AGU Fall Meeting, Washington, DC, USA, December 9-13, 2024

PDAF Parallel Data Assimilation Framework

Open-Source Developments for Community Data Assimilation Software with PDAF

Lars Nerger[1], Yumeng Chen[2], Qi Tang[3], Changliang Shao[4]

1) Alfred Wegener Institute, Bremerhaven, Germany
2) University of Reading and NCEO, UK
3) University of Neuchâtel & University of Basel, Switzerland
4) China Meteorological Administration, Meteorological Observation Center

PDAF – Parallel Data Assimilation Framework

PDAF Parallel Data Assimilation Framework

A unified tool for interdisciplinary data assimilation …

- provide support for parallel ensemble forecasts
- provide DA methods (EnKFs, smoothers, PFs, 3D-Var) - fully-implemented & parallelized
- provide tools for observation handling and for diagnostics
- easy implementation with (probably) any numerical model (<1 month)
- a program library (PDAF-core) plus additional functions & templates
- run from notebooks to supercomputers (Fortran, MPI & OpenMP)
- ensure separation of concerns (model – DA method – observations – covariances)
- first release in year 2004; continuous further development

Focus on

- Easy implementation
- Performance for complex models
- Flexibility to extend system

Open source:
Code, documentation, and tutorial available at

https://pdaf.awi.de

github.com/PDAF/PDAF

L. Nerger, W. Hiller, Computers & Geosciences 55 (2013) 110-118

PDAF – wide range of applications

PDAF Parallel Data Assimilation Framework

Applications & users, like

Operational uses:

- *Germany:* North/Baltic Seas (HBM model)
- *Europe:* Copernicus marine forecasting center Baltic Sea (NEMO)
- *China:* Arctic ice-ocean prediction system (MITgcm)

Beyond ocean

- **HydroGeoSphere** (hydrology)
- **TSMP** (Terrestrial Systems Modeling Platform)
- **WRF** (Weather forecast and research model)
- **TIE-GCM** (Thermosphere Ionosphere Electrodynamics)
- **VILMA** (Viscoelastic Lithosphere and Mantle Model)
- **Parody** (Geodynamo model)
- **HYSPLIT** (Volcanic Ash Transport and Dispersion model)
- **Cardiatic modeling** (blood flow)
- … more

Ocean and marine biogeochemistry (research applications)

- **CICE** (sea ice)
- **COAWST** (WRF-ROMS-CICE)
- **FESOM**
- **HBM** (regional HiROMB-BOSS model)
- **MEDUSA** (biogeochemistry)
- **MITgcm**
- **NEMO**
- **REcoM** (biogeochemistry, carbon cycle)
- **SCHISM/ESMF**

Coupled ocean-atmosphere & climate models

- **AWI-CM**
- **MPI-ESM** (ICON-Ocean)
- **CLIMBER-X** (paleo climate model)

PDAF – wide range of applications

PDAF Parallel Data Assimilation Framework

Applications & users, like

Operational uses:
- *Germany:* North/Baltic Seas (HBM model)
- *Europe:* Copernicus marine forecasting center Baltic Sea (NEMO)
- *China:* Arctic ice-ocean pre… (MITgcm)

Beyond ocean
- **HydroGeoSphere** (hydrolog…
- **TSMP** (Terrestrial Systems M…
- **WRF** (Weather forecast and research model)
- **TIE-GCM** (Thermosphere Ionosphere Electrodynamics)
- **VILMA** (Viscoelastic Lithosphere and Mantle Model)
- **Parody** (Geodynamo model)
- **HYSPLIT** (Volcanic Ash Transport and Dispersion model)
- **Cardiatic modeling** (blood flow)
- … more

Different models – same assimilation software

leverage synergies

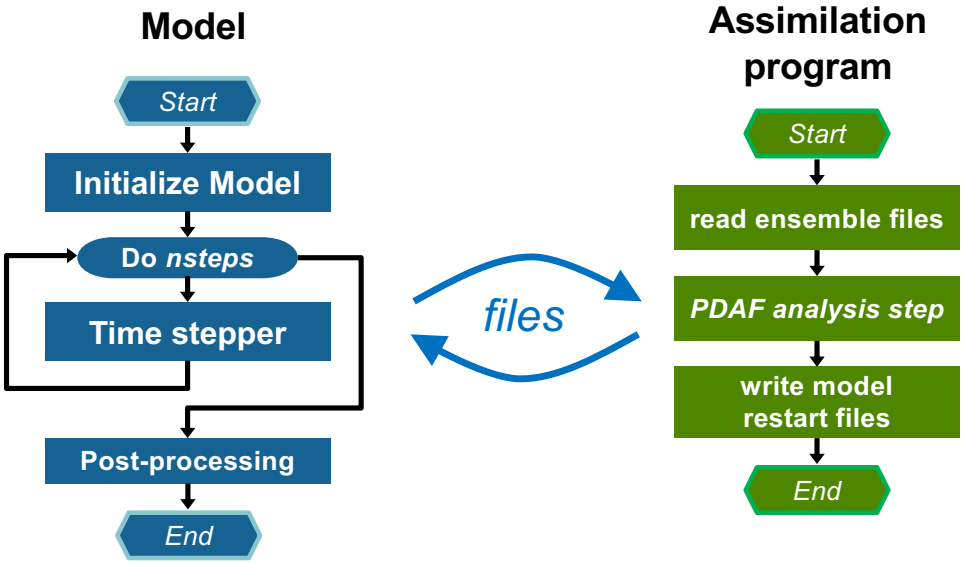
Ocean and marine biogeochemistry (research applications)
- **CICE** (sea ice)
- **COAWST** (WRF-ROMS-CICE)
- …al HiROMB-BOSS model)
- …ogeochemistry)
- …geochemistry, carbon cycle)
- **SCHISM/ESMF**

Coupled ocean-atmosphere & climate models
- **AWI-CM**
- **MPI-ESM** (ICON-Ocean)
- **CLIMBER-X** (paleo climate model)

Coupling Model and Assimilation Code: 2 Variants

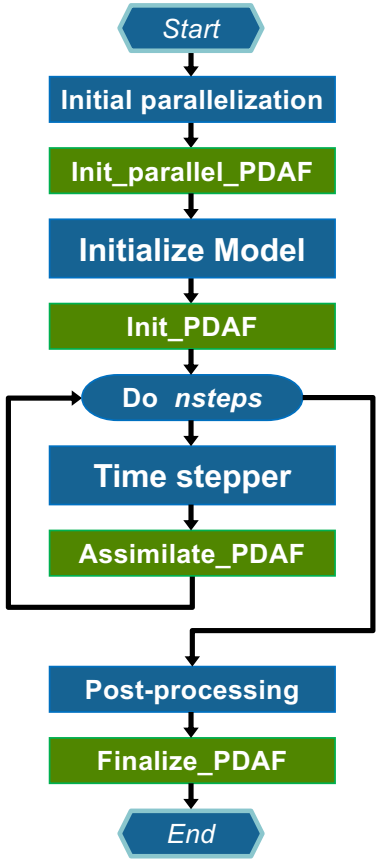
PDAF Parallel Data Assimilation Framework

Offline coupling



- separate programs for model and assimilation
- standard setup at many operational centers
- can be computed by group of compute jobs
- limited computational efficiency
  - model restarts
  - amount of disk IO

Online coupling



- augment model with assimilation functionality
- adapt model source code adding 4 routine calls
- easy to implement
- one large compute job
- in-memory data transfers
- computationally very efficient
- For coupled Earth system models: independent of model coupler

Model code | DA code

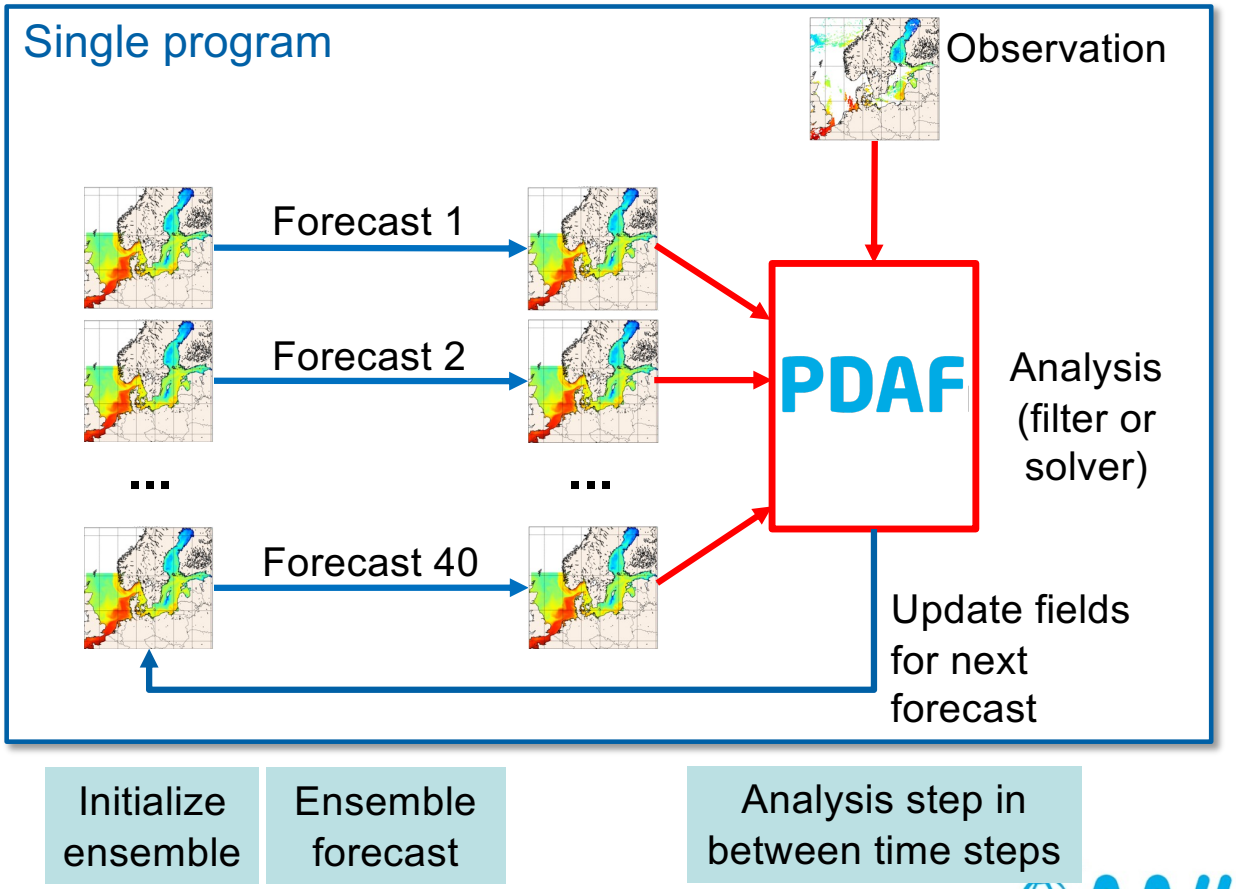
Online-Coupling – Assimilation-enabled Model

PDAF Parallel Data Assimilation Framework

Couple a model with PDAF

- Modify model to simulate ensemble of model states
- Insert analysis step/solver to be executed at prescribed interval
- Run model as usual, but with more processors and additional options (easy to use…)

- EnOI and 3D-Var also possible:
  - Evolve single model state
  - Prescribe ensemble perturbations or covariance

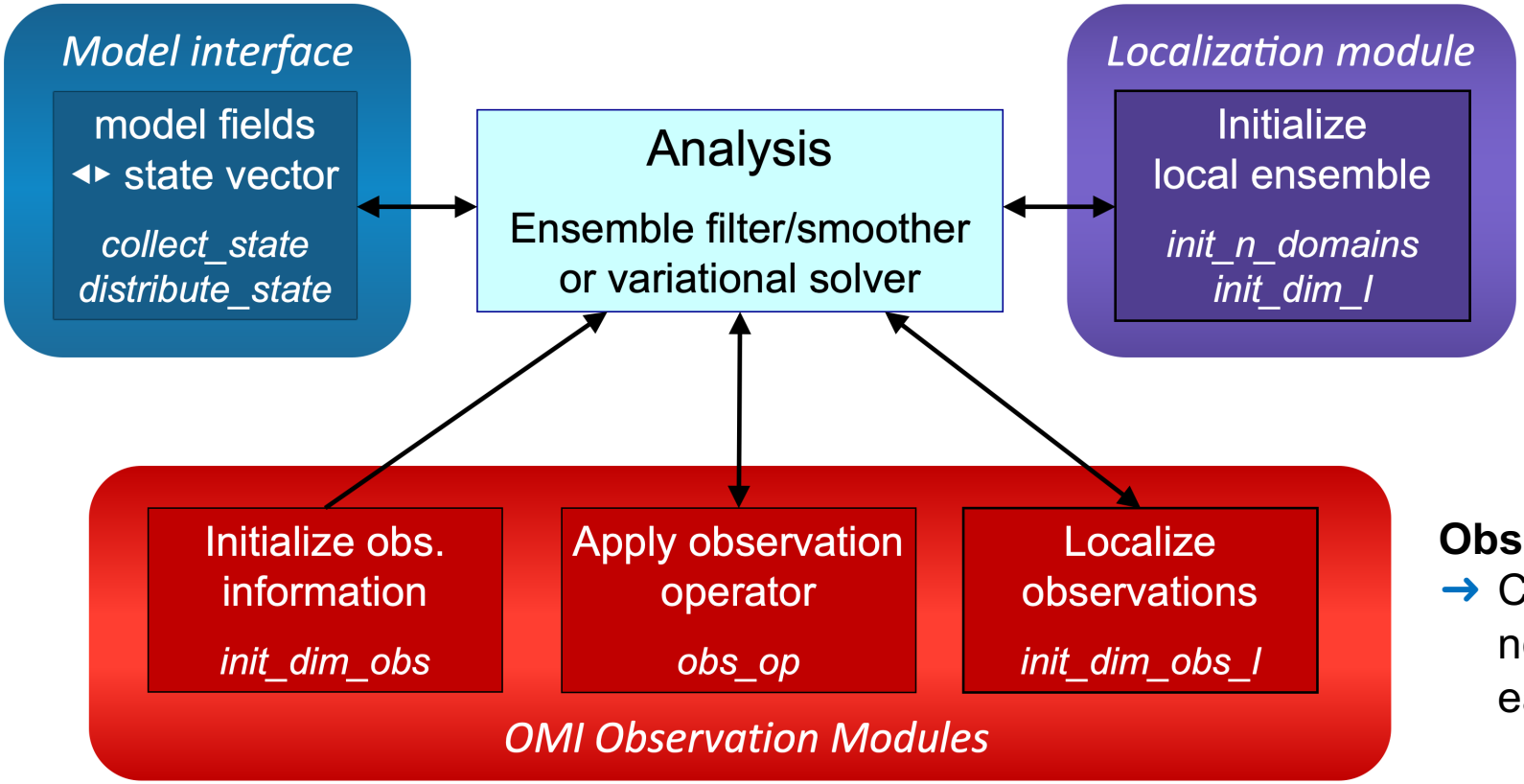


Implementing the Assimilation Analysis/Solver Step

PDAF Parallel Data Assimilation Framework

Analysis steps needs user-provided routines

- Easy to implement using templates and provided functionality



Observation modules

→ Class-like structure, but not object-oriented for easier implementation

pyPDAF

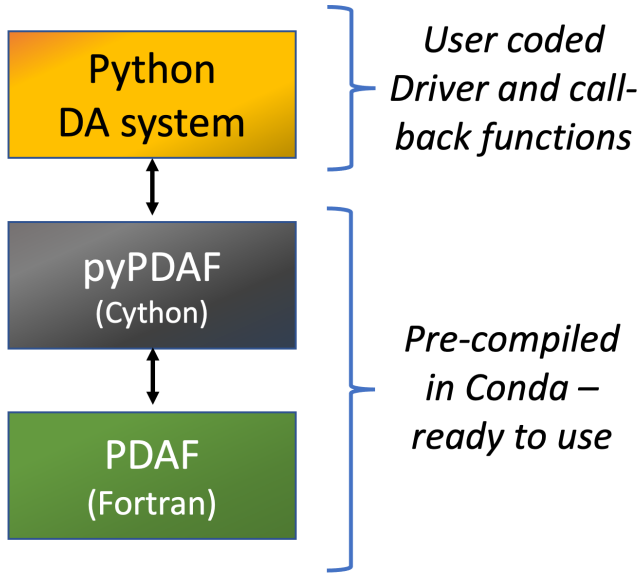
PDAF Parallel Data Assimilation Framework

Python interface to PDAF

- case-specific functions can be implemented in Python (not touching Fortran!)
- assimilation analysis computed inside PDAF (excellent performance due to compiled Fortran code)
- supports
  - Online coupling (e.g. for Python-coded models)
  - Offline coupling (using files from model runs)

- installation using Conda

  conda create -n pypdaf -c conda-forge yumengch::pypdaf



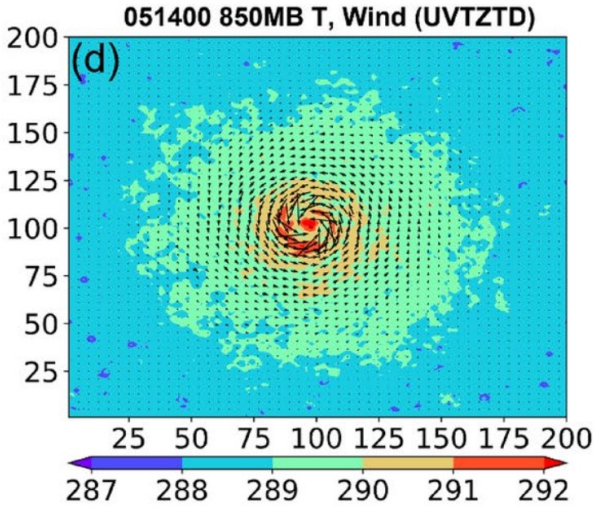
github.com/yumench/pyPDAF

New model couplings: WRF, HydroGeoSphere

PDAF Parallel Data Assimilation Framework

WRF (Weather Research and Forecast model)
Changliang Shao

Example: idealized tropical cyclone assimilating GNSS data

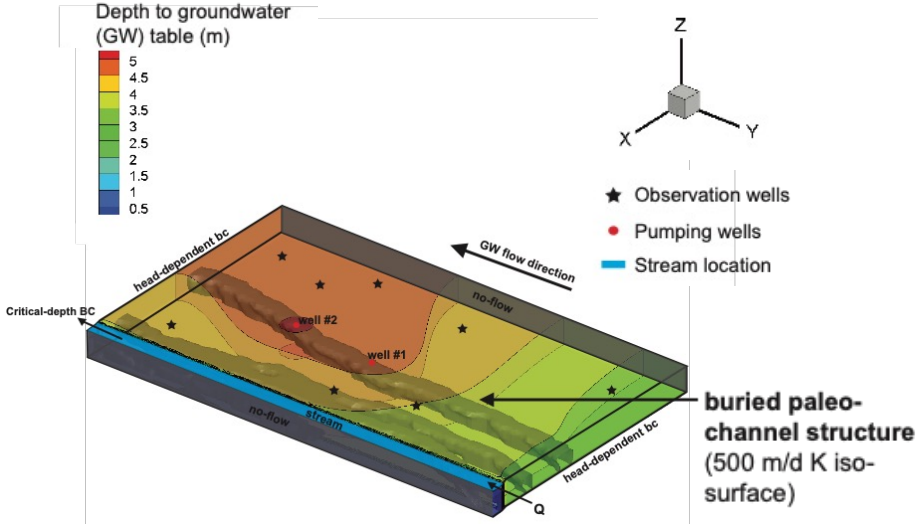


Open source:
https://doi.org/10.5281/zenodo.8367112

Shao & Nerger, *GMD,* 17, 4433, 2024, doi:10.5194/gmd-17-4433-2024

HydroGeoSphere (HGS)
Qi Tang

Example: synthetic 3-D river-aquifer model



Open source:
https://github.com/qiqi1023t/HGS-PDAF_v1.0

Tang et al., *GMD*, 17, 3559, 2024, doi:10.5194/gmd-17-3559-2024

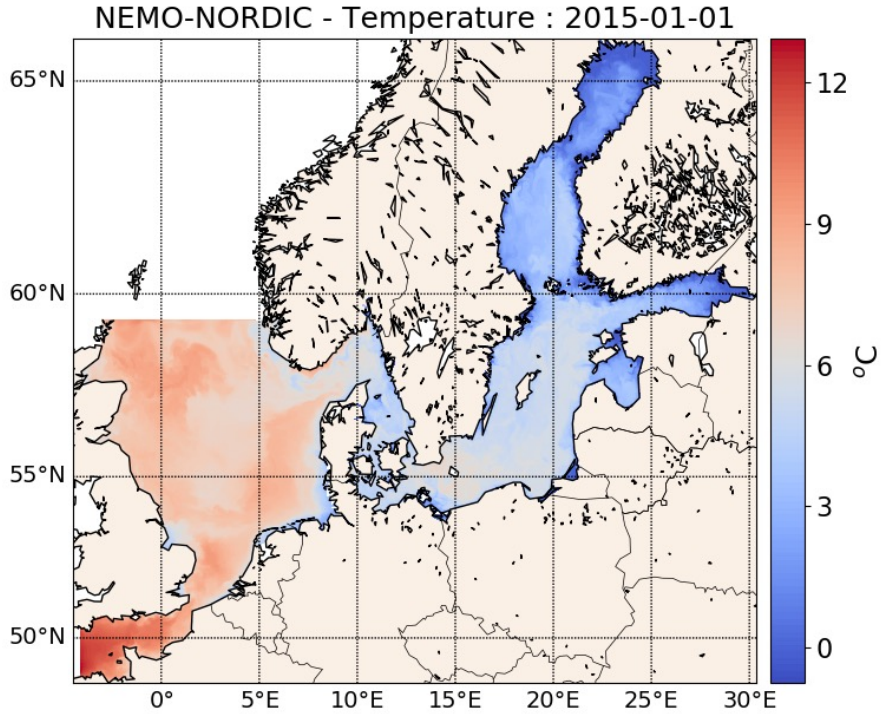
Model couplings: NEMO

PDAF Parallel Data Assimilation Framework

NEMO-PDAF

Used **operationally for ocean forecasting in the European Copernicus Marine Service.** Ocean physics and biogeochemistry for the Baltic Sea (CMEMS Forecasting Center BAL-MFC) and North Sea

Used for research regionally and in global setups (with MEDUSA biogeochemical model)



github.com/PDAF/NEMO-PDAF

This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004032.

SEAMLESS

AWI

PDAF code: DA Methods and Models

PDAF originated from comparison studies of different filters

Ensemble Filters and smoothers - *global and localized*

- EnKF (Evensen, 1994 + perturbed obs.)
- (L)ETKF (Bishop et al., 2001/Hunt et al. 2007)
- ESTKF (Nerger et al., 2012)
- NETF (Toedter & Ahrens, 2015)
- Particle filter
- Hybrid Nonlinear Kalman Filter LKNETF (Nerger, 2022)
- *EnOI mode*

3D-Var schemes

- 3D-Var with parameterized covar.
- 3D Ensemble Var
- Hybrid 3D-Var

Toy models (full implementations with PDAF)

- Lorenz-96
- Lorenz-63
- Lorenz-2005 models II and III

Model bindings

By PDAF core group:

- MITgcm
- AWI-CM / FESOM
- NEMO

Community provided:

SCHISM/ESMF
TerrSysMP-PDAF
EAT - GOTM/FABM/BGC
WRF

See full list on https://pdaf.awi.de

Summary

PDAF Parallel Data Assimilation Framework

PDAF

- unified tool for interdisciplinary data assimilation
- provide range of assimilation methods
  - ensemble and variational
- simplifies application of data assimilation
  - Easy coupling to models
  - Easy implementation of observations
- computationally very efficient
  - applicable from small models on notebook computers to complex models running on supercomputers

https://pdaf.awi.de

Thank you!

Lars.Nerger@awi.de – Data Assimilation and PDAF