Comments on revised version of “Rapid and reliable assessment of methane impacts on climate” by I. Ocko et al.

Prologue

I have entered the review process only at its final stage, replacing an earlier referee. Hence, my original intention had been just to check, whether the responses of the authors to the referee comments from the main stage have been adequate, and whether the paper text has reached a stage of maturity. This intended limitation of my referee status notwithstanding, I perceive shortcomings of this paper to an extent that I am unable to bring myself to passing over in silence.

Recommendation

My recommendation is to return the manuscript to the authors once again, at least to do some more ‘polishing’ on the presentation. I also strongly support to make a title change along the line of anonymous referee1’s suggestion: “The difficulty of using small ensembles of simulations of an ESM with large interannual variability to validate simple climate models in cases of small forcings”

Reasons

I regard the main conclusions of the paper, 1) ‘Well trained simplified models like MAGICC are able to provide a representation of the global mean climate well enough to provide assessment studies’, and 2) ‘Complex climate models have limited ability to identify the response to small forcings in cases where the expected response and the simulated internal variability have similar orders of magnitude’, as basically correct. However, the authors have allowed repeatedly to let themselves get carried away by their enthusiasm, and have inflicted into their paper a number of exaggerations of alleged complex model disadvantages that ought to be toned down for the final manuscript. (Beyond this, some confusing formulations need to be rectified.)

I also feel that the ESM simulations have not been optimally setup for the fairest possible comparison with MAGICC. Besides the possibilities, which specified dynamics simulations (e.g., Lamarque et al., 2012; Kooperman et al., 2012) would have been offered for reducing ESM internal variability, the possibility to calculate radiative forcing (RF) rather than effective radiative forcing (ERF) by radiation double calling (e.g., Chung and Soden, 2015; Dietmüller et al., 2016) has also not been exploited. This implies that when discussing radiative forcing calculations, the authors’ comparison is often not so much between a noisy ESM and a noise-free simplified model, but rather between a noisy ERF and a noise-free RF (see extensive discussion in Forster et al., 2016). However, as the main referees have not been so strict, I won’t go nitpicking either, here. Yet, I request that the distinction between RF and ERF ought to be clear throughout the paper, and that the consequences of using ERFs from the ESM, but RFs from MAGICC are openly discussed.

Specific Remarks

p. 1, l. 13 (Abstract): I recommend to phrase more carefully as follows: 'Using basic knowledge from observations and complex Earth system models, reduced-complexity climate models offer an ideal compromise in that they provide quick reliable insights into climate responses, with only a limited computational infrastructure needed. They are particularly useful for simulating the response to forcings of small changes in different climate pollutants, due to the absence of mentionable internal variability.'

p. 2, l. 23: Please, cite Fuglestvedt et al. (2010) in addition.

p. 4 l. 6, 7: '… if … the response is comparable..' ; '… and (iii) whether the lack of internal variability …'

p. 6, l. 17: 'climate sensitivity of MAGICC'; confusing as it is variable and arbitrarily set. Either recall the mean value from p. 5, l.22, or state as the reason is that CM3 has a higher sensitivity than the majority of AOGCMs used to define the MAGICC climate sensitivity).

p. 8 , l. 16 (major issue): ERF is introduced here as an extra abbreviation but is not used later when the AM3 forcings (p. 10, l. 7) are presented which are confusingly designated with RF. Please, rectify throughout the paper.
l. 18: delete one bracket behind Myhre et al. and, please, cite Shine et al. (2003) in addition.

p. 8, l. 30 Please, notice a technical error with the equation setting.

p. 10, l. 7: (major issue): Following Forster et al. (2016), ERF can hardly be expected to take robust values when calculated on a one-year or 5-year basis. Thus, the MAGICC vs. AM3 comparison must remain on a plausibility level within this paper, which should be emphasized.

p. 10 l. 25: Change CO2 to CO_2 (several further examples afterwards).

p.10, l. 20: I guess it's AM3 rather than CM3?

p. 11, l. 1-3: Confusing, as you have given another number (0.97 W/m^2) for the methane forcing on p. 10, l. 25. Is this *exclusively* due to the different reference year? If yes, please state so clearly; if no, please give other potential origin(s) for the issue.

p. 11, l. 13 : In my understanding, here we are not dealing with 'responses', but with 'forcings' and 'adjustments' from concentration changes; the 'response' is coming only in the next subsection, so please adjust the phrasing.

p. 11, l. 20: Throughout this subsection, the issue of differing climate sensitivity between CM3 and the MAGICC mean is not raised, hence creating the impression that you would regard identical surface temperature responses between the models as an optimal evaluation result. This is obviously not the case, please reconsider.

p.11, l. 29 (major issue): "MAGICC has much higher correlation coefficients [with observed data], likely through the absence of internal variability". This I an odd sentence which urgently needs to be set in an adequate context. Of course, reality has no “internal variability” because there is but one realization of it! Hence, it is absolutely necessary that such an argument must not even hint at the fact that MAGICC provides better agreement with reality than CM3 (or any other complex model) does.

p. 12, l. 6: ‘correlation of the ensembles means’, recall that the MAGICC ensemble is for 19 different model representations, while CM3 has 3 independent realizations of the same model.

p. 12, l. 19 (major issue): Given that the temperature response lags the radiative forcing, I deem it not surprising that (spurious) negative radiative forcing, while leading to temperature decrease after 1895, may not correlate with negative temperature response during the same time period, but occurs only somewhat delayed.

p. 12, l. 27 until end of this section: As the authors (correctly) claim a far-reaching influence of internal variability on the exact evolution of the temperature response time series simulated by CM3, I see no much sense in looking for mechanistic reasons to explain details of the actual evolution. If you like to keep this, please motivate it more convincingly. Another example of negative temperature responses simulated in case of CO_2 increase has been given by Huszar et al. (2013, their Fig. 10). There, too, mechanistic explanation attempts would have been obsolete.

p. 13, l 21 (major issue): “sophisticated coupled chemistry-climate models … are generally unsuitable for analysis of methane mitigation strategies”; expressed with such dogmatic universality, this statement has to be rejected. It only holds, if assessments of many mitigation options – especially with relatively little expected impact -, or a large extent of parameter sensitivities are to be investigated. Then application of complex models becomes clearly unfeasible and resorting to simplified models is doubtlessly required; best may be a combination of both as, e.g., described by Dahlmann et al. (2016) for an example of aviation impact mitigation.

References:

Chung, E.-S., Soden, B.J., 2015: An assessment of methods for computing radiative forcing in climate models, Environ. Res. Lett, 10., 074004.

Dahlmann, K., et al., 2016: Can we reliably assess climate mitigation options for air traffic scenarios despite large uncertainties in atmospheric processes? Transport. Res. Part D, 46, 40-55.

Dietmüller, S., et al., 2016: A new radiation infrastructure for the Modular Earth Submodel System (MESSy, based on version 2.51), Geosci. Model Dev., 9, 2209-2222.

Forster, P.M., et al., 2016: Recommendations for diagnosing effective radiative forcing from climate models for CMIP6, J. Geophys. Res. 121, 12460-12475.

Fuglestvedt, J., et al., 2010: Transport impacts on atmosphere and climate: Metrics, Atmos. Environ., 33, 4648-4677.

Kooperman, G.J., et al., 2012: Constraining the influence of natural variability to improve estimates of global aerosol indirect effects in a nudged version of the Community Atmosphere Model, 5, J. Geophys. Res. 117, D23204.

Lamarque, J.-F., et al. 2012: CAM-chem: description and evaluation of interactive atmospheric chemistry in the Community Earth System Model, Geosci. Model Dev., 5, 369-411.

Huszar, et al., 2013: Modeling the present and future impact of aviation on climate: an AOGCM approach with online coupled chemistry, Atmos. Chem. Phys. 13, 10027-10048.

Shine, K.P., et al., 2003: An alternative to radiative forcing for estimating the relative importance of climate change mechanisms, Geophys. Res. Lett., 30, 2047.