Review for "Towards Strongly-coupled Ensemble Data Assimilation with Additional Improvements from Machine Learning" by Kalnay et al.

This manuscript reviews the coupled data assimilation and strongly coupled data assimilation research conducted by Dr. Kalnay's group. The manuscript is well-written, with appropriate literature references. The content is very beneficial for the data assimilation community as well as the general readers who are interested in coupled data assimilation. I think it is well fit for NPG. In terms of the manuscript, I only have a few minor comments.

R: Thank you very much for your comments. All the line numbers in our response refers to the line numbers in the manuscript with tracked changes.

1. I can understand the abbr. SC and WC represent strongly-coupled and weakly-coupled, respectively. Please define them in the manuscript.

R: We have added definitions of uncoupled data assimilation (UCDA), weakly coupled data assimilation (WCDA), and strongly-coupled data assimilation (SCDA) in the first paragraph of Section 1, Introduction:

- L44: … the *uncoupled data assimilation (UCDA)* approach, which obtains independent analyses of different Earth system components based on the forecasts from uncoupled models, …
- L47: the *weakly coupled data assimilation (WCDA)* approach by creating separate analyses of the atmosphere and oceans, assimilating their domain observations based on the forecasts initialized from a coupled model.
- L52: the *strongly-coupled data assimilation (SCDA)* approach, which creates coupled analyses by assimilating the same set of the all-domain observations into different Earth system components,…

We also defined the quasi-SCDA in the second paragraph of Section 1, Introduction:

- L65: … implemented a *Quasi-SCDA* system through the "outer loop coupling", where the incremental 4D-Var atmospheric and 3D-FGAT oceanic analyses share the same outer loops so that their updated analyses will be used together to acquire the new model trajectory for the next round…

To ensure these definitions are clearly presented, we made those key words *italic* in their definition, and added the following sentence to L43:

"Different CDA strategies have been developed and summarized in Penny et al. [2017]"

And revised the section (in bold) related to quasi-SCDA from L65:

"The European Centre for Medium-Range Weather Forecasts (ECMWF) implemented the "outer loop coupling", where the incremental 4D-Var atmospheric and **3D-Var with the First Guess at the Appropriate Time (3D-FGAT, Lee et al., [2004]; Lawless [2010])** oceanic analyses share the same outer loops so that their updated analyses will be used together to acquire the new model trajectory for the next round [Laloyaux et al., 2016; 2018]. **Though cross-domain observations are not directly assimilated into separate earth components, separate earth component analyses benefit from a more coherent coupled-state through dynamical coupling at the data**

assimilation step. Based on Penny et al. [2017], outer loop coupling belongs to the Quasi-SCDA methods."

Reference:

Penny, S., Akella, S., Alves, O., Bishop, C., Buehner, M., Chevallier, M., Counillon, F., Draper, C., Frolov, S., and Fujii, Y.: Coupled Data Assimilation for Integrated Earth System Analysis and Prediction: Goals, Challenges and Recommendations. World Meteorological Organization, WWRP 2017-3, 50, URL https://www.wmo.int/ pages/prog/arep/wwrp/new/documents/Final WWRP 2017 3 27 July.pdf.

The same comment also applies to abbr. "NMC"

R: NMC method refers to "National Meteorological Center" method [Parrish and Derber, 1992].

We have revised L200: "Where $\mathbf{B}_{x,0}$ is the background error covariance of the initial ocean states estimated by the NMC methods"

to

"Where $\mathbf{B}_{x,0}$ is the background error covariance of the initial ocean states estimated by the **National Meteorological Center (NMC) method [Parrish and Derber, 1992].**"

We also added the following reference:

Parrish, D. F, and Derber J.C.: The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Weather Rev.,* 120:1747-1763.

Ln 11-12: the simple coupled Lorenz model —> a simple coupled Lorenz model

R: Corrected.

L11-12 "…, ranging from a simple coupled Lorenz model to …"

2. Ln 19: I am confused about the "full-rank" EnKF

R: The "full-rank" EnKF refers to the EnKF where we use an ensemble size greater than the size of analyzed variables. We use the full-rank EnKF in this experiment so that we don't need to apply any inflation or relaxation methods to maintain the ensemble spread. This simplifies our interpretation of experiment results.

3. Ln 24 55 upper oceans —> upper ocean

R: Corrected.

L24: "of the atmosphere and upper ocean"

4. Ln 125-126, The smallest RMSE shows at an assimilation interval of 8 time-steps that is your smallest assimilation interval. I think it is worth pointing out here.

R: Thanks for this comment. We added this point to the discussion about the EnKF results of the Lorenz model in L137:

“Singleton [2011] found that SC ETKF has the smallest analysis Root Mean Square Error (RMSE) when adopting an assimilation interval of 8 time-steps**, which is the smallest assimilation interval used in that study.”**

5. Fig. 2 From my understanding, ECCO only updates the boundary forcing and parameters, not ocean state variables. From the figure, the initial conditions of model states are updated by DA. Please clarify it.
R: In our study, the ECCO-like system updates both the initial conditions and the surface forcing. This is to mimic the ECCO approach documented in Stammer (2004), Page 4, paragraph 15:
“In the present calculation, the control vector includes the **three-dimensional initial condition potential temperature, θ, and salinity, S, fields, as well as the daily surface forcing fields of net heat, net freshwater, and momentum fluxes** over the full 10 years.”

6. Comparisons of 3/4D-Var and EnKF in a coupled QG Model
There are UC_clim, UC_3days and UC_1day. it is worth providing details on their adjustment. Which is equivalent to the regular UC applying in the atmosphere and ocean?
R: The intention of using a different forcing update for the UCDA is to investigate the benefits of using uncoupled and coupled models. As we can see in Figure (a)-(b), increasing the forcing update frequency for the uncoupled model reduces the analysis RMSE for both the atmosphere and ocean.

For the UC 3D-Var, we chose the slowest forcing update as 1 day to mimic the common approach adopted by operational centers where their uncoupled atmosphere model uses the daily SST products as the surface forcing.

The improved ocean does not enhance the RMSE in the atmosphere throng dynamic coupling, which needs some discussion.
R: Based on our results, we cannot draw the conclusion that “The improved ocean does not enhance the RMSE in the atmosphere through dynamic coupling”. Based on our calculation of the averaged analysis RMSE (scale factor 10^-5) over the last ~11 years, the SCDA atmosphere shows slightly better analysis than the WCDA atmosphere.

	WCDA	SCDA
Atmosphere	116.0	**115.9**
Ocean	5.516	**4.915**

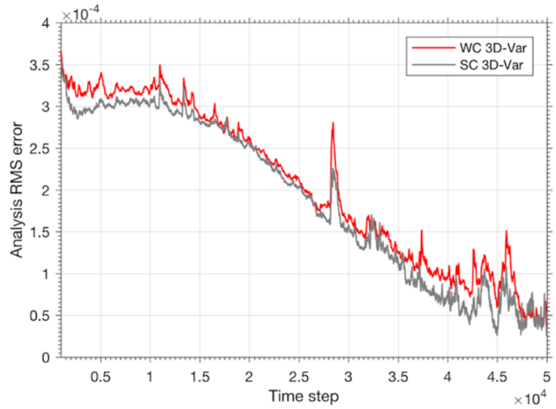
Ln 225-226 Fig. 4(a-b) only demonstrates two CDAs (WC, SC), not three methods,
R: In Figure 4(a-b), we showed WC (red), SC (gray) and UCDA with three different forcing update frequencies (climatology in blue, 3 days in green, and 1 day in yellow).

From fig. 4b, I can not conclude SC is better than WC.
R: Thanks for this comment. The conclusion that SC 3D-Var is better than WC 3D-Var is based on the averaged analysis RMSE (scale factor 10^-5) over the last ~11 years

	WCDA	SCDA

Atmosphere	116.0	**115.9**
Ocean	5.516	**4.915**

Besides, we also find that for the ocean analysis, the SC 3D-Var (gray) shows smaller RMSE than the WC 3D-Var (red) during the spin-up period, as shown below.



7. Ln 255 "lower than SC 3D-Var" should be "higher than SC 3D-Var"
R: Corrected.

L280: "Figure 5 (a)-(b) shows that when observing both the atmosphere and ocean, the SC 40-member ETKF and 4D-Var have similar accuracies for the atmosphere and ocean analyses, **higher than SC 3D-Var**."

8. Ln 257 "For 4D-Var, applying more outer loops and longer assimilation window lengths further reduces the analysis error". The state mentioned here has no support.
R: We conducted additional 4D-Var experiments with a 12-hour assimilation window and up to 4 outer-loops, and results show that they slightly reduce RMSE. Since those results are not new findings and have been shown in Kalnay et al. [2007] and Yang et al, [2012], we do not show the figures here.
We revise this sentence in L281 as:
"For 4D-Var, applying more outer loops (i.e., 3 and 4) and longer assimilation window lengths (i.e., 12 hours) further reduces the analysis error (figures not shown here), …"

9. "accuracies smaller than SC 3D-Var" should be "higher than"
R: Corrected.
L287: "For the atmosphere, ETKF, SC 4D-Var, and CERA present similar analysis accuracies **higher** than SC 3D-Var."

10. Fig.6 shows that the different RMSE between SC and WC has not reached to equivalent, especially the surface T/S, which needs to point out.

R: we add the following discussion to L325: “Longer model integration is needed to evaluate the performance of the SC and WC EnKF after the ocean surface temperature and salinity finishes spin-up.”

11. Ln 340 it is confusing for the statement “due to the missing vertical localization not used in the Ocean-LETKF.” Please rephrase.
R: We rephrase the sentence in L367 “The degradation below 25-m depth is probably due to the missing vertical localization not used in the Ocean-LETKF.”

as

“Since no vertical localization is applied in the ocean LETKF update, the degradation below 25m-depth is probably due to the sampling error caused by the small ensemble size.”

12. Ln 356 “states are assimilated by the SC EnKF”. The mean of SC EnKF here indicates the cross-model update, which is different from the other places indicating the whole DA algorithm.
R: Yes, we are performing cross-model update here, and the correlation-cutoff method [Yoshida and Kalnay, 2018] are designed specifically for the strongly-coupled EnKF. Note in our Figure 11, we also have the Standard SCDA (corresponding to the “Full” Pattern as shown in Figure 10), and it shows larger RMSE than the SCDA with the correlation cutoff method (corresponding to panel (c) and (f) in Figure 10).