Overall comments

This study emulates various carbon cycle and climate models by using impulse response functions (IRFs) and investigates how the model differences lead to differences in emission metrics such as the GWP, GTP, and iGTP. The relaxation time scales and associated weights of the IRFs estimated in this study synthesize the results of several intercomparison projects. The authors show a usefulness of IRFs to gain new insights into different models. This study is based on a substantial numerical work using datasets from several intercomparison projects.

However, I have several reservations with the manuscript as it stands.

1) I would first suggest that the authors compare their results with those of previous studies (e.g. (Joos et al., 2012; Reisinger et al., 2010; Wuebbles et al., 1995)). In the current manuscript the actual scientific contribution of this study to the literature is not very clear because the paper does not integrate previous studies in the discussion.

    - (Reisinger et al., 2010) is a major study that quantifies systematically the uncertainties in the GWP and GTP based on not only model differences but also historical constraints. This paper under review does not characterize the uncertainty by using historical observations. (Reisinger et al., 2010) is touched upon in the introduction but deserves more discussion.

    - In addition, the goal of this study looks similar to that of (Joos et al., 2012), which is not cited in this paper. (Joos et al., 2012) looks into the responses of various carbon cycle and climate models to CO_2 pulse emissions and discusses the influence of model differences on metric values. Because it appears that the same group is involved in this paper, the authors could provide in-depth comparisons between these two papers.

    - Furthermore, (Wuebbles et al., 1995) is also a relevant study that investigates the uncertainty in the GWP, which could be discussed in this paper.

2) My second comment is related to the nonlinearity. Applications of a linear IRF, which is used by the analysis here, are by construction valid within the linear range of the global carbon cycle (below about the CO_2 concentration of 550 ppm) (Hasselmann et al., 1993; Maier-Reimer and Hasselmann, 1987). Although the authors acknowledge the linear limit in the introduction, from my reading they actually fit linear IRFs on the C4MIP output in which the models are run till 2100 (reaching 700 to 1,000 ppm). I speculate this created a bias in the estimates of the C4MIP IRF parameters. To extend the applicability of a linear IRF beyond its linear range, one needs to consider the dynamic equilibrium for the ocean carbonate species under rising atmospheric CO_2 concentration, which affects the ocean CO_2 uptake (Hooss et al., 2001). A detailed biogeochemical underpinning is provided in (Tanaka et al., 2007). Or, a quicker fix in this case would be to fit the IRF on the C4MIP data only till 2050, which is when the atmospheric CO_2 concentration does not substantially exceed 550ppm.

3) The IRF based on the C4MIP dataset shows a nearly constant airborne fraction beyond just a few years after the emission (Figure 1). This strong short-term response contradicts with the behaviors of the C4MIP models and is also not consistent with the current understanding on the global carbon cycle (Archer et al., 2009). I think that the C4MIP IRF requires further investigation. This problem may be caused by how the C4MIP IRF has been calibrated (issue #2 above), but I

am not sure what the reason exactly is.

The paper attributes this peculiar behavior of the C4MIP IRF to the rising emissions (Page 953, Lines 2-8). But I think that the rising emissions do not explain the short term response of the C4MIP IRF – the rising emissions are more relevant to the uncertainty ranges of IRF parameters, as the conclusion of this paper states that "the gradual evolution of the CO_2 emission scenario in C4MIP makes it difficult to uniquely determine the CO_2 IRF". Because the IRF accounts for multiple time scales of the carbon cycle response, the emission pathway should not influence the estimates of the IRF parameters (as long as it is applied within its linear range). Note that, if one attempts to estimate a single time constant (equivalent to an IRF with just one decaying constant), the emission pathway would influence the apparent CO_2 time scale (Archer et al., 2009). Related debates are summarized in (Tanaka et al., 2012).

4) The current manuscript narrowly focuses on the IRF approach. I believe that adding some background discussions would broaden the perspective of the paper. Various types of models are used in computing metrics ((Tanaka et al., 2010); see Figures 1 and 2 for references therein). Why are the authors revisiting the linear IRF approach? What are the advantages of an IRF over a (more complex) simple carbon cycle and climate model to probe the uncertainties in metrics? Why is the linear IRF in spite of its limitation for applications? These questions do not have to be the ones to be discussed in the paper, but I think addressing this type of broad questions would benefit the paper.

5) The limitations of the metric results in terms of the type of uncertainties explored are discussed at the end of the paper, but those could be brought up upfront. It is not a problem that this study explores the uncertainties in metrics arising only from the model differences (i.e. without looking at those characterized by historical constraints). But the paper could state clearly at the beginning that this study does not fully explore the uncertainties in the CO_2 response because the focus of this study is the differences in the models used in various intercomparison projects. It could also be stated at the beginning that the analysis does not consider the uncertainties related to non-CO_2 components. It is not clear how these unaccounted uncertainties would play out and affect metric ranges. Furthermore, it may be worth pointing out the importance of the time horizon – as the metric results show implicitly, the choice of the time horizon in many cases influences more strongly the metric values than the choice of models.

I brought up several issues with the current manuscript above. However, the paper will potentially be an interesting contribution to the literature. As a final remark, I felt that there is a room for improvement in terms of the presentation of this paper. It is my overall impression that the paper (including the abstract) can be shortened by improving the wording, polishing the text, removing redundancies, and etc. Also note that, because of the issue #2, which might significantly affect the metric estimates, I did not review the part dealing with the results for metrics (Sections 4.2 to 4.4). I have detailed comments (see Supplementary pdf).

Detailed comments

Page 936, Lines 3-4: It is unclear what this sentence exactly means.

Page 937, Lines 3-5: I am not very convinced by the need to develop further the inter-comparison exercise dedicated to CO_2 and temperature IRFs at this stage. An IRF intercomparison project has just been completed (Joos et al., 2012).

Page 937, Line 7: Comparing the global climate impact does not necessarily require an emission metric. This sentence can be revised.

Page 937, Line 16: The idea of the MGTP is first proposed by (Gillett and Matthews, 2010). The iGTP is equivalent to the MGTP. The original paper also needs to be cited.

Page 937, Lines 16-18: This sentence can be revised to reflect the fact that the GWP is by far the most frequently used metric not only in research but also in climate policies such as the Kyoto Protocol.

Page 937, Line 19: A few sentences to introduce what an IRF is would be helpful for the readers, I believe.

Page 938, Lines 1-2: This is not correct. From the carbon cycle side, nonlinear IRF approaches have been put forward (Hooss et al., 2001; Joos et al., 1996).

Page 938, Line 9: The CO_2 fertilization has been introduced earlier (Page 387, Lines 5-7). But the temperature effect on soil respiration is not pointed out, which is a major component of climate-carbon cycle feedbacks.

Page 938, Line 11: (Reisinger et al., 2011; Tanaka et al., 2009) have also looked into this.

Page 938, Line 13: (Joos et al., 2012) has also shown this.

Page 938, Lines 18-19: I would rather think that the authors assume a single time scale to apply the IRF concept for the non-CO_2 components. Here a clarification is needed for the nonlinearities involving CH_4 in particular. The OH chemistry influencing the CH_4 adjustment time is taken into account as the indirect GWP (Section 6.12.3.1 of IPCC (2001)).

Page 938, Lines 22-23: An alternative is to replace a linear IRF with an energy balance model (compare (Bruckner et al., 2003; Tanaka et al., 2007)).

Page 938, Line 25: For clarification, the saturation effect for CH_4 and N_2O is also strong and considered in the respective parameterizations (Table 6.2 of IPCC (2001)).

Page 938, Line 27: Table 6.2 can be directly referenced.

Page 938, Line 29: The size dependency is also due to the nonlinearity in the carbon cycle.

Page 939, Lines 23-24: I think that an original purpose to compute an IRF is to avoid running the AOGCM, which is computationally expensive.

Page 940, Lines 11-16: Related the issue #5, I suggest that the discussion also touches on the limitation of this analysis and clarify what types of uncertainties are considered in the metric results (compared to (Reisinger et al., 2010)).

Page 941, Lines 7-8: Related to my earlier comment (Page 938, Lines 18-19), I do not think that one always imposes a single time constant on the models of all the non-CO_2 components (for example, see (Reisinger et al., 2010; Tanaka et al., 2009)).

Page 942, Line 8: See my earlier comment (Page 938, Line 27).

Page 943, Line 8: (Hooss et al., 2001) also finds n=2.

Page 948, Line 8: Why is "total climate sensitivity" rather than "equilibrium climate sensitivity"?

Page 948, Lines 10-12: Are the climate sensitivity estimates used as prior estimates or directly prescribed to the IRFs?

Page 949, Line 11: It is stated that an inverse estimation technique of (Tarantola, 2005) is used for the estimation of the IRF parameters, but I wonder how it is actually applied and what the specific assumptions are. For example, what are the prior for the IRF parameters? Is there any correlation assumed between the IRF parameters? How does the objective function look like? To use the inversion theory of (Tarantola, 2005) in the context of optimization, (Tanaka et al., 2007) extract and discuss relevant assumptions.

Page 950, Lines 12-13: Why is a delta-pulse experiment more difficult for climate models than carbon cycle models?

Page 953, Lines 27-29: How strong does this discrepancy in the long-term constants affect the metric values analyzed in this paper (with a time horizon of < 100 years)?

Page 954, Lines 6-8: Why is there a substantial difference between the climate sensitivity estimated in this study and the one in IPCC (2007)?

Page 962, Lines 13-14: What are the examples of metrics comparing two non-CO_2 species?

Page 962, Lines 16-17: I would think the other way. This study shows that the IRF parameters for a long time scale are less well constrained due to the limited length of model runs available. Thus, results with a time horizon of 500 years would not easily be obtained. Even if obtained, one would need to take it with caution.

References

Archer, D., Eby, M., Brovkin, V., Ridgwell, A., Cao, L., Mikolajewicz, U., Caldeira, K., Matsumoto, K., Munhoven, G., Montenegro, A., Tokos, K., 2009. Atmospheric lifetime of fossil fuel carbon dioxide. Annual Review of Earth and Planetary Sciences 37, 117-134.

Bruckner, T., Hooss, G., Füssel, H.-M., Hasselmann, K., 2003. Climate System Modeling in the Framework of the Tolerable Windows Approach: The ICLIPS Climate Model. Climatic Change 56, 119-137.

Gillett, N.P., Matthews, H.D., 2010. Accounting for carbon cycle feedbacks in a comparison of the global warming effects of greenhouse gases. Environmental Research Letters 5, 034011.

Hasselmann, K., Sausen, R., Maier-Reimer, E., Voss, R., 1993. On the cold start problem in transient simulations with coupled atmosphere-ocean models. Climate Dynamics 9, 53-61.

Hooss, G., Voss, R., Hasselmann, K., Maier-Reimer, E., Joos, F., 2001. A nonlinear impulse response model of the coupled carbon cycle-climate system (NICCS). Climate Dynamics 18, 189-202.

Joos, F., Bruno, M., Fink, R., Siegenthaler, U., Stocker, T.F., Le Quélé, C., Sarmiento, J.L., 1996. An efficient and accurate representation of complex oceanic and biospheric models of anthropogenic carbon uptake. Tellus B 48, 397-417.

Joos, F., Roth, R., Fuglestvedt, J.S., Peters, G.P., Enting, I.G., von Bloh, W., Brovkin, V., Burke, E.J., Eby, M., Edwards, N.R., Friedrich, T., Frölicher, T.L., Halloran, P.R., Holden, P.B., Jones, C., Kleinen, T., Mackenzie, F.T., Matsumoto, K., Meinshausen, M., Plattner, G.-K., Reisinger, A., Segschneider, J., Shaffer, G., Steinacher, M., Strassmann, K., Tanaka, K., Timmermann, A., Weaver, A.J., 2012. Carbon dioxide and climate impulse response functions for the computation of greenhouse gas metrics: a multi-model analysis. Atmospheric Chemistry and Physics Disccussion 12, 19799-19869.

Maier-Reimer, E., Hasselmann, K., 1987. Transport and storage of CO_2 in the ocean -- an inorganic ocean-circulation carbon cycle model. Climate Dynamics 2, 63-90.

Reisinger, A., Meinshausen, M., Manning, M., 2011. Future changes in global warming potentials under representative concentration pathways. Environmental Research Letters 6, 024020.

Reisinger, A., Meinshausen, M., Manning, M., Bodeker, G., 2010. Uncertainties of global warming metrics: CO_2 and CH_4. Geophys. Res. Lett. 37, L14707.

Tanaka, K., Berntsen, T., Fuglestvedt, J.S., Rypdal, K., 2012. Climate effects of emission standards: the case for gasoline and diesel cars. Environmental Science & Technology 46, 5205-5213.

Tanaka, K., Kriegler, E., Bruckner, T., Hooss, G., Knorr, W., Raddatz, T., 2007. Aggregated Carbon Cycle, Atmospheric Chemistry, and Climate Model (ACC2) – description of the forward and inverse modes, Reports on Earth System Science. Max Planck Institute for Meteorology, Hamburg, p. 188.

Tanaka, K., O'Neill, B.C., Rokityanskiy, D., Obersteiner, M., Tol, R., 2009. Evaluating Global Warming Potentials with historical temperature. Climatic Change 96, 443-466.

Tanaka, K., Peters, G.P., Fuglestvedt, J.S., 2010. Policy Update: Multicomponent climate policy: why do emission metrics matter? Carbon Management 1, 191-197.

Tarantola, A., 2005. Inverse Problem Theory and Methods for Model Parameter Estimation. SIAM.

Wuebbles, D.J., Jain, A.K., Patten, K.O., Grant, K.E., 1995. Sensitivity of direct global warming potentials to key uncertainties. Climatic Change 29, 265-297.