

Atmos. Meas. Tech. Discuss., author comment AC4
<https://doi.org/10.5194/amt-2021-122-AC4>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC3

Sheng Li and Ke Du

Author comment on "A minimum curvature algorithm for tomographic reconstruction of atmospheric chemicals based on optical remote sensing" by Sheng Li and Ke Du, Atmos. Meas. Tech. Discuss., <https://doi.org/10.5194/amt-2021-122-AC4>, 2021

Response to comments of reviewer 3

We thank the reviewer for the helpful comments and suggestions, and for careful reading of the manuscript. Listed below are our responses, with the original comment/question displayed in italics.

Anonymous referee report

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.

Thank you for the comments and suggestions. We have done a major revision according to your suggestions. Please see the responses below for specific changes we have made to the manuscript.

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.

There are several things which need to be mentioned about this comparison.

(1) The performance of the algorithms is problem dependent. The beam configuration and underlying concentration distribution also affect the reconstruction quality. There is no one algorithm that performs absolutely better than others all the time. Even the NNLS algorithm could give the best result under some specific conditions (Wu and Chang, 2011). This is the common situation for an optimization algorithm.

(2) The bi-variate Gaussians are good choice to simulate the diffusion process. It has been commonly used in the comparisons of the algorithms in the ORS-CT problems. The generated maps better simulate distributions which have several hotspots. It is true that the real distribution map may be complicated depending on the specific dispersion conditions. Other methods to generate the test maps can also be used. For example, the bivariate lognormal distribution was used in Wu and Chang (2011) to simulate a non-symmetric distribution. But that only represents one situation, and the distribution function may not be better than others in other cases.

(3) Another reason that the test maps were generated in this way is for the purpose to compare with the results from the GT-MG algorithm, which used the bi-variate Gaussians. And the beam configuration was also set to be similar to that in the GT-MG approach. This is a common set for comparisons of tomographic reconstructions and was used in practical applications (Dobler et al., 2017).

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).

Thanks for the good suggestions. We have added the averaging kernel matrix into the manuscript to determine whether the concentration can be independently predicted and how regularization limits reconstruction accuracy. In the 2-D tomographic reconstruction, the averaging kernel is considerably affected by the beam geometry and is better to be used as a measure to evaluate the beam configuration. But it also reflects the regularization error given the same beam geometry in this study. We used a measure called fitness which was defined as the average Frobenius distance between the resolution matrix and the identity matrix to predicate the reconstruction error. The MC algorithm shows slightly better performance than the LTD algorithm with a fitness value of 1.3878 comparing to 1.4411.

The off-diagonal elements are not zeros. The reconstructed concentration at each pixel is a weighted average of the concentration of the surrounding pixels because of the smoothness regularization. Each row of the resolution matrix can be regarded as smoothing weights. The 2-D display of the row of the averaging matrix gives a clear dependence of the beam configuration, while the diagonal elements may not provide much information in this case.

Please refer to section 2.5, 3.5 in the supplement for the detailed description.

Specific comments

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

The reference has been added (Radon, 1986; Herman, 2009; Censor, 1983).

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

That is correct. For the traditional pixel-based methods, the pixel is required to be passed by at least one beam, otherwise the problem will be ill-posed. The regularization technique is the method to solve the ill-posed problems by adding a priori information into the original problem. This is shown in this study by using high-resolution grid division, in which case many grids are not passed by the laser beams. But we still got a good reconstruction through smoothness regularization.

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

"pixels" has been used.

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.

It is correct that there were 20 retroreflectors installed (including the ones at the corners). We can see from beam configuration (Fig. 1) that each retroreflector reflects two beams coming from different directions. Therefore, the total beam number will be $20 \times 2 = 40$. After removing the overlapped beams along the diagonals, the total beam number is 38. We have revised the texts to give a more detailed description of the setup. Please see section 2.1 in the supplement.

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.

For each laser beam, the path-integrated concentration (PIC) is measured by the analyser. The predicated PIC for one beam equals to the summation of the multiplication of the pixel concentration and the beam length passing it. In generality, let us assume that the site is divided into $N_c = m \times n$ pixels, which are arranged as a vector according to the left-to-right and top-to-bottom sequence and indexed by j . The average concentration for j -th pixel is c_j . The total number of laser beams is N_b , which are indexed by i . The length of the laser beam passing the j -th pixel is L_{ij} . Then for the i -th beam, the measured PIC b_i is contributed by all the pixels that it passes. A system of linear equations can be set up for all the beams: $\mathbf{b} = \mathbf{L}\mathbf{c}$, where \mathbf{L} is the kernel matrix that incorporates the specific beam geometry with the pixel dimensions, \mathbf{c} is the unknown concentration vector of the pixels, \mathbf{b} is a vector of the measured PIC data. This section has been rewritten to give a detailed explanation of the experimental setup. Please refer to section 2.1 in the supplement.

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?

Yes. A weight needs to be assigned to each equation depending on the uncertainty of the observation. Assuming the analyzers have the same performance, the uncertainty is mainly related to the path length. Therefore, equations are assigned weights inversely proportional to path length to make sure different paths have equal influences. In this study, the lengths of the laser paths are approximately equal to each other. Thus, their

weights were all set to be the same value which is scaled to be 1. The weights for the third-derivative prior equations were assigned to be the same value of w because they were all based on the same grid length. Please see section 2.2 in the supplement for the revised description.

Line 139: With equation (3) you require a solution with "small" L_1 norm, whereas, usually one requires a small $L_1(c - c_a)$ norm, where c_a is some prior estimate of c . Please explain the rationale behind your choice of $c_a = 0$.

Using $(c - c_a)$ means that we have some predefined values c_a which we use to regulate the solutions to make it close to the prior estimations. This is the case in the atmospheric sounding problem, where some modeled values are used as the prior estimations. But for the reconstruction of the 2-D concentration distribution of arbitrary area site. There is no such prior information about the underlying concentrations. The only information we know is that the values are non-negative.

Line 148: If the regularization parameter μ 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 μ ? Which is the solution of the LTD algorithm? Please specify the equation.

We have revised the equation. Now the grid length is part of the third-derivative prior equation. The regularization parameter is independent of the grid size and determines the balance between data fidelity and regularization terms. Determination of optimum regularization parameter is an important step of the regularization method. However, the regularization parameter is problem and data dependent. There is no general-purpose parameter-choice algorithm which will always produce a good parameter. For simplicity, we use the method based on discrepancy principle. The regularization parameter is chosen from a finite section of a monotonic sequence. For each value of μ , an optimal solution is derived by solving the inverse problem. Then the discrepancy can be calculated. The regularization parameter is determined to be the highest value that makes the discrepancy equal to $N_b \sigma^2$, where σ is the standard deviation of the noise. In this study, the reconstructions varied only slowly with the regularization parameters. Therefore, precise selection of the parameter was not necessary. The regularization parameter was selected from four widely varying values. The one produced the smallest discrepancy was used. Please see section 2.2 in the supplement for the equations and revised description.

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.

The non-negative least square (NNLS) optimization algorithm is the first widely used optimization method proposed by Lawson and Hanson (1995) in their textbook to solve the non-negative constrained least square problem. In this paper, the "NNLS algorithm" to the tomographic reconstruction refers to solve the original system of linear equations using the NNLS optimization algorithm without adding other a priori information. The LTD approach generates a new system of linear equations with non-negative constraints, which can also be solved by the NNLS optimization algorithm. Please see the revised description section 2.1 in the supplement.

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.

We have revised the formulations in the manuscript to give a more accurate definitions of the quantities and distinguish the continuous and discrete functions. The seminorm is defined as an integral of a continuous function, which is the total squared curvature in the MC algorithm. Then it is discretized to be the sum of the squared curvature at each pixel. Also, c is defined as a 1-D vector of the unknown pixel concentrations, while C is defined as a 2-D matrix of the pixel concentrations when they are arranged in the 2-D form. $f(x,y)$

is used to describe the underlying concentration distribution. Please see section 2.1, 2.2, and 2.3 in the supplement for the detailed definitions.

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.
The seminorm is defined as the total squared curvature in the MC algorithm. It is first discretized to be the sum of the squared curvature at each pixel. Then in order to minimize the seminorm, its partial derivative with respect to the concentration at each pixel needs to be zero, resulting in a difference equation. This equation is appended at each pixel as a smoothness prior equation and a new combined system of linear equation is set up. Please see the revised section 2.3 in the supplement for the equations and the detailed description.

Line 169: the same comment I made for μ (line 148) here applies to ω . Which constant for the inverse proportionality did you use?
The weight w is determined in the same way as determining the regularization parameter of μ , which was described in above answer for question of line 148. Please also refer to section 2.2 in the supplement for the detailed description.

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 σ , it contains σ_x and σ_y ...
The ppm unit has been changed to mg/m³. σ_x and σ_y have been defined.

Lines 212-213: this sentence is not clear to me.
This sentence has been revised as "They did not measure the peak location error and used a different way to calculate the exposure error by limiting the calculation domain in a small area near the peak instead of the whole map. Therefore, the results of the GT-MG algorithm are listed as a reference and only the measure of nearness is compared. "

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.
There is a mistake in the sentence. It has been revised as "The smaller the nearness value, the better the reconstruction quality". The trend of nearness with the increase of the source number is discussed in that paragraph.

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.
When there are two or more peaks on the map whose peak magnitudes are comparable to each other, the reconstruction algorithm may not identify the correct location of the highest peak. Therefore, a large error may happen when the highest value on the reconstructed map is located on the wrong peak.

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 than the error with which they are retrieved. Why the "source number" counts so much?
The reason is given in the previous answer. The large error is mainly because the reconstructed highest peak is the wrong one. When there are more peaks, this kind of error is more likely to happen.

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.
Actually, the NNLS algorithm is used as one representing method to solve the unregularized inverse problem. The main performance difference is due to the unregularized and regularized reconstructions. The unregularized inversion uses coarse grid

division and produces fewer pixel concentrations representing constant values in the grids. Therefore, this method has difficulty dealing with the large variation in a scale smaller than the grid size. It fits the true distribution better when the distribution becomes more uniform. We have updated the manuscript to include this explanation.

Technical corrections

Line 50: summarising ??

This sentence has been revised as "The path integral is approximated by the summation of the product of the pixel value and the length of the path in that pixel."

Line 85: necessary ?

The word has been replaced by "need".

Line 136: what is the superscript "21" ?

The number has been replaced with a reference.

Line 174: "well-posed"

Corrected.

Line 212: it (?)

Corrected.

Line 224: maybe "complexity" ?

Corrected.

Line 228: are slightly better ...

Corrected.

Line 241: derivation ??

It is the standard deviation of the results from 100 tests.

References

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, in: Series on Atmospheric, Oceanic and Planetary Physics, Vol. 2, edited by: Taylor, F. W., World Scientific, 2000.

von Clarmann, T., De Clercq, C., Ridolfi, M., Hoepfner, M., and Lambert, J.-C.: The horizontal resolution of MIPAS, Atmos. Meas. Tech., 2, 47–54, <https://doi.org/10.5194/amt-2-47-2009>, 2009.

Dobler, J. T., Zaccheo, T. S., Pernini, T. G., Blume, N., Broquet, G., Vogel, F., Ramonet, M., Braun, M., Staufer, J., Ciais, P.: Demonstration of spatial greenhouse gas mapping using laser absorption spectrometers on local scales. *Journal of applied remote sensing*. 11(1), 014002, doi: 10.1117/1.JRS.11.014002, 2017.

Wu, C.F., Chang, S.Y.: Comparisons of radial plume mapping algorithms for locating gaseous emission sources. *Atmospheric Environment*, 45, 1476-1482. DOI: 10.1016/j.atmosenv.2010.12.016, 2011.

Please also note the supplement to this comment:

<https://amt.copernicus.org/preprints/amt-2021-122/amt-2021-122-AC4-supplement.pdf>