Atmos. Chem. Phys. Discuss., 6, S2020–S2025, 2006
www.atmos-chem-phys-discuss.net/6/S2020/2006/

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 #3

Received and published: 21 July 2006

1 General comments

The paper "Validation of remotely sensed profiles of atmospheric state variables: strategies and terminology" sets up a general framework for validation of remotely sensed atmospheric profiles. It generalizes the idea to compare coincident measurements of two independent instruments to the case of less than perfect coincidence and even further to the case of completely independently taken data sets. Especially the rigorous treatment of the error arising from non-perfect coincidence between data pairs is new and will help scientists to deal with this frequently occuring situation. The paper is very well structured and gives precise recipes how to perform the validation process. However, for my feeling the application of the χ^2 test needs some further explanation close to the begin of the paper, because it is the central tool for the validation process. In addition there are several mistakes in the formulas, but they can be easily corrected.

2 Specific comments

p. 4974, l. 10

The null hypothesis seems to be defined opposite to the common definition: Usually the null hypothesis assumes that the difference between two measurements is not significant, that means that the two measurements are drawn from the same distribution (see e.g. Chap. 14 in *Numerical Recipes in C: The Art of Scientific Computing*, W.H. Press et al., Cambridge University Press., 2nd edition 1992, and see also Chap. 27 in *Kleine Enzyklopädie Mathematik*, W. Gellert, H. Küstner, M. Hellwich, and H. Kästner (editors), Verlag Harri Deutsch, 2nd edition, 1984).

p. 4976, l. 17

The precision p as defined in Eq. (4) is the well known standard deviation of the measurement; in Eq. (9) this quantity is expected to equal the diagonal elements of the random error covariance S_{random}, which usually should be provided along the remotely sensed data. However, if the scatter of a sample of measurements is assumed to be composed of both the measurement random error and the natural variability – as discussed later in section 5.2 – and if in addition the random error covariance of the data set is derived just from instrumental noise by means of error propagation through the data analysis process, then the two quantities do not need to be equal. In this case we would have

$$p_n^2 = \sigma_{random;n}^2 + \sigma_{nat;n}^2$$

ACPD

6, S2020–S2025, 2006

Interactive Comment

p. 4977, l. 12

I think one of the key features of this paper is the systematic application of the χ^2-test in the validation process; to my understanding the purpose of this test is first to check the consistency of a specific error and the covariance matrix describing the size of this error, and second to test if a the bias as defined in Eq. (2) is statistically significant. This should be explained briefly when introducing the χ^2-test in Eq. (10); otherwise the reader might not understand, why in the following part so much effort is put into the carefull estimation of the coincidence error covariance matrix.

p. 4984, l. 5

Eq. (25) to (27): Using $S_{diff,random;k}$ as weights in the bias determination: Throughout the paper a certain error source is characterized by usually a single covariance matrix. However, assigning an index k to the covariance matrix $S_{diff,random;k}$ means that a separate covariance matrix exists for each pair of coincident measurements. Introducing this weighting feature at this point of the text doesn't help in understanding the main idea of this section; instead it even may confuse the reader. My feeling is to better leave this feature completely away and instead lay particular stress on the very important χ^2 test shown in Eq. (28).

p. 4985, l. 4

The extension of the statistics of an additive bias to a multiplicative bias is not so straightforward: I don't understand why the mean relative deviation has been chosen by Eq. (29) and not by Eq. (31). In other words: Which idea makes Eq. (29) to be preferred to Eq. (31)? As said in the text the mean relative deviation is always different from the relative mean deviation. Actually Eq. (29) gives a result which depends on the variance of $\hat{x}_{ref}$ because the devision by $\hat{x}_{ref}$ is a non-linear operation and causes a non-Gaussian distribution of the quantity $\hat{x}_{val;n,k}/\hat{x}_{ref;n,k}$. Thus it makes more sense to first average the additive bias and the mean reference value in order to get the true numbers as close as possible

ACPD

6, S2020–S2025, 2006

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

EGU

and afterwards devide these two numbers to get the correct ratio.

p. 4985, l. 20

The correlations in Eq. (32) must be considered only if the errors of $\hat{x}_{val;n}$ and $\hat{x}_{val;n}$ are actually correlated – not because the nominator and denominator include common terms.

p. 4987, l. 8, 10, and 14

In Eqs. (38), (39), and (40) the notation with angle brackets is confusing; with the exception of this section the sum over k has been always explicitly written in the formulas, thereby making very clear what to sum up. I suggest to keep this style also for this subsection.

p. 4988, l. 19

Some reference about Bayesian statistics and its meaning would be helpful for a reader being new in the field of retrieval techniques.

p. 4989, l. 3

If there is only a single reference profile available then it is better to write "If only a single profile measurement is available which does not coincide ...".

p. 4990, l. 12

The introduction of the "sufficient validation" may be confusing, because it sounds as if it is possible to statistically prove that two distributions are the same. It is only possible to show that two distributions are very likely not the same. All what is suggested here is to split a large data sample into several smaller subsamples and to check the χ^2 distribution. Of course then this distribution gives more information than a single χ^2 from only a sub-sample. Indeed this is clearly said in the text: "Alternativly one can also perform a single χ^2 test for the entity of measurements.". So the sufficient validation doesn't give more confidence than the standard statistics already provides; it is just a different view of the same data

ACPD

6, S2020–S2025, 2006

sample. To be clear: There is nothing wrong with this section, just the name "sufficient validation" may be misleading.

3 Technical corrections

p. 4983 Bias determination

probably typing error in the line before Eq. (22): "... the bias b is estimated as" (instead of "at")

p. 4984, l. 16

In Eq. (28) S_{bias} must be replaced by its inverse S_{bias}^{-1}.

p. 4985, l. 21

In Eq. (32) it must be devided by $\bar{\hat{x}}_{ref;n}^{4}$ instead of $\bar{\hat{x}}_{ref;n}^{2}$.

p. 4986, l. 18 and p. 4987, l. 1

In Eq. (37) $x_{val;m}$ and $x_{val;n}$ need an additional index k in order to be consistent with Eq. (23); in the same way $x_{ref;m}$ and $x_{ref;n}$ need an additional index l.

p. 4988, l. 3

In Eq. (41) the squared covariance $s_{bias;m,n}^{2}$ must be the simple covariance without squaring.

p. 4988, l. 13 and 15

In Eqs. (43) and (44) the elements $S_{m,n}$ should have an additional index val and ref. In addition there is a typo in Eq. (43) ("nal" instead of "val").

p. 4989, l. 7

In Eq. (45) the covariance matrix in the round brackets must be replaced by its inverse (an exponent -1 is missing).

ACPD

6, S2020–S2025, 2006

Interactive Comment

p. 4989, l. 7 and l. 11

In Eqs. (45) and (46) the overbar should be consistent with Eqs. (43) and (44) and should not extend over the index val.

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

ACPD

6, S2020–S2025, 2006

Interactive Comment