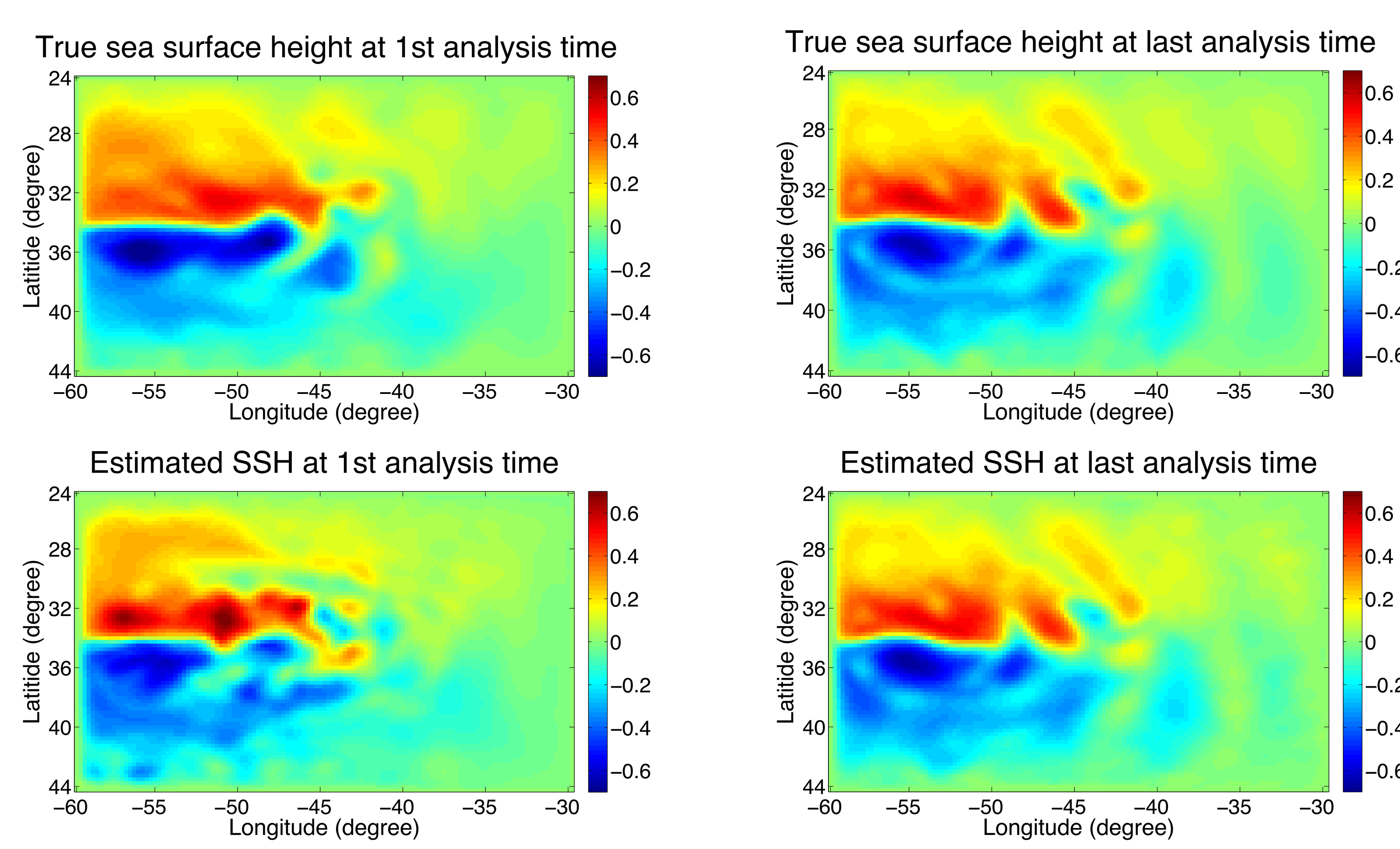
http://pdaf.awi.de

Assimilation Experiments

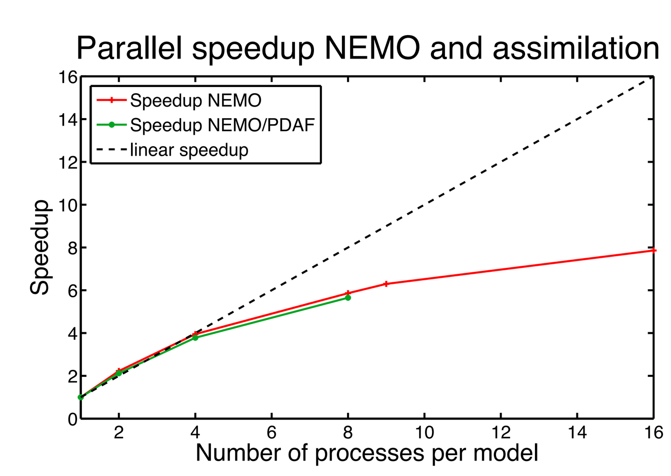
Assimilation experiments are performed to validate the assimilation system. A box configuration of NEMO ("SQB") is used that simulates a double-gyre. The SQB-configuration is one of the benchmarks of the SANGOMA project [http://www.data-assimilation.net]. The grid has 121×81 grid points at a horizontal resolution of 0.25° and 11 layers. Synthetic observations of sea surface height at ENVISAT and Jason-1 satellite tracks and temperature profiles on a $3^\circ\times3^\circ$ grid are assimilated each 48 hours over 360 days. Observation errors are respectively set to 5cm and 0.3°C. The assimilation uses the ESTKF filter [4] with localization [5]. An ensemble of 32 states is used. The errors in all fields are significantly reduced by the assimilation (see SSH below).







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 leads only to a small reduction of the speedup.



References

[1] Cosme E., Brankart J.-M., Verron J., Brasseur P. and Krysta M. (2010). Implementation of a reduced-rank, square-root smoother for high resolution ocean data assimilation. *Ocean Modelling*, 33: 87–100

[2] Nerger, Hiller, and Schröter (2005). PDAF - The Parallel Data Assimilation Framework: Experiences with Kalman Filtering, in *Use of High Performance Computing in Meteorology - Proceedings of the 11th ECMWF Workshop* / Eds. W. Zwieflhofer, G. Mozdzynski. World Scientific, pp. 63–83

[3] Nerger, L. and W. Hiller (2012). Software for Ensemble-based Data Assimilation Systems – Implementation Strategies and Scalability. *Computers & Geosciences*. 55: 110–118

[4] Nerger, L., T. Janjić, J. Schröter, J., and W. Hiller (2012). A unification of ensemble square root Kalman filters. *Mon. Wea. Rev.* 140: 2335–2345

[5] L. Nerger, S. Danilov, W. Hiller, and J. Schröter (2006). Using sea-level data to constrain a finite-element primitive-equation ocean model with a local SEIK filter. *Ocean Dynamics* 56: 634–649