Biogeosciences Discuss., 11, C7215–C7220, 2014
www.biogeosciences-discuss.net/11/C7215/2014/


Biogeosciences Discussions
Open Access



# *Interactive comment on* "Audit of the global carbon budget: estimate errors and their impact on uptake uncertainty" *by* A. P. Ballantyne et al.

**G. Peters (Referee)**

glen.peters@cicero.uio.no

Received and published: 9 December 2014

The paper gives a good overview of the Global Carbon Budget and details key aspects of its uncertainty. The paper has a synthesis feel, but at times could have a stronger integration with existing literature. The paper is well structured (Methods, Results, Discussion), but this has a consequence of repetition as each component of the carbon budget is discussed three times. In some sections I could not quite follow how some of the uncertainty analysis was done, and this might require some small changes to the descriptions. I do have a few concerns about some of the assumption on uncertainty distributions, but these may have good explanations that I missed.

Overall, I only have minor individual comments, but there are quite a few which may give the feel of a "major revision". In the list below, I make with a * some of the more

significant comments.

1. Abstract, line 3. In the first instance write out carbon (C)

2. The abstract has a feel that fossil fuel emission uncertainty has “come to dominate”, but this seems to contradict Figure 11? It seems LUC still dominates, but FF will dominate soon? 3. Page 14934, line 19. What about process emissions (other than cement)

4. Section 1.2 discusses atmospheric and ocean, and a paragraph for each. Wouldn’t it make sense to split to a section on atmospheric and a section on ocean?

5. A sentence which is mentioned a few times “Because fossil fuel emissions are often estimated from energy consumption or production statistics, they are a fairly well constrained economic variable”. I don’t understand this. Are FF an economic variable? What is a constrained economic variable? And why is something estimated from production statistics well constrained (is that a casual statement, is there a reference?). I think the energy statistics have quite some uncertainty, and may be less bound then differences in emission factors (or perhaps even energy contents in some cases/countries). I think this statement needs to be reconsidered (also in other places in the paper).

6. Section 2.2.1. Since this is talking about atmospheric concentrations, it would be useful to give numbers here in both ppm and PgC.

7. Section 2.2: “Because fossil fuel emission estimates are derived from economically-constrained energy consumption statistics, errors in these emission estimates are relatively small”. As before, how is this economically constrained and how big is small (5%, 10%, 20%)? This also seems to contradict other parts of the text saying that emission uncertainty now dominates.

8. Page 14941, line 10: Ok to reference Francey et al, but it may be worth also referencing the comment and response to that paper.



BGD

11, C7215–C7220, 2014

9. Section 2.2.1. The word “error” is used here a lot. Some of the uses are not really “error”? As an example (“accounting practices”) if one country uses a sector approach and another reference approach, is one of them in “error”, which this is just a different method to estimate emissions? If cement production is not included then I would only call it an “error” if they wanted to include it, but didn’t. Really, not including it is a system boundary question and hence a structural uncertainty?

10. Page 14942, line 1-2: “due to social and political pressures”. I don’t think Guan et al were that strong, but suggested it as a possible reason.

11. * Page 14942, line 6+: I am not sure I completely followed this. Countries are grouped to regions, and each region has a specific uncertainty. Ok (though, it would be good to give a table of the uncertainties for each region, helps for reproducibility). I didn’t understand the weighting bit. This is since you take random subset of countries from the region in the bootstrap, and then you need to rescale to replicate the regional total? What is the link to the errors of the largest emitters? I see you reference Andres et al, but I think adding an extra sentence of clarification may help [Incidentally, I have read Andres, and I searched for “Monte Carlo”, “bootstrap”, “weight”, and none of these words came up]. On the constant error “factors” are constant over time, is this the relative error?

12. * I am not an expert on bootstrap methods, but perhaps you need to give a few words on why you are using bootstrapping in this case (or paper). One way to generate samples would be to assume that you would have a relative error for each region (say 10%, specifying a standard deviation) and then assume a distribution (say log normal) and apply a random distribution to generate different samples. Are you doing this, and then resampling? I did not really see how you came up with a distribution.

13. Page 14943, line 1. Ok, I am perhaps a little slow. But what is El Camino? Google came up with some interesting results, so I guess this is not a standard term? Why did you use it?



BGD
11, C7215–C7220, 2014

14. Equation 4. I think it is great to include the temporal correlation. But why 0.95? Ok 20 years, but why 20 years as opposed to 10 or 30? I realise there is no data, but some explanation may help. The correlation will basically give a decaying correlation over time. The correlation with the adjacent year will be 0.95ˆ2, with an inventory 20 years ago 0.95ˆ20=0.35? Is that how I should interpret?

15. * Page 14943, line 9+. Ok to include CDIAC and EDGAR. But why BP. BP has crude estimates with no methodological description. The estimates can sometimes differ substantially at a national level. I would suggest it is better to use IEA, and better still, use IEA sectoral and IEA reference to make a subset of 4 emission estimates. Did you include cement with BP? If not, you will introduce a bias to the results.

16. * Equation 6. I will echo my point equation 4, but why 0.05. That is a tiny correlation. It is basically no correlated, and that correlation diminishes over time. Surely the correlation should be larger, even 0.95 as for FF. And how does 0.05 translate to 5 years? From the Global Carbon Project work a change in method can result in a complete change in the time series from 1959. I would expect the uncertainty in LUC to persist much longer than 5 years and certainly no less than the FF.

17. Page 14946, line 1+. The AF is introduced here, and mentioned a few times throughout. But, there seems to be no reference to the detailed analysis of AF in the literature. In the last 5 years so, several papers have been discussed on this topic, and I think it is worth linking to that literature here.

18. Equation 9. My first reaction was that this was a correlation matrix (use of Sigma), but this just represents combinations of different datasets? (3 FF and 3 LUC leads to 9 combinations?) For each cell in the matrix you have 500 samples (it is like a 3D matrix) and you have 52 years? I guess I am repeating what you are writing, but this suggests the explanation needs a slight tweak. . .

19. Page 14949, line 5+. “difficult to determine dC/dt was in fact increasing”. This is a little confusing, and I think a bit of care is needed. It is not that you have written anything



BGD
11, C7215–C7220, 2014





wrong, but you are talking about the rate of change of a rate of change (dC/dt). C is clearly increasing (dC/dt is positive), but it is unclear if dC/dt is increasing (d2C/dt2). In other words, it is unclear whether the growth in C is accelerating over time? I would just be a little more explicit on some of these distinctions.

20. Section 3.2. There is again the term “error” used here, and am not sure it is correct. Is “uncertainty” better?

21. Page 14951, line 24+. There is improved detection of changes in C update, but a recent change in that trend. Is this just a trade-off between the constantly reducing uncertainty in dC/dt but the growing uncertainty in E? This “trough” in the last decade may be more a coincidently combination of the uncertainties, rather than anything more physical in the climate system?

22. Page 14952, line 24. Ok, 122 simulations had a decreasing trend in N? That would mean that atmospheric growth (dC/dt) grew faster than emissions? This sounds unphysical, or I misinterpreted. It would be quite interesting to see a plot of the 122 sets of emissions and dC/dt to see if they look physical in any way!

23. * Discussion. It is ok to have a discussion, but I must admit I had a feeling of deja vu. I think I read some of this before! Perhaps one weakness of the paper is that it does not link to the existing literature. The Global Carbon Project also does quite some work on understanding the global carbon cycle, yet this work is barely mentioned (only mention is to the ocean data?). I think the discussion would be a good place to compare with the work of the GCP. What new is added with your analysis? E.g., “others have underestimated X”, “we find that there has been insufficient emphasis on Y”, etc. That would greatly improve the discussion

24. “The greatest source of error in fossil fuel emission estimates is derived from national energy consumption statistics that can be as high as 20% of total emissions for some nations”. But earlier this was not uncertain as it was economically constrained?



BGD
11, C7215–C7220, 2014





25. Figure 3. There is a missing something “All inventories also include cement production as”? What did you do for BP?

26. * Figure 4-6. The figures are generally nice, but these ones make it difficult to get an idea of the distribution. For example, in Figure 4 it looks like a value between 0 and 2.5 is equally likely. Is it possible to plot with shading to give some idea of the distributions? Where is the median? Where are the 1 sigma values, 2 sigma, etc. Alternatively, a set of histograms could be placed under Figures 4-6 to show the distributions.

27. * Figure 4-6 (4,6 in particular). This figures show large “spikes” every year. This I imagine is a lack of temporal correlation. If you put in a strong temporal correlation (0.95) then those spikes will disappear. This means that if I plotted an individual realisation in these figures, they would be rather random (the emissions in year t+1 will have no link to the emission in year t). This effect should be much smaller in the fossil emissions. I think it is worth exploring individual realisations a little to see if they make sense. Ultimately, I would consider increasing correlations in the LUC data (as mentioned earlier). One would also expect correlations in the ocean data. Each measurement or model run is not independent of the previous value, in which case I would expect some temporal structure in that data.

28. Figure 8B, why is it so skewed?

29. Figure 8C, D. It would be good to show the 0 value on these figures.

30. Figure 9. I like this, it would be good to have colours that contrast more than blue and green (though I see why you chose those colours).

31. Figure 11. Nice summary of the paper. These seems to contradict the finding in the abstract? LUC is still the largest source of uncertainty, but FF is growing very fast.

---

Interactive comment on Biogeosciences Discuss., 11, 14929, 2014.