**Atmospheric
Chemistry
and Physics
Discussions**

Interactive
Comment

Interactive comment on "Validation of remotely sensed profiles of atmospheric state variables: strategies and terminology" *by* T. von Clarmann

T. von Clarmann

Received and published: 27 July 2006

I feel embarrassed about the errors which have found their way into the ACPD version of the paper and gratefully acknowledge the reviewer's contribution to remove them. In the following, the issues raised by the reviewer will be treated point by point.

General comment:

As recommended by the reviewer, the application of the chi-square test will be discussed already in Section 3.1

Specific comments:

p. 4974, l.10: Indeed the usual terminology in testing is that the null hypothesis is the hypothesis of zero difference (hence the name!) and a significant difference is assumed if the null hypothesis is rejected. In validation the problem is posed in the

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

EGU

opposite way: The zero difference shall be verified, which leads to some confusion in terminology. For this reason the term "null hypothesis" will be avoided in the revised paper.

p. 4976, l. 17: Agreed, but I prefer to subtract the variance of natural variability in the right-hand side of Eq. 4 rather than to adjust Eq. 9. This makes things consistent while maintaining the original meaning of precision.

p. 4977, l.12: Some lines of justification and motivation why the correct error estimation is important will be added.

p. 4984, l.5: I think that without mentioning of the weighted mean in bias determination the paper would be somewhat incomplete but I agree that this feature causes confusion because it is mentioned before Eq. 28 where the chi square test is discussed. I will reorganize the section such that first the standard bias determination scheme is fully discussed (including the chi square test), and the aspect of inversely variance-weighted contributions to the bias is mentioned at the end of the section. After this reorganization, the logical flow of the standard scheme of bias determination will not be interuppted by this side aspect. I feel that, since the rationale of chi suare testing will be discussed in the beginning of Section 3.1, this needs not be repreated here again.

p 4985, l4: agreed; this paragraph will be reworked.

p 4985 l20: agreed, will be corrected

p 4987 l8, 10, and 14: In my current working version of the manuscript, the subsection has been entirely rewritten in reply to review #2 and should now be less confusing and more consistent. Indeed, there were errors in this section.

p 4988 l19, agreed, reference will be added.

p 4989 l3, agreed, will be changed.

p 4990 l12: indeed it is not possible to prove that two distributions are the same. How-

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

EGU

ever one can give evidence that the probability that they are different is below a certain threshold. The difference between "necessary validation" and "sufficient validation" basically is the threshold, not the method. A probability that the distributions are different of slightly less than 95 % is small enough to reject the hypothesis that the profiles are different, but to give evidence that the profiles are probably the same, this probability should better be below 5 %. The section will be rewritten to make this clear. The splitting of the large data set will be used as an example that a subset may allow only "necessary validation" (= failed falsification) while a larger data set may allow "sufficient validation"). The splitting of the data set certainly does NOT give more confidence and does NOT make the difference between "sufficient" and "necessary" validation.

Technical corrections: all agreed; will be corrected.

I would like to take this opportunity to mention another error in the ACPD version which will be corrected: Error propagation of differences of uncertain quantities requires that 2 times the natural variability is subtracted, not only one.

Further, in the context of the F-test, a caveat will be added that this test is particularly sensitive to deviations of the actual distributions from normal distributions.

Interactive comment on Atmos. Chem. Phys. Discuss., 6, 4973, 2006.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

EGU