Atmos. Chem. Phys. Discuss., 6, S1535–S1543, 2006
www.atmos-chem-phys-discuss.net/6/S1535/2006/

Atmospheric Chemistry and Physics Discussions

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

Anonymous Referee #2

Received and published: 5 July 2006

1 General comments

The paper "Validation of remotely sensed state variables: strategies and terminology" describes the theory, the formal equations and the terminology that should be used in the validation of remotely sensed atmospheric profiles. While the statistical methods illustrated in the paper are, in several cases, described also in statistics textbooks, some important generalizations are presented, which are required when handling correlated measurements and in the intercomparison of measurements which do not refer exactly to the same atmospheric state. I think that this paper will be very useful to the scientific community validating atmospheric measurements: this is especially because the scientists taking care of validation are usually experimentalists, experts of the *refer-*

ence measurements used for the validation, but are not very familiar with statistics and with the formalism required for handling the various error components of the involved measurements.

In my view, for the importance of the subject dealt with and for its own nature, this paper belongs to the category of “review papers" rather than to the category of “technical notes" as stated by the author itself both in the abstract and in the conclusions.

The paper is well structured and I would recommend it also for publication on ACP, however I found several errors in the presented equations (as specified below) and these should be corrected before the final publication. I hope that the author itself will carefully check once again the presented equations as I cannot be sure that I got 100% of the errors.

I also feel that including in the text a few more explanations as outlined below would improve significantly the readability of this paper which deals with a rather difficult subject. In particular, I feel that for each considered validation approach it should be stated more clearly which are the errors determined with the statistical estimators (i.e. using validation and reference measurements), and which are the errors we are validating and of which we know an estimate prior to the validation work. Also it would be a good idea to mark e.g. with a circumflex $(\hat{})$ all the errors derived from the estimators presented in the paper.

2 Specific comments

1. p. 4977, l. 10: the smoothing error of the difference, $\mathbf{S}_{smooth,diff}$ has to be estimated (Rodgers and Connor, 2003) and eventually minimized or nullified (see e.g. M. Ridolfi, S. Ceccherini and B. Carli, “Optimal interpolation method for intercomparison of atmospheric measurements", *Opt. Lett.*, **31**(7), 855–857, (2006)).

2. p. 4977, l. 15: I suggest to define $\mathbf{S}_{coinc}$ immediately after equation (11), not far away like now.

3. p. 4977, l. 17: This may apply ... please change this sentence into something like: “This applies e.g. when the same (or correlated) temperature profiles or spectroscopic data are used to derive both measurements".

4. p. 4978, l. 9: d is introduced here as a scalar quantity, however, as mentioned later in the paper, in general it is a vector. I think it would be better to introduce $\mathbf{d}$ as a vector since the beginning and then to treat the “easy" cases in which it can be handled as a scalar.

5. p. 4978, l. 24, 25: *... and systematic sampling errors may inadvertently be treated as random coincidence errors*. I was not able to understand this statement, what are the *sampling errors* ? In which domain ? Please explain more clearly what is meant here.

6. p. 4979, l. 5, 6: *... sufficiently fine resolved typical reference distribution ...* In which domain should be “fine resolved" the distribution ? (space, time, or ...). To what kind of *distribution* is the author referring to ? A statistical distribution or simply the spatial distribution of a selected atmospheric parameter ? Please specify.

7. p. 4979, Eq. (14): from Eq.s (15) and (16) it seems that the experimental data are not used here to calculate average values, therefore the denominator of Eq. (14) should be I and not $I-1$.

8. p. 4979, l. 15, 16: If I understand properly, $\mathbf{S}_{distribution}$ is the covariance matrix characterizing the error of the reference distribution $\widehat{\mathbf{x}}_r$ (not of $\widehat{\mathbf{x}}_c$ as stated at l. 16 ...), this covariance should be known a-priori, it does not matter the method that was used to determine it. I.e. this covariance could have been determined

ACPD

6, S1535–S1543, 2006

Interactive Comment

e.g. with model sensitivity tests or could be just a guess, it is not necessarily determined through error propagation. The important issue here is that $\mathbf{S}_{distribution}$ is a known quantity and it is not determined on the basis of the measurements we are comparing.

9. p. 4980 l. 2: What is $\mathbf{S}$ without subscripts here ? I guess it is meant $\mathbf{S}_{coinc}$.
10. p. 4980, Eq. (17): The lower left element of the matrix should be either $\mathbf{C}_{coinc;2,1}$ or $\mathbf{C}^T_{coinc;1,2}$. I. e. $\mathbf{C}^T_{coinc;2,1}$ seems too much ...
11. p. 4982, Eq. (20): according to the conventions established earlier, $\mathbf{x}_r$ represents a vertical profile. However, the horizontal averaging kernel $\mathbf{A}_{hor}$ should be multiplied by an horizontal field ... therefore Eq. (20), and consequently also Eq. (21), are wrong (or at least misleading) as written now.
12. p. 4983, Eq. (23): Eq. (23) is rather “known" as it is the estimator for the covariance of an average value. Therefore this equation should be understandable and accepted even if reported without the intermediate step which I found misleading, i.e. I was not able to understand the reasoning indicated by the intermediate step.
13. p. 4984, Eq. (25): the coincidence error is a characteristic associated with the difference between two measurements. This kind of error is not associated to an individual measurement, therefore I recommend using the symbol $\mathbf{S}_{coinc}$ rather than $\mathbf{S}_{val,coinc}$ to indicate the covariance matrix of the coincidence error.
14. p. 4984, Eq. (26) and (27): although these equations are correct I do not recommend using them as they are based on the covariance $\mathbf{S}_{diff,random,k}$ that is being assessed in the validation work.
15. p. 4984, l. 13: The consistency to be checked is between $\widehat{b}_n$ and the a-priori known estimate of the systematic error of the profiles difference. Please state clearly this concept. I also recommend to indicate this latter error with the symbol

ACPD

6, S1535–S1543, 2006

Interactive Comment

$\sigma_{diff,sys;n}$ and not with the symbol $\sigma_{sys,n}$ which, according to Eq. (8) is already assigned to the systematic error of an individual measurement.

16. p. 4984, Eq. (28): please replace $\mathbf{S}_{bias}$ with $\mathbf{S}_{bias}^{-1}$ here.
17. p. 4985, Eq. (31): In the summation at the denominator, please replace $\widehat{x}_{ref,k}$ with $\widehat{x}_{ref;n,k}$.
18. p. 4985, Eq. (32): I was not able to understand the reasoning used to derive this equation. Anyhow, I have derived this equation by applying the error propagation to Eq. (31) re-written as a function of the available independent measurements $\widehat{x}_{ref;n,k}$ and $\widehat{x}_{val;n,k}$. I found that Eq. (32) is correct, provided that $\bar{\widehat{x}}_{ref,n}^{2}$ is replaced by $\bar{\widehat{x}}_{ref,n}^{4}$ at the denominator.
19. p. 4986, Eq. (34): Please make sure that the index k numbering the measurements within the sample is always reported as the rightmost index in all the equations. In this equation as well as in Eq.s (38) and (41) the position of the index k is not consistent with the convention used for the other equations of the paper. Please note also that the bias should not have the index k, therefore, please replace $\widehat{b}_{k,n}$ with $\widehat{b}_{n}$ in Eq. (34). Finally, to be precise, r should have a sub-script n to indicate that it depends on the altitude.
20. p. 4986, l. 17: Please define also L here.
21. p. 4987, Sect. 5.1: The symbol $\langle x \rangle$ is used in statistics to denote the *Expectation Value* of a random variable x. If we do not know the probability density function of x we can try to estimate $\langle x \rangle$ from experimental measurements, using a statistical estimator formula, but it should be clear that this is an approximation. This is in fact the case explained in this Section of the paper, therefore I would not recommend the use of the symbol $\langle\rangle$ in the equations of this Section. In general

I found this Section quite confusing, I would modify it according to the following guidelines.

First we want to verify that the a-priori estimated total error on the profile difference ($\mathbf{S}_{diff}$) is consistent with the estimate derived from the intercomparison. This can be done by testing the probability of the following chi-square (that should replace Eq. (39) of the paper, I would remove Eq. (38) which contains also errors):

$$\chi^2 = \sum_{k=1}^{K} (\widehat{\mathbf{x}}_{val;k} - \widehat{\mathbf{x}}_{ref;k})^T \mathbf{S}_{diff}^{-1} (\widehat{\mathbf{x}}_{val;k} - \widehat{\mathbf{x}}_{ref;k}) \tag{1}$$

whose expectation value is: $\langle \chi^2 \rangle = NK$. This chi-square test fails if the bias is different from zero. In this case the bias should be preliminarily evaluated using Eq. (22) and, subsequently, the a-priori estimated random error on the difference $\mathbf{S}_{diff,random}$ should be tested for its consistency against the r.m.s. fluctuation of the profiles difference about the bias. This can be done by checking the probability of the following chi-square (that should replace Eq.(40)):

$$\chi^2 = \sum_{k=1}^{K} (\widehat{\mathbf{x}}_{val;k} - \widehat{\mathbf{x}}_{ref;k} - \widehat{\mathbf{b}})^T \mathbf{S}_{diff,random}^{-1} (\widehat{\mathbf{x}}_{val;k} - \widehat{\mathbf{x}}_{ref;k} - \widehat{\mathbf{b}}) \tag{2}$$

whose expectation value is $\langle \chi^2 \rangle = N(K-1)$ i.e., compared with the previous case, one full profile of N elements is lost for the determination of the bias profile.

22. p. 4988, Eq. (41): This equation also seems to have some errors, I think it should read exactly as Eq. (23) and therefore it is unnecessary here. Note that this conclusion (as well as the one in the previous comment) are based on the assumption that Eqs. (22) and (23) are used for the determination of the bias and of its error. I do not recommend the use of Eqs. (26) and (27) in the context of validation.

ACPD

6, S1535–S1543, 2006

23. p. 4988, l. 10: It should be stressed that the samples considered for this type of validation should be very large and really *random*. If the considered samples refer to different seasons or latitudinal bands $\mathbf{S}_{val;nat}$ and $\mathbf{S}_{ref;nat}$ can easily be different. Could the author cite an example in which this validation approach was used ? Alternatively it should be stated clearly that this approach is valid theoretically but in practice it will be very difficult to use.

24. p. 4988, Eqs. (45) and (46): I argue that $s_{m,n}$ should read $s_{val,sample,m,n}$ in Eq. (43) and $s_{ref,sample,m,n}$ in Eq. (44). Furthermore in Eq. (43) the index k should be added to $\widehat{x}_{val,m}$ and to $\widehat{x}_{val,n}$. Similarly, in Eq. (44) the index l should be added to $\widehat{x}_{ref,m}$ and to $\widehat{x}_{ref,n}$.

25. p. 4989, l. 3, 4: I would reword here: *If only a single reference profile is available and it does not coincide with any of the measurements to be validated, we can check whether this single reference profile belongs to the statistical distribution defined by the sample (of size K) of the measurements to be validated*.

26. p. 4989, Eq. (45): write $\chi^2 =$ ..., furthermore the inverse sum of the covariances should be used here in place of the sum of the covariances.

27. p. 4990, l. 8: The multiplication axiom holds only for independent measurements. However, often the measurements we are dealing with are not independent, see e.g. the case of profiles retrieved with the same set of spectroscopic line data .... Therefore I understand that, before undertaking the χ^2-test and applying the multiplication axiom, first we must make sure that the considered measurements are really independent from each other or, alternatively, a debiasing should be applied.

28. p. 4990, l. 26 and p. 4991, l. 7, 8: The integrals of the χ^2 probability density function are usually tabulated in statistics textbooks only for number n of degrees of freedom of the χ^2 less than 20 or 30. This happens because for the theorem

of central limit, already for $n > 20$ the χ^2 probability density function cannot be distinguished from the Gaussian normal distribution and therefore the searched probability can be easily obtained from tables reporting the integral of the normalized Gauss function. In conclusion, in the paper I would not mention this problem of lack of tabulated χ^2 probabilities because it is not a real problem.

3 Minor suggested corrections

1. p. 4974, l. 5: change “single" into “individual".
2. p. 4974, l. 18: and *of* a reference instrument.
3. p. 4977, l. 7: regridding of one *or* both profiles.
4. p. 4978, l. 5: Comparison of *two individual* profiles ...
5. p. 4978, l. 11: Here and in several places in the paper, change “criterium" into “criterion".
6. p. 4979, l. 1: *ECMWF*: expand the acronym the first time it is used.
7. p. 4981, l. 14: change “independent" into “independently".
8. p. 4982, l. 7: perturbational
9. p. 4982, l. 21, 22: I would say: *..., this latter approach can be applied in both directions*.
10. p. 4982, l. 24: remove “relative".
11. p. 4983, l. 6: *... measurements of each measuring system, respectively*.

12. p. 4983, l. 13: estimated *as*...
13. p. 4983, l. 16: *... the elements of which are estimated as ...*
14. p. 4988, l. 17: change “particular" into “particularly".
15. p. 4989, l. 9: I suggest rewording like: *Its elements are estimated as:*
16. p. 4989, l. 15: “accidently" ??
17. p. 4990, l. 7: Replace “5-%" with “5%". Do the same at l. 11.
18. p. 4990, l. 7: Replace “than" with “then".
19. p. 4990, l. 5–15: please use consistent notations for $P(K, \chi^2)$, i.e. K should always be the first argument and always uppercase.
20. p. 4990, l. 17: change “F-test using" into “F-test to be used".
21. p. 4990, l. 19: change “single χ^2-test for the entity of measurements" into “single χ^2-test for the full set of measurements".
22. p. 4991, l. 7: large *number of* degrees of freedom. See also the more general comment above.
23. p. 4991, l. 16: *If* the sufficient validation fails, ...
24. p. 4993, l.9: The more *conservative* approach to give evidence ...
25. p. 4993, l. 10: I would be more quantitative here by changing “is small" into “is less than 5%".
26. p. 4993, l. 13: here and in the abstract, p. 4974, l. 2, I would replace “technical note" with “review work" or simply “work". See also the general comment above.

ACPD
6, S1535–S1543, 2006

Interactive Comment

EGU