

Reply to Reviewers (Compo, Reviewer 2, Reviewer D)

Reviewer comments are small and indented, replies are normal text

We greatly appreciate the reviewers time and effort in reviewing this manuscript.

Review by Gilbert Compo

Review of Bosilovich et al. "On the Reprocessing and Reanalysis of Observations for Climate". The present paper gives an overview of the current status and challenges associated with obtaining atmospheric and oceanic observations of climate quality and with their reprocessing into climate data records, gridded climate data sets, and reanalyses. The paper is technically sound. The sections are well-written. I believe that the paper will be a signpost for important directions in these areas. With some relatively small modifications, I recommend that the paper be published in the OSC monograph.

There are a few substantive issues that can be easily overcome but will require some work.

1. Some sort of road map paragraph or subject paragraph that states the point or aim of this paper is needed. We are into the thicket of MSU issues and then on to data biases without knowing the overall scope of the paper. The road map should also point out that "oceanic reanalyses" are only peripherally being considered here, except as coupled to the atmosphere.

At this time, we didn't want to add an "introduction" section, as it may turn somewhat redundant with the subsequent sections. The abstract has been revised to call out the purpose of the paper, and also the limitation to atmospheric reanalyses.

2. The paper lacks a clear section with the requested "recommendations and priorities to the international climate research community." While some recommendations are present, it is not always clear what is recommendation and what is an issue that is being raised. In addition to their current placement, it would be beneficial to the reader to bring the specific recommendations together in a few paragraphs near the end of the paper.

The Manuscripts recommendations have been isolated and emphasized more in the revision

3. I was surprised that there was no clear call for a meteorological data set that is the equivalent of ICOADS but for land stations. The current text mentions that none exists. I suggest that a recommendation be made on this hole in the meteorological database.

Such a call is now included

Specific issues:

1. abstract: The connection between "reprocessed observations", "reanalyses", and "climate datasets" is unclear. If possible, establishing more of a connection between these three areas would help the reader.

Ultimately, both were organized together in the OSC session. The relationship is that both aim to improve long term data sets. Reprocessing creates pure observation record, while reanalysis uses multitudes of observations, data assimilation and a forecast model to create the record. We expect that the revised manuscript is more clear on this.

2. pg. 2 The sentence starting "While subsampled data may be less useful for time series analysis..." does not seem to fit with a paragraph on the construction of climate analyses.

This has been extensively rewritten and revised, with clarity a partial concern

3. pg. 2-3. I suggest that “analyses” and “climate data sets” be defined somewhere early in the paper. Particularly in its first instance, there is a potential confusion between climate studies and gridded fields. This confusion is seen in the 2nd paragraph where “climate analyses” and “time series analysis” are discussed. The use of “climate data sets” on page 3 also has the potential to cause confusion if these are not what was meant by “climate analyses” on the previous page.

The paper has been reviewed for consistent use of terminology

4. pg. 2. I suggest referring to the data sets as “observational” not “observed”. Alternatively, “data sets derived from observations”. This is stated clearly on pg 3 and “observational” is used on pg 10.

The paper has been reviewed and revised for a consistent description of the data sets

5. pg. 3. I suggest changing “all extant” for “most extant”. There is the possibility that the outlier dataset is wrong.

Agreed, the change is made.

6. pg 4. References to the “various groups” should probably be included in the XBT discussion.

The section has been revised considering this comment

7. pg. 5. start of 2nd full paragraph. I think you mean “Addressing the above concerns....”

The change is made.

8. pg 6. Reference or support for the sentence “Indices attempting....” is needed.

Citations have been included

9. pg 7. I’m not sure what “evolve their quality” means.

“improve” is the right sense, but not the right way to say, so, now written as improve the maturity of”

10. pg. 8 to 9. In an good discussion of ICOADS and ISTI, mention of at least two additional data sets would be appropriate. a. Mention of the WOCE dataset (<http://woce.nodc.noaa.gov/wdiu/>) on which the ocean reanalyses and coupled reanalyses rely, should be made.

b. Mention of the International Surface Pressure Databank would be appropriate.

ISPD v2 (<https://reanalyses.org/observations/international-surface-pressuredatabank>)

already includes feedback information from one reanalysis (20CR) that is the sort of information that is still in the planning stages for IVAD and ISTI. 11. pg 9. Mention of the Atmospheric Circulation Reconstructions over the Earth (ACRE, Allan et al. 2011) initiative would be appropriate. ACRE encompasses the data rescue and recovery mentioned earlier through partners such as oldweather.org, the utilization of the observations in reanalyses, such as the 20CR, and the application of the observations and reanalysis fields to climate needs, such as energy, agriculture, and ecology.

This has been addressed in the current revisions

12. pg 10. I suggest that the key sentence: “Reanalyses differ from reprocessed observational data sets in that sophisticated data assimilation techniques are used in combination with global forecast models to produce global estimates of continuous data fields based on multiple observational sources.” appear much earlier in the suggested road map of the paper.

Copied to first paragraph of the first section, with a brief definition of reprocessing

13. pg 10. Referencing the "State of the Climate Report 2010 (Blunden et al. 2011)" for climate monitoring would be appropriate at the end of 1st full paragraph.

Citation added.

14. pg 10. Reference or support is needed for the statement, "Estimates of the basic dynamic fields..."

Citation added.

15. pg 12. Perhaps the time span of each reanalysis could be mentioned, or reference table at [reanalyses.org](http://reanalyses.org/atmosphere/comparison-table)

Included.

16. pg. 18. I suggest changing "reanalysis.org" to "reanalyses.org". Both work, but the former redirects to the latter, which may confuse some.

Changed.

17. pg 18. You may want to mention that reanalyses.org includes web-based tools that allow users to interactively compare reanalysis fields and time series.

Included.

Minor points:

1. pg 4. Minor typo "chases" to "cases".d.

corrected

2. In many sections, there are missing commas joining two sentences or missing commas before an introductory clause. Examples:

a. pg 6 "In order to assess..." - comma is needed after "CDR".

b. pg. 6 "Validated uncertainty estimates... and a lack of data..." - a comma should appear before the "and".

The document has been proofed for this issue

3. pg 7 to pg 8 - redundant sections on "With such a large range of data products..."

Good catch, second shorter paragraph deleted.

4. pg 11. "conventional observing system" should be defined.

Conventional observations tends to mean in situ and retrieved observations, assimilated in a conventional method, as opposed to radiance assimilation. It was defined later in the manuscript, around page 17. Now elaborated here as well.

5. pg 13. Current write-up of 20CR makes is seem as though the dataset is only hemispheric. Despite large uncertainties, it is globally complete.

Well, the “hemispheric” is mainly a sideways reference to the increasing uncertainty in the tropics, and higher quality when dynamics and flow dependence have influence on the analyzed field. The sentence has been modified to show that 20CR is indeed global, and the word hemispheric removed.

References:

Allan, R., P. Brohan, [G.P. Compo](#), R. Stone, J. Luterbacher, and S. Brönnimann, 2011: [The International Atmospheric Circulation Reconstructions over the Earth \(ACRE\) initiative](#). *Bull.*

Amer. Met. Soc., 92, 1421-1425. doi: 10.1175/2011BAMS3218.1.

Blunden, J., D. S. Arndt, and M. O. Baringer, Eds., 2011: *State of the Climate in 2010*. *Bull.*

Amer. Meteor. Soc., 92, S1-S236. doi:10.1175/1520-0477-92.6.S1.

Respectfully,

Gilbert P. Compo

Thanks much!

Reviewer 2

I was asked to review the above manuscript, which is intended to be a community--- based scientific position paper addressing the above title. After reading the manuscript I have to conclude that the manuscript is not addressing the title, but is very much focused and limited to atmospheric reanalysis, required observations and shortcomings in the field. I find this very disappointing but can see this is partly related to the list of authors. My recommendation therefore is to change the title to reflect that the paper is limit to aspects of atmospheric reanalyses and atmospheric data (instead of climate data sets), or, alternatively, ad experts form other parts of the climate system and make this an all---inclusive paper. As it stands it would not suffice as community paper on aspects of climate observations and climate---oriented reanalysis at large.

You are absolutely correct in that the form of the paper reflected the community response to WCRPs call to review and contribute to the white papers, as well as the author’s efforts to attract further contributions. While the tint of the paper is certainly more toward atmospheric reanalyses, the fundamental methodology and problems with data density and quality are quite similar. Nonetheless, now that the 4th WCCRP international reanalysis conference has past, we have integrated some of the ocean, land and cryosphere discussion and needs into this manuscript.

few specific comments:

There are now figures. This is in principle OK. But adding a few could illustrate greatly the status of climate data and reanalysis and problems encountered in both.

The citations in the manuscripts have figures. The authors found no need to reproduce previously published figures, nor were any new figures generated to support the discussions in the manuscript. In the revisions, however, A figure on uncertainties was included. We would be interested in specific suggestions for figures that would be both original and help convey the key messages to the reader.

Abstract:

What do the authors mean by "Climate data sets are generally adequate for process studies and large--scale climate variability. Beyond that careful investigation of the data ... are required ..."?

It should be made clear what "reanalysis data" means. Here it implies that those are results from atmospheric reanalyses. However, this is not necessarily the case.

It is not clear, as commented here, what the difference between climate data, atmospheric reanalyses and more general reanalyses needs to be articulated. For example, historical reconstructions of SST exist for the 20th century, despite low data densities. Likewise, atmosphere and ocean reanalyses both exist for the 20th century. So its not clear that there is a discrepancy between what is meant by "reanalysis data" compared to "atmospheric reanalyses" or "ocean reanalyses", at least in the abstract.

Chapter 1:

Ad a reference to first sentence of 2. Paragraph.

The section has been extensively revised including these comments

The paper uses jargon in several places. Examples are "climate---ready analyses" on p. 2 or "one size fits all". Probably not everyone outside the atmospheric community knows what "background" means.

Jargon has been removed or clarified

One-size-fits-all would be better said as singular or sole approach. Simply, it means that developers may have to adapt for special circumstances.

3rd paragraph: explain what "potential problems" with observations mean. I am at a loss as to what the rest of the paragraph means.

This section has been revised extensively

p. 6: I find it interesting that in the context of CDR a reference to a paper from 1974 is being used. Is this still the most up---to---date paper?

Additional citations are now provided

p. 7: not sure of OCR is being defined. What is meant by "With >> such << a large range of data products currently available ..."?

Optical Character Recognition (OCR) is defined and the sentence has been clarified.

p. 8: The first paragraph is a straight repetition from p. 7.

The shorter paragraph has been removed.

Statement at end of page is not correct: outside of the atmospheric community such archives do exist!

This has been revised. However, it would have been constructive to provide supporting information and a list of comparable archives.

p. 11, bottom line: it is only here that the first time "atmospheric reanalysis" is being stated explicitly, although almost exclusively the atmosphere is being addressed.

See the previous response to this issue

p. 15: what are reanalysis observations?

This has been revised for clarification

p. 18: WOAP does not exist any more

During the development of the submitted draft, WOAP existed, or it was clear there would be a follow on committee. So, WOAP reference was maintained, at least as a place holder through the review process. This has been revised to reflect the current WCRP structure.

Thank you for your time and effort in reviewing the paper.

Reviewer D:

I am mostly for the description of the paper. Most of my comments are minor ones.

Please note that I am not English native. A few descriptions in this draft are not understandable for me.

Line 7-8 ... by GCOS (GOOS etc

This is wrong. GOOS is a different framework from GCOS. GOOS is Global Ocean Observing System. GCOS has 3 panels; AOPC, OOPC and TOPC, for atmosphere, ocean and terrestrial respectively. OOPC is supported by GOOS.

The GOOS acronym has been removed. The link is to principles developed by GCOS.

Page 3 Line 3-5 "For example, satellite typically measure radiances which can be related to sea-surface temperature only by...."

I cannot understand the meaning of this sentence. Why do you say "sea-surface temperature only"? Each channel of measuring radiance has its own weighting function. Some channels do not depend on SST.

The sentence is using sst as an example, the word "only" is not needed to make the point, so, hopefully the revised sentence is clear.

Line 19

What do you mean "one-size-fits-all" ?

One-size-fits-all would be better said as singular or sole approach. Simply, it means that developers may have to adapt for special circumstances. There have been substantial revisions to content and clarity

Page 4 Line 1-18 Sorry this paragraph is difficult. In particular I cannot understand well the description from line 7 to 14.

Revised

Page 6 Line 8 What do you mean "covariability of errors" ? I do not understand the meaning well.

Revised

Page 7 Line 6 "OCR" should be written "Optical Character Reader (OCR)"

This has been corrected.

Page 10 Line 9-11 "An important disadvantage"

I think the largest problem is model bias (background bias) given as first guess field for data assimilation. I think jumps appeared in time series mainly depend on the "existence or no existence" of observational data, although I agree the strength of the observational constraint.

Actually, I think we are in agreement, perhaps said differently. The appearance of the model bias changes based on the strength of the observational constraint, even just going from some observations to many more observations will change how the model bias appears. Ultimately, it is the model bias that causes the problems.

Page 12 Line 1 25 year ---> 25-year

Line 2-3 and their..... ---> and their second, the Japanese 55-year reanalysis (JRA-55) is underway

Line 13 remains --> remains ?

Line 18 MERRA, and CFSR --> MERRA, CFSR, and JRA-55

JRA-55 also uses the variational bias correction method for satellite radiance.

These have been amended in the present version.

Line 19 It is true that CFSR is the first reanalysis to use a coupled ocean/atmospheric model, but it uses intermittent coupling. It is not fully coupled system. Fully coupled analysis is not easy because time scales of atmosphere and ocean phenomena are quite different. Here 'intermittent' is an essential word.

Your comment on CFSR is accurate, and we have revised the sentence, however, using the Saha et al terminology "semi-coupled".

Page 15 Line 11-13

Please emphasize that "reanalysis" is not "observation". I know many papers confused "reanalysis" and "observation". Many researchers who do not know observations well do not realize the difference. Observation is not always truth. Reanalysis is not always represent the truth field. Nobody knows the truth.

It is true that this issue cannot be overstated. Section 2.3 has been revised to show that reanalyses take observations as input, but are not observations themselves.

Page 16 Line 3(Haimberger, 2007) is not listed in the References.

This oversight has been corrected.

Page 17 End of Chapter 2

If possible, please mention reprocessing past satellite data. For example, EUMETSAT and JMA completed to reprocess AMV's of METEOSAT and GMS each. AMV data is effective data to improve reanalysis quality because they cover globally. If we could use reprocessed AMV over the US region, it must be quite useful and would contribute greatly to produce a more homogeneous and higher quality reanalysis. There must be archives of raw original satellite data somewhere in the US organizations / institutes.