Comment on amt-2021-122
Anonymous Referee #3

The paper from Sheng Li and Ke Du proposes a new minimum curvature (MC) algorithm to apply smoothness constraints in the tomographic inversion of optical remote sensing measurements, to reconstruct the spatial distribution of atmospheric chemicals in a given domain (a 40 x 40 m square area in the example given). The authors compare the performance of their new proposed method to that of other existing methods, such as the non-negative least squares (NNLS) and the low third derivative (LTD). The performance is assessed on the basis of a few test maps containing one or several (up to five) bi-variate Gaussian sources. Apparently, the MC algorithm performs significantly better than the NNLS method, and shows almost the same performance of the LTD algorithm in terms of reconstruction accuracy. Compared to this latter, however, the MC algorithm allows to save from 27 to 35% computation time, depending on the number of sources in the domain.

The subject of the paper is interesting, comprehensively presented in the introduction and put in the context of the existing literature on the topic. The method used for the assessment, however, is not sufficiently general and could be improved. The presentation of the algorithms assessed is not sufficient, the actual equations used should be included in the paper. Regarding the language of the text, I am not native English speaker, thus I cannot provide a reliable feedback. However, the language sounds a bit “strange” to me at several instances. Therefore I recommend a review by a language Editor. Also, I would suggest to avoid flooding the text with acronyms. Several of them are not really necessary and make reading the paper uncomfortable.

In conclusion, I am very sorry but I can recommend this paper for publication in AMT only after major improvements, as outlined in the comments below.

General comments

My main concern is that the authors have compared the field reconstruction errors of the NNLS, LTD, MC (and GT-MG) algorithms on the basis of a set of only five test distribution maps. I would say that they have verified some necessary conditions which, however, are
not sufficient to assess the relative efficiency of the considered methods. Each solution depends both on the measurements and on the constraints applied. Here it is not clear whether the LTD and MC solutions perform better than the NNLS because of the smarter applied constraints or because of the specific experimental distributions (bi-variate Gaussians) considered in the examples given.

Error covariance matrices and averaging kernels (see e.g. Rodgers, 2000) are broadly used tools in the atmospheric remote sensing community to characterize the recovered spatial distribution (yes, also 2D distributions ...) from the point of view of the retrieval error, and of the spatial resolution (width of the Point Spread Function) of the measurement chain (measuring plus inversion systems). Applying these tools to the inversion methods considered in the paper is possible, thus the authors should use them. For example, from the analysis reported in the paper, it is not self evident that the spatial resolution of their measuring system changes strongly depending on the \(x,y\) position within the squared field considered: there are grid-elements which are crossed by 2 or 3 beams, and others (near the sides of the field) which are not sounded at all. Thus, the spatial resolution must be very poor near the sides of the squared field and much better near the center. I believe this feature would be self-evident from maps of the diagonal elements of the 2-dimensional averaging kernels of the different solutions considered (see e.g. von Clarmann et al, 2009).

Specific comments

Lines 42-44: please include references for the mentioned techniques. They are not standard for the whole atmospheric remote sensing community.

Lines 73-74: not only, I guess. The chosen pixels should be crossed at least by one beam, otherwise the NNLSF is ill-posed.

Line 122: I would name the small squares as “pixels” instead of “grids”.

Line 127: if the PIC is measured at the retro-reflectors, then you have only 16 measurements, or 20 measurements if retro-reflectors are installed also at the corners of the square. Instead, I guess that for the NNLSF you need at least 36 measurements. Please make an effort to describe more thoroughly the experimental setup.

Section 2.1: it would be interesting if the authors could explain the details of the experimental setup, I could not understand which is exactly the measured quantity. This would be useful also to understand to which degree the linear formulas (1) and (2) are accurate.

Equation (3) assumes that all the measurements have the same precision, which may not be the case if the signals observed are very different in intensity (e.g. because of the different absorption paths). Could-you please add a comment?

Line 139: With equation (3) you require a solution with “small” \(L_i\), \(c\) norm, whereas, usually one requires a small \(L_i (c - c_a)\) norm, where \(c_a\) is some prior estimate of \(c\). Please explain the rationale behind your choice of \(c_a = 0\).

Line 148: If the regularization parameter \(\mu\) of eq. (3) is grid-dependent, then I would expect it to appear in some vector form in equation (3) rather than as a scalar. How do you establish the actual value of \(\mu\)? Which is the solution of the LTD algorithm? Please specify the equation.
Line 149: In principle, the NNLS algorithm does not use constraints, correct? Here you are explaining the LTD algorithm, so it cannot be solved with the NNLS approach. Maybe you refer to the Newton method? Please explain.

Lines 158-165: Up to eq.(6) and later also in eq. (9), \( c \) is a vector. In eq.s (7) and (8) "c" seems a function. Please improve the notation, it would be difficult to implement your MC algorithm based on your description.

Line 167: I have understood that you are finally using eq.(9), that is the discretized form of eq.(8). Is eq.(9) more or less equivalent to eq.(7)? This description is very confusing.

Line 169: the same comment I made for \( \mu \) (line 148) here applies to \( \omega \). Which constant for the inverse proportionality did you use?

Line 187: If \( c(x,y) \) is a concentration, then \( Q \) cannot be measured in ppm (that is a mixing ratio). Eq.(10) does not contain \( \sigma \), it contains \( \sigma_x \) and \( \sigma_y \)...

Lines 212-213: this sentence is not clear to me.

Line 220: “The smaller the number of sources, the better the reconstruction quality”. I think this is intrinsic to the definition of the nearness quantifier. Please comment.

Lines 231-233: do you mean that the "peak-location" quantifier could mistake peaks with similar amplitudes? In this case it would be advisable to refine eq.(12) or to apply it with some caveats.

Line 233: I think that, as it is, eq.(12) is reliable only if the peaks to be reconstructed have amplitudes that differ from each other by much more that the error with which they are retrieved. Why the "source number" counts so much?

Line 239-240: I am skeptical about your general statement regarding the NNLS performance. I would suggest adding some details regarding how you actually computed the NNLS solution.

**Technical corrections**

Line 50: summarising ??

Line 85: necessary ?

Line 136: what is the superscript “21” ?

Line 174: “well-posed”

Line 212: it (?)

Line 224: maybe "complexity" ?

Line 228: are slightly better ...

Line 241: derivation ??
References
