



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

**Anonymous Referee #2**

Received and published: 16 April 2012

This is an unusual paper in that it does not present any new science, or new tools, but rather it formalizes a generic protocol, defining good practice, by which to evaluate and grade models. The basic ideas expressed in the paper are very general in nature. The authors define an 8-step procedure which a model evaluation should follow. The procedure could be applied to any kind of complex scientific problem amenable to modelling. The protocol is well thought out; an evaluation project following the protocol would satisfy a lot of formal requirements that could be demanded of such an evaluation. On the other hand, I do not think the community had been waiting for this paper to tell it how to conduct a model evaluation. The wealth of literature on this subject is proof of the contrary. However, considering a recent community move towards grading and ranking of climate models, rules around defining such grades, and clarity about what they mean, and how they can be meaningless, are now needed with increased

C1773

urgency. In this context, I think the paper is a valuable addition to the literature, and recommend it for publication in GMD with the technical changes listed below.

Details:

P3601L7: Remove first colon.

P3601L13: Remove “been”

P3602L11: Replace “although,” with “however”.

P3603L20: Cut out “such as”. Technically, I don’t think emissions are “boundary conditions”.

P3604L1: “hints at”

P3604L21: Replace “i.e.” with “namely”.

P3604L25f: This sentence is a bit awkward. Please rephrase, such as “It is important to indicate (a) why there is a need for these observational data, and (b) in case high-fidelity model data is used, its range.”

P3605L18: “Estimates of uncertainty arising from . . . are more difficult to derive.”

P3606L4: “that the models underestimate wintertime European NO<sub>2</sub> by a factor of 2.”

P3606L5: Replace “whereas” by “However”.

P3606L6 “statistically” (adverb)

P3606L16” Insert question mark

P3606L21: Are these really “frequency distributions” or are they PDFs?

P3607L2ff: Awkward sentence, please rephrase.

P3607L5ff: Some duplication here, please correct.

P3609L13: Please rephrase, such as “allows one to infer which parameters are simu-

C1774

lated with deficiencies.”

P3609L23: I don't understand this sentence. Precisely which ratio is formed?

P3610L7: “corresponding quality indicators” (plural)

P3611L1: replace “is” with “are”

P3611L3: “of including a statistical test among the quality indicators”

---

Interactive comment on Geosci. Model Dev. Discuss., 4, 3599, 2011.

C1775