J. D. Annan (Editor)

While an editor comment is not formally required at this stage, I would like to encourage the authors to prepare a revised manuscript that takes account of the suggestions of the two helpful reviews.

Anonymous Referee #1

We thank the referee for her/his comments, and give our responses in detail below.

General Comments: This paper presents methodology for the ECCO v4 reanalysis with specific attention paid to changes in the current version. It also tries to address the controllability of the system, which is a very interesting and important question.

Since ECCO is a well-established reanalysis product that the authors rightly point out has been used in many scientific studies, the refinement of the product is very relevant to GMD. The paper presents advances that are extremely suitable for addressing questions within the scope of EGU. The controllability analysis in this paper is novel and potentially useful in addressing scientific questions as well as improving future state estimation projects. As the authors documented, this iteration of ECCO has enough modifications and improvements from the previous version to be considered a substantial advance in modeling science. Furthermore, since ECCO v4 will undoubtedly be used in future studies, it is important for this documentation to be published.

The new changes to ECCO are well documented and changes are clearly outlined. The system is obviously very large and complex, so there is excellent direction to references if the readers want to learn more about different aspects. In this way, what is new and what is previously done is clearly stated and documented. I believe that the description is complete enough that others can use the freely available code that the authors describe and reproduce the author's results. Tools are also provided to apply the framework to other applications.

The paper is well written and largely free of errors. The title is appropriate and it is clearly stated which model version is being described. The abstract is sufficient in summarizing the contents of the paper. Furthermore, the appendices are thorough and appropriate for this type of paper. I have a few minor suggestions for clarifications and a few typos that I will list below, but I think this is a very good paper that is very appropriate for this journal and should be published with only technical corrections.
Thanks.

Specific Comments:
Page 3679, line 25 – It would be nice to have a little more elaboration on why it is tempting to attribute global mean sensitivity to discrete choices to omission of hydrology modeling over other explanations.
Sentence now refers to the "omission of atmospheric, continental, etc. modeling" more generally and as "merely a working hypothesis". The unclear reference to "hydrology" was expanded and is now stated between parentheses as an example to explain why we favor this hypothesis.

Page 3688, Line 28/Figure 10 – Fig. 10 shows misfits in ECCO v4 that are lower than other ECCO versions and other products. For the broad misfits that are still present, are there thoughts on how these can be addressed in future versions?
Further optimization (adjoint iterations) may simply be needed as stated in section 5.4, which now includes an explicit reference to Fig10.

Minor corrections:
Page 3681, Line 1 – commas around "for example"

Done.

Page 3686, Line 10 – Add the appendix number
Done.

Anonymous Referee #2

We thank the referee for her/his comments, and give our responses in detail below.

This contribution is appropriate to the journal, it reports significant new work and new concepts, and I recommend that it be accepted, although I ask for some minor revisions.

This contribution has two aspects.

The first is a noble effort to document a very complicated modeling and assimilation system so others can both understand and work with the output and even reproduce the runs. It provides a comprehensive (and mostly well-written) description of the model and state estimation framework, and it is hard to suggest much that has been left out or needs more explanation. I am not enough of an expert to challenge, or even deeply check, many of the detailed model equations, but I trust that they are correct.
Thanks.

The system is working, as the fits show, and the recent improvements to the system are very impressive. Their discussion of weaknesses is particularly interesting, and it is good they acknowledge that error analysis is a weakness. This is in contrast to the "competing" ROMS assimilation system, which has well-developed uncertainty tools, even though it is not automatically-generated. This sparks a comment that the only reference to ROMS is Shchepetkin and McWilliams for the modified Adams-Bashforth time step, but I guess that given the purpose of this contribution, it is not necessary to provide a survey of models.
The two efforts maybe complement each other rather than compete with each other. In particular the MITgcm adjoint is typically run continuously over multiple decades to carry continuous state estimation, while the ROMS adjoint is geared toward sequential data assimilation. Uncertainty quantification may indeed be further developed in data assimilation than in state estimation as suggested by the referee, but the two problems are quite different. Not unlike Moore et al 2011 for ROMS, the present paper focuses on presenting one system rather than comparing them. A reference to Moore et al 2011 was added in the concluding section.

The second aspect is a detailed analysis of many runs of the same assimilation system to look at the sensitivity to a wide variety of model changes: "structural" and "parametric". This is made possible by the impressive the modesty of the v4 computation requirements. I had somehow thought that this computation would be beyond reach, but it is not.

quoting from the text: "Advanced usage of ECCO v4 may include re-running forward model solutions (the state estimate in particular) or its adjoint. Computational requirements are modest –the 20 year forward model integration typically takes between 6 and 12 h on 96 processors."
As noted by the referee, this is an advantage of ECCO v4. This point is now stressed in the concluding section, with a reference to the Appendix containing this quote.

Although the analysis here is welcome and very well done, there are a few things left out. A main question I would suggest be addressed is whether the adjustments of the mixing parameters are needed to compensate for the resolution of the model, in which case their spatial patterns may have little to do with real mixing. I agree that they are a

very convenient way to allow model errors, and, combined with adjustment of fluxes at the ocean surface, presumably allow the elimination of biases in temperature or salinity. While these adjustments produce a large reduction in the misfit to hydrographic data, it would help the curious reader to have the author's discussion of the likelihood of compensating errors.
The curious reader was already referred to the separate paper that assesses the mixing parameters geography in section 4.4. This reference is now placed at the end of section 5.2 and extended to discuss error compensation. The need for additional investigation is also now further stressed in the concluding section.
Phrased another way, the inclusion of these controls moves the ECCO framework closer to the "weak constraint 4D-Var" used elsewhere, including ROMS, where error (control) terms can be included in every model equation at every grid point and time. This can make it easy to fit observations by deprecating the model dynamics, so a cautious approach is needed.
A major difference with weak constraint 4D-var however is that the chosen approach does not introduce source/sink terms of unknown nature in the model equations. Modifications of model parameters within acceptable ranges does not equate with deprecating model dynamics. The estimated parameters are of course uncertain, which is now further emphasized at the end of section 5.2.

In fact, the entire discussion of the model metrics responding to various controls and configurations could be made clearer. (and please write "Tables 3 and 8" instead of "Tables 3–8", which implies all the tables in between, too.) For example, is global mean temperature and salinity surface or total depth? If total depth, it's hard to see how mixing parameters could affect them much, since they redistribute the heat and the model is forced with fluxes. In addition, the conclusions drawn from the comparing the response magnitudes across different metrics were not convincing. I am especially skeptical of the bootstrap correlation calculations. So a bit more clarity here would help the dumber readers like me.
Section 5.1 was streamlined and clarified. Global means are computed over "total depth" but can react to any ocean model change since the surface buoyancy and freshwater forcing is computed through bulk formulae, which is now recalled in section 5.1. The use of bootstrapping is now further motivated.

Detailed suggestions:

first, please spell-check the manuscript. e.g. "abscence" on page 3662. "estime", and so on.
Done.

also there are many mismatched plurals, e.g. "slow models drifts" on page 3681. or page 3693: "internal-waves dynamics" (should be: internal-wave dynamics)
Most of these are trivial, but can slow down the reading.
Done.

Finally, the Conclusion section seems to be a bit redundant in places, repeating "expert choices" when it doesn't seem necessary.
One of the two instances of "expert choices" in the conclusion was removed.

Abstract:

I found the following phrase confusing: "Both components are publicly available and highly integrated with the MITgcm" I suggest: "Both components are publicly available. The ECCO framework is highly integrated with the MITgcm"
Sentence was revised accordingly.

I also found the phrase "model-data constraints" to be confusing. I suggest "observational and model constraints", or something more specific.
Sentence was revised accordingly.

In fact, the word "constraints" is often added to "observational" in the text, where "observations" alone might be better. Much of the time "observational constraints" is correct, referring to constraining the model to match observations, but sometimes it seems to just mean observations. There are many variations, e.g. "observed data constraints", and I suggest going through and regularizing the terminology so that it is always the same, and observations are just "observations".
`constraint' was removed when un-necessary (in section 4.3 and elsewhere). `data' was generally preferred to `observations'. Terminology was regularized throughout the paper.

for example on page 3684: "A permanent issue is the need for additional data constraints, particularly in the abyss (Wunsch and Heimbach, 2014)." This implies that the observations exist, but were not used as constraints. I don't think this is what is meant, however. I suggest: "A permanent issue is the need for additional observations, particularly in the abyss (Wunsch and Heimbach, 2014)."
Sentence was revised accordingly.

The description of the grid could be a little clearer: Page 3658: "Poleward of 57 N, LLC is topologically equivalent to CS minus one cube face (Fig. 1)." then, later: "Between the LL sector and the Arctic cap, the grid makes a gradual, conformal transition that is evident in Fig. 3 between 57 and 67 N."

I suggest these two sentences be put adjacent, and perhaps a little more detail on how the grid transitions from LL in a sector to something like the CS, if that is indeed what happens. And/or just move up the reference to Appendix A so it comes at the start instead of end, and the reader can look at that before puzzling about the short summary.
Paragraph was revised accordingly.

This also applies to the Southern Cap, where the text says: "To the South of 70 S, LLC is topologically equivalent to a pillow case that would have two vertices in each hemisphere (Fig. 2, right panel)."

I first thought "hemisphere" meant east and west, and was confused, but they mean north and south; this could be rephrased just to say there are two vertices for the conformal grid on the southern cap and they are on land. No need for the pillow case.
Paragraph was revised accordingly.

just a comment: The downside of the specialized grid is the complication of the analysis (page 3656), and I would be surprised if the Matlab toolbox could eliminate that entirely, but it is a very important part of the package.

page 3665: the "C-D" scheme is referenced without defining it. Please move up the definition, or refer to the defining section.
Done.

Figure 4: caption for lower panels should be clarified. Is it streamfunction at the surface or at 2000m?
Caption now specifies "for the vertically integrated flow".
The text should perhaps also say explicitly that the streamfunction is changed more by the C-D scheme than the customized viscosity, although the rms vertical velocity is also reduced more by the C-D scheme.
More explicit Fig. references were added where this point was already made.

I was surprised that the v4 solution was forced directly by wind stress instead of a bulk formula. I had thought that the ECCO framework had previous specified atmospheric fields and let bulk formulae compute the fluxes. Is this the case? If so, does the new approach not take account of stress modifications by surface currents or temperature

fronts? Please give a little more detail.
As already stated in the text, bulk formula are used to compute buoyancy and mass fluxes. For wind stress, both options are available as part of the model as now noted in section 3.5. The simple approach chosen in ECCO v4 accounts for wind stress errors directly, which is now more clearly stated. The referee's point regarding neglected stress modifications by surface currents was added.

section 4, page 3668: J is a SQUARED distance from the obs. same comment on next page, section 4.1
'squared' is now always appended before `distance'.

figure 6 caption needs some editing for grammar (e.g. "bottom panel show" (add s), "all biweekly period" (add s))
Done.

in fact, all the captions could do with a little proof-reading. It is tedious to list the typos.
term: "regression" tests: This is not regression as I usually see the word. I would characterize these as checksum experiments, or reproducibility checks...? Maybe this could be explained a little.
Captions were proof-read. We have further qualified `regression test' and added a textbook reference for this term in Appendix F. We prefer to retain the term "regression test", since it appears to be well understood to clearly distinguish between testing new features (not part of this) and testing that nothing has unexpectedly been changed (the goal of these tests). "Reproducibility tests" is not an ideal phrase, since the tests operate in the context of numerical and computational optimizations that may change reproducibility at the bit-wise level, but not change macroscopic behavior of a solution.

I particularly like Appendix G: the "solution history" section, as that is often a dark secret in state estimation. The model, parameters, and observations evolve during the iterations, and so reproducibility is difficult or impossible. I applaud the authors for listing the details, as part of their documentation of the solution. This disclosure could be helpful to others attempting related projects.
Thanks. Reproducing the detailed optimization and revision steps that led to this state estimate would admittedly be challenging. We should stress however that any ECCO v4 user can easily reproduce the state estimate solution. As far as we know this is a unique capability amongst ocean or atmospheric `re-analysis' products.

page 3702: "Revision 4 iteration 10 consisted in a filtering of atmospheric control parameters adjustments to reduce irregularities in the forcing that had appeared during adjoint iterations."

I suggest rephrasing or adding details to be clearer: was it space and/or time filtering?
`Filtering' was replaced with `trimming', which seems more accurate. This was done by subtracting leading EOFs, as now stated in the paper.