Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2019-32-RC2, 2019

Atmospheric Measurement Techniques Discussions

Interactive comment on "The cost function of the data fusion process and its application" *by* Simone Ceccherini et al.

Anonymous Referee #1

Received and published: 9 March 2019

This paper is a continuation of the series of papers authored by Ceccherini et al. over the last few years. It continues the good work describing the features of the complete data fusion (CDF) method. As such, it is worthy of publication, provided the authors address the following general and specific comments.

The main general comment is the suggestion that the authors transfer even more of the technical discussions (e.g., in Sect. 3), including representation of equations, to an appendix (I have not checked all the mathematical workings, but they seem correct to me). Whether this should take place is a decision between the authors and the editor.

Generally, the authors have done a good job of writing the paper. I have a minor correction:

C1

P. 10

L. 2: probably -> likely.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-32, 2019.

Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2019-32-AC1, 2019

Atmospheric Measurement Techniques Discussions

Interactive comment on "The cost function of the data fusion process and its application" *by* Simone Ceccherini et al.

Simone Ceccherini et al.

s.ceccherini@ifac.cnr.it

Received and published: 30 April 2019

The comment was uploaded in the form of a supplement:
https://www.atmos-meas-tech-discuss.net/amt-2019-32/amt-2019-32-AC1-supplement.pdf

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-32, 2019.

Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2019-32-RC1, 2019

Atmospheric Measurement Techniques Discussions

Interactive comment on "The cost function of the data fusion process and its application" *by* Simone Ceccherini et al.

Anonymous Referee #2

Received and published: 8 March 2019

The paper by Checcherini et al. covers a topic that fits well within the scope of AMT. It seems scientifically sound to me in most of its parts. The paper is well written, and the text between the equations guides the reader safely through the maths. The notation chosen in some places seems a bit odd to me. I suggest publication after clarification of the following issues.

p1 l18 ff: The first paragraph of the intro is very informative and clear. If helps to pick up the thread. Congratulations!

p1 l26: To me it is counter-intuitive that the equivalence holds also in nonlinear cases. Does this statement need a qualification "equivalent ... in the linear case, and

approximately equivalent in the moderately non-linear case"? I have mentioned this already in the access review and since the authors did not change anything I suspect that this has been done by intention and I may have missed the main argument. Just to make sure that things are correct, I suggest to either add "within linear theory" to the text, or to provide (in the rebuttal, not in the paper) a short justification why this statement holds also in the non-linear case.

p1 l35: It seems to me to be ambiguous what the interpolation error shall represent. Is it (a) that fine structures of the true profile are not represented by the profiles sampled on a coarse grid, and that interpolation to a finer grid will not recover these fine structures? If so, I would object that these issues are already covered by the averaging kernel of the coarse grid representation, and it is not clear to me why an extra treatment is needed. Or (b) is the source of this inconsistency an inconsistency of different interpolation methods. Here it is important that data points of a profile retrieval do not represent a single point in the atmosphere but that the value retrieved at this point depends on the profile assumed for the entire layer between two levels. This is because the forward solution of the radiative transfer equation is basically an integration along the line of sight, and thus, assumptions of the atmospheric state between the levels are used (regardless if an Curtis-Godson approach or a direct integration method is chosen). For this purpose, the forward model typically uses an internal interpolation in the retrieval. The solution of the retrieval based on this forward model is conditional to this internal interpolation or interpolation scheme or assumption. Assuming a different behavior between the layers would result in different retrieved values at the model levels. I think it needs to be elaborated why an interpolation uncertainty is introduced if an interpolation scheme is used which is consistent with that in the forward model. Or do the authors assume that the interpolation of retrieved profiles is made with another interpolation scheme than that used in the forward model?

p2 l20: Here it is not clear what the term 'error' means. Is it a statistical estimate or is it the actual difference between the actual retrieved value and the true (or expectation) value? In other words, do the sigmas represent a statistic or an instantiation (or realization) of it? Normally, the symbol σ is used for a statistic, namely the standard deviation. But this seems to make no sense in this context because the sigmas themselves are arguments of the expectation value function. If the errors (sigmas) are understood to be actual differences, then the paragraph makes sense to me but I challenge the adequacy of the notation. The symbol σ is, in the context of error analysis, always used for a statistic, namely, a standard deviation, and it took me multiple readings of this paragraph to find out that the authors seem to mean something else. Our alphabet has so many letters available, each of them as capital and small character, and we have multiple alphabets available (Latin, Greece, ...). For the actual differences, d or delta would be an option but there are many others. Why choosing without need a notation which will almost certainly lead the reader astray?

p2 Eq 3: Here it becomes clear that sigma is not a standard deviation but a single instantiation of noise mapped from the signal space into the atmospheric state space. I consider this very unfortunate and misleading because sigma is usually reserved for the standard deviation; see my comment above.

p2 Eq 4: I consider $\mathbf{F}$ as an unfortunate choice of symbol for the Fisher information matrix, because in Rodgers' notation this symbol stands for the (vectorial) forward model. Since such a formal paper, even without notation-related complication, is hard enough to digest, I think, one could make the reader's life easier by choosing another letter in order to avoid confusion.

p3 Eq 11: According to p 2 line 25, $\mathbf{S}_\mathrm{a}$ seems the a priori covariance represented on the grid on which the retrieval is performed. It does not contain any variability on any finer scale (see von Clarmann, AMT 7, 3023-3034, 2014 for similar considerations regarding the role of $\mathbf{S}_\mathrm{a}$ in the calculation of the smoothing error.) The $\mathbf{C}^{(\mathrm{f})}$ matrix involves a finer grid. Is thus the $\mathbf{S}_\mathrm{a}$ matrix, which is evaluated on the retrieval grid, sufficient for the

purpose of Eq. 11? This is not immediately intuitive and needs some discussion. Or have you tacitly redefined $\mathbf{S}_\mathrm{a}$? If so, this should be stated explicitly. This equation seems to give the answer to my question about the meaning of the interpolation error (p1 l35), and the interpretation of the interpolation error seems to be option (a). Given the problems caused by the grid-dependence of $\mathbf{S}_\mathrm{a}$, wouldn't it be safer to dismiss the idea to characterize the interplation effect as a covariance matrix and to provide an averaging kernel of the resulting fused profile which includes also the interpolation effects instead?

Section 3 ff: I cannot guarantee that all the manipulations of equations are flawless but the general rationale behind seems compelling to me.

Technical issues: Subscripts of matrices which are neither variables nor indices but denote simply the name of the matrix should not be printed italic but roman (e.g. the a in $\mathbf{S}_\mathrm{a}$, the 'int, coin, other' in the $\mathbf{S}_{i,\mathrm{int}}$, $\mathbf{S}_{i,\mathrm{coin}}$, $\mathbf{S}_{i,\mathrm{other}}$, respectively, etc).

The similarity index appears to be quite high but I consider the detected similarities as insignificant. The index seems to be very sensitive to standard formulations that everybody uses.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-32, 2019.

Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2019-32-AC2, 2019

Atmospheric Measurement Techniques Discussions

Interactive comment on "The cost function of the data fusion process and its application" *by* Simone Ceccherini et al.

Simone Ceccherini et al.

s.ceccherini@ifac.cnr.it

Received and published: 30 April 2019

The comment was uploaded in the form of a supplement:
https://www.atmos-meas-tech-discuss.net/amt-2019-32/amt-2019-32-AC2-supplement.pdf

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-32, 2019.