

Interactive comment on “Developing a common, flexible and efficient framework for weakly coupled ensemble data assimilation based on C-Coupler2.0” by Chao Sun et al.

Anonymous Referee #1

Received and published: 29 June 2020

The present paper reports the development of a framework for a coupled data assimilation (CDA). Works related to CDA require important technical preparation where many obstacles may occur. This might be related to numerous problems ranging from the compilation of a coupled model code along with a coupler code, communications between different models to purely IT-related problems like MPI libraries and so on. All these problems have to be addressed prior to scientific problems of choosing an optimal coupling scheme between models, and finally the CDA problems like accurate parametrization of error covariance matrices in a coupled framework etc. That is why the present work is a necessary first step towards a CDA system. And only technical aspects of the development are discussed here. However, I suggest that the authors

C1

work more on the paper text. The current text is difficult to understand. The authors present their system as a framework towards future CDA systems describing the system using words like a DA system, a numerical model etc. Conversely, the PDAF system is mentioned, which is somehow used along with the existing DA system that authors call GSI/EnKF. But the use of PDAF may be limited to models that can be incorporated inside the PDAF Fortran code and called as subroutines which is not the case for every model. That is why the authors might clarify early in the text what is the configuration of their future CDA system and what are the purposes of the system.

These are my specific comments.

First, please increase the dpi for the figures, it's hard to read them.

Line 21: please properly introduce the acronyms: "...system GSI/EnKF and the coupled model FIO-AOW".

Lines 26-31: Here you are talking about two widely used DA algorithms: ensemble techniques derived from the Kalman filter and variational data assimilation. But there are methods that you don't mention. Please reformulate the paragraph accordingly. Besides, I would suggest to remove the sentence ", can be viewed as a special case of ensemble-based methods with only one member in the ensemble when we attempt to design and develop a software framework for data assimilation." (lines 30-32).

Line 34: Please add the Environment and Climate Change Canada's (ECCC) hybrid 4D-EnVar DA used operationally in the weather prediction system: Buehner, M., McTaggart-Cowan, R., Beauvine, A., Charette, C., Garand, L., Heilliette, S., Lapalme, E., Laroche, S., Macpherson, S.R., Morneau, J. and Zadra, A. (2015) Implementation of deterministic weather forecasting systems based on ensemble-variational data assimilation at Environment Canada. Part I: the global system. Monthly Weather Review, 143, 2532–2559

Lines 36-37: With the rapid development of science and technology, numerical fore-

C2

casting systems with only an individual component model (such as an atmospheric model) have reached a predictability limit. Please reformulate it or add appropriate references.

Lines 37-38: Coupled models have been widely used in numerical forecasting to break the bottleneck of the limited predictability. . . I would also reformulate it in a way that it was established that coupled models may provide better results with respect to uncoupled models. Besides, you may wish to add references for coupled models that used operationally in several operational centres (Met Office, ECMWF, ECCO, etc.)

Lines 42-43: add references to coupled data assimilation (CDA) systems developed in JAMSTEC, NOAA, Met Office, ECMWF, ECCO: Sugiura et al., 2008, Mochizuki et al., 2016, Yang et al., 2013, Zhang et al., 2014, Lea et al., 2015, Laloyaux et al. 2016 and 2018, Browne et al., 2019, Skachko et al. 2019

Lines 45-49: It is not clear what do you mean. Reformulate or simply remove this sentence.

Line 49: You don't mention the NASA GMAO system, for example. I would suggest changing "most" by "several".

Lines 49-60: The ensemble size of current global operational systems used in NWP generally exceeds 256 members. The global models used in the ensemble are run on grids with horizontal sampling of several tens of kilometers. So, the I/O operations are currently necessary. If the aim of the paper is not to discuss real systems, please state it more clearly in the abstract and the beginning of introduction. It will help to understand what kind of CDA system you are developing and its future use.

Line 62-63: reformulate the meaning of WCDA. Not the models are assimilated independently, but data from model components are assimilated independently by two separate DA systems.

Line 76-78: It's obviously more efficient. However, you don't mention the feasibility of

C3

this approach when a model (or several models and couplers) is called in the PDAF Fortran code is a subroutine.

Line 332: again, GSI/EnKF has not been defined yet. What is GSI?

Lines 330-343: It's difficult to understand this section. First, you have an atmosphere-ocean-wave coupled model. Second, you have a DA system that you call GSI/EnKF. Does this system compute atmospheric analyses? How ocean analyses are computed? Do you compute also analysis for the waves?

Line 345: What do you mean by "combines"? Is it a hybrid system? Or you may choose between two options: a variational (which one?) and ensemble technique? You may also state explicitly that both options share parts of codes, for example, observation operators.

Lines 345-350: What DA method is used in your experiments: EnKF, EnSRF or ETKF? Please clarify.

Instead of using "pure ensemble DA", you may simply say that the EnKF (or ETKF or whatever) mode was chosen.

Line 354-371: atmospheric DA? So, do you use EnKF to perform atmospheric analyses?

L373: what step you are talking about? An atmospheric model time step?

L374-395: As shown. . . It's not clear.

L397: where it was used? And who has upgraded the coupled model to the c-coupler 2.0? After I read the section 5, I realized that there is no ocean data assimilation in your system. So I wonder why you call your prototype WCDA? Your system is an atmospheric DA using a coupled model to compute background states.

Interactive comment on Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2020-75>,

C4

2020.

C5