We thank the reviewer for the useful comments. In the following, we answer the specific comments (included in "**boldface**" for clarity) and, whenever required, we describe the related changes implemented in the revised manuscript. Indicated page and line numbers refer to the original version of the paper published on AMTD.

Anonymous Referee #2

Review

GENERAL COMMENTS

The paper is well written and describes the employed methods in sufficient detail. With the last section, the method is applied to a practical example and the differences are examined.

The topic fits the journal well.

The method, as extended in Sect. 2.3 is particularly useful when joining satellite measurements taken on different grids and with co-location errors, where only certain diagnostic matrices are provided.
I suggest publication after properly positioning the CDF method as multi-variate inverse CM-weighted mean and addressing the major and specific comments below.
We have considered all the major comments listed below in order to show that, as suggested by the reviewer, the CDF can be described as a multivariate inverse CM weighted mean. Therefore, in the revised version of the paper, we have used the equations suggested by the reviewer, which provide useful information on the physical meaning of the CDF. However, these equations cannot be used to perform data fusion and we preferred to put this discussion in an appendix (Appendix B), in order to not interrupt the flux of the reasoning.

MAJOR COMMENTS
==============

line 76

I think it might help here the understanding to introduce the relation of
\hat{x} = A x_true + (I-A) x_a + G \epsilon
as this shows more readily the nature of the formula: a weighted average of the true state transformed by the different measurement characteristics of the involved instruments:
x_f = (\sum S^-1_i A_i + S_a^-1)^-1 (\sum S^-1_i (A_i x_true,i + G_i \eps_i) + S_a^-1 x_a).
which also leads pretty naturally to the derivation of the aggregated averaging kernel matrix.
In the revised version of the paper, we introduced these two equations in Appendix B.

The formula above as well as (10) - (12) can also be simplified drastically by exploiting that
S_i^-1A_i = F_i,
which mathematically is very reasonable and fits well to the general framework of optimal estimation and Kalman filtering.
In the Appendix B we wrote the CDF formulas as suggested by the reviewer and discussed them.

Is the whole method, in its given form, not fully identical to a "simple/straightforward" linear optimal estimation/maximum likelihood estimate of all involved instruments *linearized* around the individual solutions? Which is indeed very reasonable, but not really a "new" method.
The new mathematical description makes this pretty obvious in contrast to the original, more convoluted formula.
We agree with the reviewer that the method can be seen as an optimal estimate obtained by all the considered measurements linearized around the individual solutions. We wrote this in the Appendix B.

The given mathematical notation can be argued for due to the information supplied by typical retrieval products, but both forms, the "standard" form using the (inverse) Fisher information matrix as weight in a weighted mean and the given form should be described and compared against each other.
In the Appendix B of the revised version of the paper we wrote the equations of CDF using the Fisher information matrix as weight in a weighted mean and compared these expressions with the original ones given in the manuscript.

The authors should discuss this and how it differs (or not) from the method described, e.g., by Rodgers in Sect. 4.1.1.
In The Appendix B of the revised version of the paper we discussed the connection between CDF and the method described in Section 4.1.1 of Rodgers (2000).

SPECIFIC COMMENTS
=================

line 23

You stated that the method delivers the same result as a simultaneous retrieval, so in what respect or in relation to what can its quality be better?
The implementation of the CDF is much simpler than that of the simultaneous retrieval. Indeed, the simultaneous retrieval requires to integrate into a single inversion system the radiative transfer models capable to simulate the measurements of the different sensors involved in the simultaneous inversion. Furthermore, the simultaneous retrieval requires the simultaneous access to all the (Level 1) measurements used in the inversion, thus implying the need to handle relevant data volumes. These characteristics complicate the implementation of the simultaneous retrieval and increase the necessary computational resources. The CDF overcomes these complications by combining the Level 2 products supplied by the individual retrieval processors of the measurements.
We added this consideration in the revised version of the paper.

line 48

Didn't you just state that the formula was introduced by Ceccherini (2021)? So it isn't introduced here, "only" discussed in greater detail?
In Ceccherini (2022) the formula was derived only for the fusion of two profiles, here we generalize the formula for any number of profiles and discuss it in greater detail.

In the revised version of the paper, we change the sentence specifying better what we have done here. Furthermore, we change also line 44, specifying that in Ceccherini (2022) the formula was derived only for the fusion of two profiles.

In fact, Ceccherini (2021) seems to suggest that the formula was introduced by Schneider (2021)?
The correct statement is that: starting from the Kalman filter method as applied in Schneider (2022), Ceccherini (2022) derived the new formula of the CDF for the fusion of two profiles.

I think the historical development and relationship between the papers and methods should be discussed in slightly more detail than given here, taking into account in particular other peoples contributions.
In the revised version of the paper, we changed the sentence and added the reference to Schneider et al. (2022).

line 233

Is the Python code with a reference implementation available? I.e. can the results of Section 3 be reproduced?
The Python code used in the paper is not user friendly and, accordingly, is not made available, but the results in Section 3 can be easily reproduced implementing formulas in Eq. (33-34). In the revised version of the paper, we provide the references to the Python functions used to calculate the eigenvalues and the generalized inverses.

MINOR REMARKS
=============

line 7

Who has proposed it?
It was proposed in Ceccherini (2022), but since in the abstract reference citations have to be avoided, the reader finds this information in the introduction, that in the revised version has been changed to better clarify the historical development and relationship between the papers and the method (see answer to the comment above).

line 30

Performances ... have -> performance has
In the revised version of the paper, we changed the sentence as suggested by the reviewer.

line 39/43

I would say that while Rodgers provides a very useful discussion on the use of Kalman filters for the use case at hand, it is not a suitable reference without also giving (Kalman, 1960; see Rodgers). Are the references in lines 39 and 43 switched?

In the revised version of the paper, we added the reference to Kalman (1960).
The reference in lines 39 and 43 are both correct.

line 66

The readability of the formulas could be greatly improved when the "^-1" notation of the involved matrices would be above the index, not after it.
In the revised version of the paper, we moved the “^-1”, “*T*” and “#” notations of the involved matrices from after the index to above it.