Atmos. Chem. Phys. Discuss., 12, C10123–C10126, 2012
www.atmos-chem-phys-discuss.net/12/C10123/2012/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Comparison of three vertically resolved ozone data bases: climatology, trends and radiative forcings" *by* B. Hassler et al.

T. Shepherd

tgs@atmosp.physics.utoronto.ca

Received and published: 4 December 2012

It is certainly useful to document the differences between these three ozone data sets and the reasons for those differences, and the authors are to be commended for the thorough fashion in which they have done this.

The problem I have with this paper, from a user perspective, is that it is a mixture of apples-to-apples and apples-to-oranges comparisons, and it is important for the reader to know which is which. Yet they are treated in essentially the same fashion by the authors.

The apples-to-apples comparisons involve issues such as using different ozone observations, different vertical resolution, etc. There is only one atmosphere and un-

certainties in reconstructing the past behaviour of the atmosphere are uncertainties of an apples-to-apples nature. The well-known issues with RW07 in representing the Antarctic ozone hole certainly fall into this category.

However, the use of different basis functions is an apples-to-oranges comparison, because the explained signal and the trend mean different things in each case. For example, it is said several times that BDBP fits the observations better in an RMS sense than the other data sets; well of course it does, because it includes more of the variability!

Most seriously, the use of different basis functions leads to different conclusions with respect to the 1979-1996 trends. In describing the ozone changes, the authors use "decrease", "depletion", "loss", and "trend" synonymously. It is very confusing to use multiple words for the same thing, and there seems to be an implicit assumption that all these quantities are deterministic responses to forcings rather than stochastic noise. The authors need to be crystal clear on what they mean by these terms, and if they think they represent deterministic effects that should be reproducible by models. Especially because the paper is ostensibly being written for modelers, it really needs an upfront discussion of the philosophy behind the different choices of basis functions, and how these might affect the use of the different data sets.

In particular, by including ENSO and volcanic basis functions, BDBP obtain a much larger Arctic ozone trend over 1979-1996 than the other two data sets. What is the user to make of this? If the effect was due to ENSO, which is known to affect Arctic ozone, then one might wish to include that effect in a model run with prescribed SSTs, but one certainly would not want to include it in a coupled model run or a run with prescribed modeled SSTs. If, on the other hand, the effect was related to the volcano term, then this raises the issue of whether the observed very low Arctic ozone values in the mid-1990s were really a deterministic response to the volcanic forcing. The authors seem to imply this (p. 26572, lines 8-11), but none of the three studies cited addressed polar ozone changes. It is well established (e.g. the last few Ozone Assessments) that the low Arctic ozone values in the mid-1990s were due to a combination of a series of very cold winters combined with high chlorine loading (Tegtmeier et al. 2008 GRL is probably the most comprehensive attribution study), so the argument would need to be made that the series of very cold winters was the result of the volcanic forcing. I am not aware of any solid literature basis for such a claim. At the very least, the reader needs to be told that this is the implicit assumption behind the claim that the BDBP Arctic ozone trend is really a trend and not just some natural variability that happened to alias onto the volcanic forcing term in the MLR model.

In a similar fashion, the authors introduce the use of a "hockey-stick" fit as an alternative to EESC (p. 26579, lines 4-5), arguing that perhaps the dependence on EESC is nonlinear and therefore not properly captured by an EESC term in MLR, even though there has never been any solid evidence that this is the case. But then it seems a bit disingenuous to only examine the pre-1996 and not the post-1996 part of the fit; one cannot ignore "inconvenient truths"!. Vyushin et al. (2007 JGR) compared EESC and hockey-stick fits to extra-polar total column ozone. There was hardly any difference for the SH, but for the NH, there were big differences between the two, similar to what is seen here (i.e. with the hockey-stick fit picking up much more of the pre-1996 decline). One then needs to think very carefully about the implications of the differences. Vyushin et al. did not make a judgement as to which was better, but noted that if the hockey-stick fit was to be believed, then ozone recovery (in the sense of a statistically significant increase) was on the verge of being detected in 2005, which seemed physically unlikely. None of this sort of discussion occurs here, which leaves the user helpless in deciding which data set to use.

What I would have found interesting would be to compare the component of the trend that was congruent with EESC in each case, as this would be an apples-to-apples comparison. This would be very useful for the modeling community.

p. 26564, lines 17-19: I find this statement odd. MLR does indeed specify the time evolution of different components of the signal, but it does not specify their amplitude, so the overall trend is not imposed, as you suggest. Note that this statement contradicts

the recent past practice of the WMO Ozone Assessment, which expresses trends in terms of EESC, so you are basically saying, without any evidence to back it up, that what they have done is wrong ("should not be used" is very strong language). In fact, aren't you also saying that the detection-attribution approach used heavily in IPCC is also wrong? Frankly I don't see anything wrong with asking what component of past behavior is attributable to a particular forcing. The key in any statistical analysis is to also look at the behavior of the residuals, which confirms the goodness of the fit. If, for example, the actual dependence on EESC is nonlinear, then this should become clear by looking at the residuals.

Interactive comment on Atmos. Chem. Phys. Discuss., 12, 26561, 2012.