

Interactive comment on “Impacts of physical data assimilation on the Global Ocean Carbonate System” by L. Visinelli et al.

Anonymous Referee #2

Received and published: 25 April 2014

General comments

The paper “Impacts of physical data assimilation on the Global Ocean Carbonate System” discusses the impact temperature and salinity data assimilation has on a low resolution global physical/biogeochemical model. This topic is of clear relevance to Biogeosciences and will be of interest to the research community involved in ocean biogeophysical modelling, and to the ocean data assimilation community.

In general I am happy that the ideas presented in the work are of value and seem to be backed up, with some reservations, by the results given. The scientific methodology pursued in testing the authors' hypothesis (that physical assimilation improves the carbonate fields) seems to be reasonable, although it is in need of improvement. Thus I believe that this work should be published. However, the paper in its current form

C1264

does possess substantive weaknesses, outlined below, that will need to be addressed before formal publication. I thus recommend that this paper should be published after moderate revision with further review.

Specific comments

1. All of the other efforts of which I am aware in which physical data assimilation has taken place alongside a biogeophysical model have experienced severe problems with spurious upwelling. A particularly good example of this problem, which used a similar model and assimilation scheme to your experiments, is shown in ‘A nutrient increment method for reducing bias in global biogeochemical models; 2010; JGR; 115, C10036;’. This paper should be referenced in your work and shows upwelling of nutrients in the boundary currents which then diffuse into the gyres. I think it is likely that your runs have something similar occurring; indeed evidence can be found in your plots 8b and 8c where TSREAN has more DIC in the gyres than CTRL, particularly in the Pacific.

As this is one of the key problems when assimilating physical data with biology you need to investigate the issue within your runs and comment upon it within the paper.

2. Your description of the data assimilation system (section 2.3) is inadequate and lacking in important details. In particular, please address the following: a) The description of the EOF analysis is confusing. It is not clear whether the EOFs just describe correlations between different variables or include vertical correlations. Also you need to state how many EOFs you use and the percentage variance they explain. b) If they are not included in the EOFs, then please provide details of the vertical correlations. c) Please provide more details of the inflation factor applied to the observation errors. A quantitative statement about the size of the inflation needs to be given, as well as details of the spatial variability – a figure may be useful here. d) The statement “rejects observations with a too large departure” is vague and not scientifically rigorous. Please provide a quantitative description of the quality control criteria.

3. I am not convinced by the methodology used to compare model pCO₂ to the SOCAT

C1265

data. As your model and SOCAT climatologies are of substantially different periods (1993-2010 and 1968-2007 respectively) therefore simply differencing them is scientifically dubious. A far better approach would be to compare the model and data directly and calculate the mean and standard deviations of the differences.

4. In section 3 you describe calculating the RMS of the difference between observations and run TSREAN. Your description implies that you take the difference between the observations and the analysis; i.e. you compare against the observations after they have been assimilated. These observations are not independent of the assimilation and should not be used to assess skill. What you should give in the paper are the statistics of the observations minus background; i.e. the statistics of the innovations. These sorts of statistics are commonly used in assessing the skill of assimilation systems. While not as good as using genuinely independent data, using the innovation statistics is far more robust than using the observations minus analysis.

5. Your equations (4) imply: $DpCO_2 \sim \exp(\theta \Delta T)$ and $DpCO_2 \sim DS$ (here I have used D for Delta); thus I cannot see how you get the expressions in equations (5).

6. Your argument between lines 20 and 25 on page 5414 seems to rely on the idea that alkalinity is not affected by physical transport, particularly vertical transport. However, this is not true and alkalinity is advected and mixed just like all other properties of seawater.

Technical corrections

p5400, l15: "once physical" not "once the physical"

p5400, l16: "in pCO₂" not "in the pCO₂"

p5400, l22: The atmospheric concentration is 'measured' not 'estimated'.

p5400, l23: The "increase" is a greater than 30% growth; it is nonsensical to say it "resulted" in it.

C1266

p5400, l25: Change "has been mainly imputed to the changes..." to "has been mainly attributed to changes..."

p5401, l1: The value of 9.7 PgC/yr needs a reference, or at least make it clear which reference it comes from.

p5401, l2-3: Reword these lines. Try "carbon emission from combustion was 9.7 +/- 0.5 Pg C/yr [(reference)], with the fraction of this carbon going into the ocean estimated at 0.25 PgC/yr by Sarmiento et al (2010) and 2.9 +/- 0.5 PgC/yr by Le Quere et al (2013)".

p5401, l3-8: I don't understand the sentence beginning "Assessing the oceanic...", please reword. Splitting the information into shorter sentences will help.

p5401, l12: Reword to: "...flux that span inverse modelling techniques (refs), prognostic physical-biogeochemical OGCMs (refs), and global field reconstruction using sparse..."

p5401, l14: OGCM is acronym and should be defined when it is first used.

p5401, l18: It is not clear that you are talking about atmospheric pCO₂, rather than the oceanic equivalent. Please clarify the wording.

p5401, l18: "regulated by atmospheric" not "regulated by the atmospheric".

p5402, l1-4: This is a very long sentence and hard to read, try splitting it up.

p5402, l1-4: I'm not sure I agree that lack of constraints leads logically to a large number of models. I don't think you need to make this argument, just state that the biogeochemistry is uncertain – I'm sure that there are a lot of references that you could use to back up the assertion.

p5402, l7: "may bring" not "may already bring"

p5402, l14: There is no comma after recently.

p5402, l16: Remove "to help improving the reanalysis of the pCO₂".

p5402, l26: "collecting pCO₂ data" not "collecting pCO₂".

C1267

p5403, l4: "OGCM" not "OGCMs"

p5403, l5: Start a new sentence after "components".

p5403, l6: OceanVar has not been defined at this point in the paper, define it here or leave it out of the sentence.

p5403, l15-l16: It's personal preference, but I have rarely seen "Section" abbreviated to "Sect.". I suggest writing out the full word.

p5403, l19: You should also list the discussion/conclusions section.

p5403, l25: It might be worth putting in a figure of the ORCA2 grid, rather than describing it.

p5403, l25: The cos(theta) term should only appear if you are talking about the grid spacing in actual units of distance, such as kilometres, and not when you are using degrees.

p5404, l23 and elsewhere: You cannot refer to PELAGOS before you have defined what it is.

p5405, l3: "in marine systems" not "in the marine systems"

p5404, l18: Replace "...years, as we have tested..." with "...years, found by..."

p5406, l12: "is 10 days" not "is of 10 days"

p5407, l2: "checks" not "check"

p5408, l1: Salinity is also affected by advection and mixing.

p5408, l4: "of gaseous" not "of the gaseous"

p5411, l23: "...stations, both at the surface..." not "stations, and both at surface".

p5412, l3-10: You need to state how much you are filtering the data here, and not leave it until several paragraphs later.

C1268

p5412, l13: Define the acronym JGOFS.

p5412, l12-15: Why choose 0.2K for the mixed layer depth? You need to justify this better. You could just state that it is the same definition as used by De Boyer.

p5412, l21: "fact" not "facts"

p5412, l25: "partially attributed to" not "partially resulting from"

p5413, l4: Use "we remind the reader" not "we remind here".

p5413, l10: "in ALK" not "for ALK"

p5413, l14: "worsens" not "worsen"

p5413, l25: "a significant mismatch" not "significant mismatch"

p5413, l26: "poor performance" not "worst performance"

p5414, l3: "the data" not "data"

p5414, l6: "seen in Fig. 6" not "shown in Fig. 6"

p5414, l7: "we remind the reader that" not "we remind here that"

p5414, l12: "...values has a large impact..." not "values largely impacts".

p5414, l13: Use "especially" rather than "particularly"

p5415, l6: "largely covered" not "vastly covered"

p5416, l16: You need to define G here, not leave it until after equation 10.

p5417, l4-7: This sentence is too long, try breaking it up into shorter, easier to read, sentences.

p5417, l21: Gi is the area of the elements of the grid, it is not the grid itself.

p5418, l9: "show" not "showed"

C1269

p5420, l17: "...to that obtained in While..." not "to what obtained by While"

p5421, l5: "...being as the biology is very..." not "...being the biological very..."

p5421, l11: Start a new sentence after "variables".

p5421, l12: "such a coupled" not "such coupled"

Tables 2 & 4: Both of these tables actually contain two tables. However, this is not obvious from looking at them. Please make the distinction between the two tables clearer. There are many ways to do this: use a double line, use a thicker line, white space etc.

Table 5: In the caption use "(MEAN)" and "(STDEV)" not "(here MEAN)", "(here STDEV)".

Table 5: It would be far better to give a figure showing the different regions, than listing them in the caption text.

Interactive comment on Biogeosciences Discuss., 11, 5399, 2014.