

Interactive comment on “Three-dimensional variations of atmospheric CO₂: aircraft measurements and multi-transport model simulations” by Y. Niwa et al.

Anonymous Referee #2

Received and published: 17 August 2011

The investigators use 4 atmospheric transport models (3 online and 1 offline) driven by 2 different datasets of surface CO₂ flux (Flux 1 and Flux 2). The outputs from different numerical experiments are then compared to aircraft measurements taken at mid to upper troposphere as part of the CONTRAIL project. I find the paper informative and contributes to the further understanding of the behavior of various atmospheric transport models we use to interpret atmospheric CO₂ in terms of carbon sources and sinks. I recommend acceptance with major modifications.

Overall, the study presents some interesting model-to-model intercomparison results. The differences amongst these models in simulating radon (Figure 3) and CO₂ (Fig-

C7911

ure 4) are attributed by the authors to differences in the vertical mixing and cumulus convective parameterization schemes employed in various models used in this study. In this context, I have some questions/comments based on what is stated in the paper (1) For Figure 3, the authors state that the “low radon concentration suggests that vertical transport of MJ98-CDTM is slower than those of the other models.” Now, based on what is presented in Section 2.2.2 (MH98-CDTM0 and in Section 2.2.4 (NIES), it is my impression that the convective schemes employed in these two models are similar, yet the difference (particularly at 300 hPa) in Figure 3 between the two models is quite noticeable. Why is that? I would also like to suggest switching the 300 and 500 hPa columns of diagrams, putting the 500 hPa column to the left. (2) In the sentence following the one quoted above, the authors state that “ the simulated radon concentrations are rather comparable with each other” I strongly suggest deleting the word “rather.” (3) In Figure 4, why does NIES look quite different from MJ98-CDTM for JFM? Isn't the convective parameterization scheme used in each of the models very similar? (4) Also in Figure 4, unless I missed it in the main text, which flux (Flux 1 or Flux 2) was used to generate the diagrams in the figure? This information should at least be mentioned in the caption.

In regards to the comparison between the model output and the observation, I make the following questions/comments/suggestions: (1) In Section 3.2, the authors calculate average correlation coefficients in their attempt to establish the “reasonableness” of the model reproduction when compared to the observation (e.g., 1st paragraph, 2nd sentence: “. . . . models reasonably reproduced the observed vertical profiles: average correlation coefficients are 0.62 and 0.71, respectively” for Flux 1 and Flux 2.) How statistically significant are these numbers, and from each other? Every time one performs a statistical calculation comparing one variable with another, it is essential to establish statistical significance level. Here is another example from the paper (and there are others) in Section 3.3 (1st paragraph, 3rd sentence) where the authors state that “seasonal amplitudes simulated from Flux 2 are large and closer to the observed one than those from Flux 1.” Is the difference between Flux 1 and Flux 2 statistically

C7912

significant at, say 95% confidence level? (2) In Figure 5, there are large differences, depending on the geographical location, between the model output and the observed vertical profiles. When the authors state that (Section 3.2, 1st paragraph, 2nd sentence) “the models reasonably reproduced the observed vertical profiles” I think it might be helpful to put this in the context of between-model differences. (3) In the caption for Figure 5, there is a mention of two panels showing time-altitude cross-sections, one for 2006 and another one for 2007. This needs to be explained and discussed in the main text. (4) Section 3.2.1 (3rd paragraph, 4th sentence): I have dealt with many atmospheric models and I find deficiencies in these models troublesome. When the authors make a statement like “we cannot completely attribute the model-observation discrepancy to the model deficiency” how much can one actually attribute the discrepancy to incomplete or defective model representation of the real atmospheric dynamics? The authors need to provide sufficient proof. (5) Section 3.2.3 (1st paragraph, last sentence): Please delete the phrase “more or less” and use a more acceptable word like “generally” if one has to. (6) Section 3.2.3 (2nd paragraph, 3rd sentence): How do you know that the “model-observation mismatches” are not due to a shortcoming in the atmospheric model dynamics? (7) Section 3.2.3 (4th paragraph, 1st sentence): The phrase “marginally failed” needs to be quantified. Or just simply take out the word “marginally.”

Interactive comment on Atmos. Chem. Phys. Discuss., 11, 12805, 2011.