



Interactive comment on “The ACCENT-protocol: a framework for benchmarking and model evaluation” by V. Grewe et al.

V. Grewe et al.

volker.grewe@dlr.de

Received and published: 18 April 2012

We are grateful for the suggestions and have included them all.

Reviewer General Comments: This paper summarises a comprehensive framework for undertaking model evaluation. The use of example questions with which to test the models is very useful as is the discussion of uncertainty. It would be nice to see the results of this framework put into practice as a proof of context. Perhaps that is to follow in a future paper.

Answer: Thanks, and yes there are 4 protocols in preparation within the ESMVal project. However, as long as they are not submitted I refrain from explicitly refer to them in the paper.

C1786

Reviewer: Specific Comments: Table 1. I don't think it is necessary to include this as a table as these questions and their abbreviations are clearly stated in section 2.1.

Answer: Agreed. We had this discussion already among the authors and now follow the reviewer's recommendation to delete table 1.

Reviewer: Section 2.1: It might be worth stating here that asking whether a model is 'good' in the general sense is not as useful or perhaps even sensible as there are too many variables to assess, with likely too much uncertainty to constrain any results. Having a specific specific aspect (e.g. ozone layer recovery) provides much greater power for the validation. You quite clearly use very specific questions as examples of 'purpose' but its probably worth saying that a more general question is not a useful 'purpose' if that is the case.

Answer: Included in Sec. 2.1 (second last sentence).

Reviewer: Section 2.5: Can you include some brief discussion on a useful format in which to provide quality information? Is it a text description or a machine readable set of fields on model grid resolution or something entirely different? It is vital that this information is easily accessible and understandable to the scientific community and perhaps even those end users of products for decision making. Many different institutions could produce a variety of different formats of assessment. This could lead to confusion and in the worst case scenario it may lead to users ignoring such assessments when using the data. This is perhaps not something you can go into detail with here but a little discussion of the issues would be useful, especially if there is a preferred format that you are aware of. Measurement uncertainty in in situ observations of climate variables such as temperature, precipitation, pressure etc. is not well documented. There are many different instruments with differing levels of precision. Calibration of these has not always been regularly conducted and any adjustments not always documented. Any metadata that do exist are rarely in a digital format. I hope the situation for atmospheric composition data is much better and as this protocol is

C1787

aimed at that data perhaps the original statement is ok. However, its worth being aware that the picture is not so rosy for all climate data.

Answer: I think there is the need for two kinds of data. First Meta data describing as a text the basis of the uncertainty estimate, which can range from 'Expert judgement' to a short description of the algorithm and further reference. Second the uncertainty in digital form in the same way as the measurement data itself.

Reviewer: Page 3607 line 22 - Do you really mean 'smaller' here? For the models to be significantly different from reality would they not have to lie outside the uncertainty range given by the observations? Perhaps rewording this sentence may make its meaning clearer - I may have misunderstood.

Answer: Thanks, Re-phrased

Reviewer: Section 2.6: I'm not quite clear on the equations here but that might be my lack of familiarity with the specifics. Is 't' time?

Answer: Yes it is. I added it to in the text now.

Reviewer: Section 2.8: Given the subjectivity of benchmarking can you add a comment here stressing the importance of clarity in how the assessment was conducted. Anyone using the quality information or grading should understand exactly what was assessed and how. Also, if different model products are to be compared using this assessment then it is important that the same assessment was conducted on all products otherwise a comparison of the validation/assessment is not really useful. If assessments were to differ then this should be clear.

Answer: A sentence is added in section 2.8: "Hence, given the subjectivity of benchmarking, it is important that any assessment includes a clear description on how the assessment was conducted."

Reviewer Technical Comments: (suggested changes are capitalised) Section 2.2 I don't think you need 'As discussed in the last section' as this follows very nicely anyway.

C1788

Perhaps just saying 'The IDENTIFIED purpose...' would suffice.

Answer ok.

Reviewer Page 3603: line 11 - 'extend' should be 'extent' line 20 - 'input parameters AND boundary conditions such as emissions...' sounds more logical?

Answer: ok

Reviewer: Page 3607: line 3 - 'has impacts' I think instead of 'have an impact'

Answer: ok

Reviewer: lines 5-8 - I don't understand this sentence. Please can you reword it.

Answer: done! No idea what happened with this sentence. - thanks

Reviewer: Colons needed before lists: p3602 l26, p3605 l12, p3608 l5

Answer: hm. Anyway, changed accordingly.

Reviewer: Page 3609: line 12-13 - 'leads to the conclusion OF which parameters' sounds better. line 21 - 'feedbacks AND altitude of aircraft emissions...'

Answer: done.

Reviewer Page 3611: line 1 - 'ozone grade and a temperature grade ARE not comparable' lines 14-18 - 'an example which' no comma needed lines 14-18 - 'for the quality indicators AND THEN using its lower...' Page 3612: line 15 - no 'Especially' needed just 'The uncertainties.' line 27 - Perhaps use 'BENCHMARKING data' here as opposed to 'observational data' for consistency.

Answer: done.