PDAF - Recent Developments for Open-Source Community Data Assimilation

Lars Nerger[1], Yumeng Chen[2], Armin Corbin[3], Johannes Keller[4]

[1]*Alfred Wegener Institute, Bremerhaven, Germany*
[2]*University of Reading and NCEO, Reading, UK*
[3]*University of Bonn, Bonn, Germany*
[4]*Research Center Jülich, Jülich, Germany*

Abstract

We will provide an overview of the Parallel Data Assimilation Framework (PDAF, https://pdaf.awi.de), and its recent developments, in particular related to the major revision upgrade V3. PDAF is a widely used community open-source software framework for data assimilation (DA). PDAF supports algorithmic research and the application of DA with complex models, and is also used operationally. PDAF provides functionality for ensemble forecasting, DA methods, diagnostics, and a sophisticated approach to handle a wide variety of observation types. PDAF can be used offline with a separate DA program or online by augmenting the model code to turn it into an ensemble forecasting model with DA functionality, which leads to very high computational efficiency. The online coupling to a model with minimal changes to the model code usually takes less than a month. PDAF now provides algorithms ranging from domain-localized ensemble Kalman filters (e.g. LETKF, LESTKF) and smoothers to EnKFs with serial observation processing (e.g. EAKF), nonlinear filters, and 3-dimensional variational DA methods. The PDAF infrastructure available for these DA methods can be leveraged to add further DA methods and make them readily available to a wide user community. With the major release 3, PDAF was overall modernized and new functionality was added. A universal interface enables the use of all EnKFs and nonlinear filters without code changes. Further, incremental analysis updates can be applied without changes to the model code. For coupled data assimilation, PDAF supports the case where different model components are simulated on the same model grid, like to ocean physics and biogeochemistry, as well as for e.g. coupled atmosphere-ocean or Earth system models in which different components of the Earth System are simulated on distinct grids and perhaps separate model programs.