Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2020-403-RC2, 2020

Atmospheric Measurement Techniques Discussions

Open Access

Interactive comment on "The Adaptable 4A Inversion (5AI): Description and first XCO_2 retrievals from OCO-2 observations" *by* Matthieu Dogniaux et al.

Anonymous Referee #2

Received and published: 29 December 2020

The paper of Dogniaux et al. reports on the development and testing of a carbon dioxide retrieval algorithm (5AI) for spectroscopic solar backscatter measurements such as those conducted by the OCO-2 satellite. The algorithm is deployed and evaluated for an evaluation data set from the OCO-2 mission. The paper is well written and the analyses are sound. But the serious drawback of the study is that the proposed algorithm does not account for particle scattering i.e. the algorithm is not (yet) what is typically called a full-physics algorithm. And, the algorithm is also not built as a computationally efficient approximation, for which lower accuracy would be acceptable.

The neglect of particle scattering essentially reduces the retrieval problem to a trans-

Printer-friendly version

Discussion paper

mittance calculation which will induce substantial errors even if the particle load in the atmosphere is low (e.g. AOD <0.5, termed "clear-sky" by the study). I actually wonder why a sophisticated radiative transfer model such as (V)LIDORT is required at all to perform such calculations. Isn't it just Beer-Lambert's law? Maybe, molecular Rayleigh scattering is included?

For example, the differences in spectral residuals between 5AI and ACOS (illustrated in Fig. 10) most likely stem to a large part from the differences in particle treatment, which is corroborated by the finding that 5AI and ACOS retrievals become more similar if the spectral differences are synthetically added to the 5AI processing (section 4.4). Essentially, the ACOS particle treatment is "added" to the 5AI calculations.

Given that the remote sensing community has been working on the simultaneous retrievals of greenhouse gases and particle properties for many years, the study lags behind current developments. On the other hand, the paper appears to be one of the first presenting a new algorithm to be applied to the problem and, it is always scientifically interesting and important to compare to new approaches.

If to be published, the authors should clearly discuss the drawback of neglecting particle scattering and they should underline that this assumption is most likely the leading error term. In particular, they should mention it in prominent places such as abstract and conclusions and they should add a section where they describe the (lack of) scattering treatment. Currently, the manuscript reads like (V)LIDORT (with standard scattering treatment) is used and only quite late it becomes clear that particle scattering is neglected.

I understand that full-physics retrievals are tough to get to work, but the authors might want to consider to support their evaluation by some refined assessments wrt. particle scattering:

- For the analysis in section 4, it might be insightful to examine correlation plots of the 5AI-ACOS (and/or 5AI-TCCON) differences with scattering and radiative transfer parameters (e.g. aerosol optical thickness retrieved by ACOS, surface albedo).

- The authors could add a forward modelling sensitivity study where they assume a particle scattering scenario (e.g. the one suggested by the ACOS retrievals) for which they actually perform full-physics forward calculations (without the derivatives) and then compare the spectral residuals to clear-sky calculations. Such an assessment could serve as a means to single out the effect of neglecting particle scattering in a clean way.

- ACOS (and other algorithms) previously reported on systematic challenges. For example, there was a bias in earlier ACOS retrievals towards the Southern higher-latitudes which was attributed to stratospheric aerosols (O'Dell et al., 2018). There was a land-ocean bias in early RemoTeC/GOSAT retrievals (Basu et al., 2013). For a future refined assessment, it would be interesting to evaluate 5AI in terms of these specific findings. [O'Dell, C. W, https://doi.org/10.5194/amt-11-6539-2018, 2018, Basu et al., https://doi.org/10.5194/acp-13-8695-2013, 2013.]

Other comments

General: The manuscript repeatedly emphasizes that the 5AI retrieval is a "Bayesian" concept. Is there more to it than just using the Bayesian/optimal estimation formalism (as most regularization concepts do)? A truly Bayesian retrieval would require a careful setup of the prior covariance (to represent the true covariance). Table 1 suggests that the prior covariance is rather a reasonable ad-hoc choice than the true atmospheric covariance. For example, 4 hPa surface pressure uncertainty is certainly a tight constraint compared to the actual atmospheric variability. Choosing prior covariances ad-hoc is common practice and thus, I do not criticize the procedure per se, but I do recommend being humble when highlighting the Bayesian and "optimal" nature of retrievals under such assumptions.

Introduction: The introductory discussion of other algorithms includes too much of detail. Many of the details (e.g. spectroscopic data, covariance setups, radiative transfer speed-ups) evolve over time without being published. Some key conceptual aspects are missing e.g. the fact that RemoTeC has chosen to not retrieve surface pressure because the problem becomes too ill-posed with respect to the microphysical particle parameters. I would recommend summarizing some of the algorithm features to avoid the paper being already outdated when it gets published. Some important early conceptual work by Oshchepkov et al., 2008, could be added. [Oshchepkov et al., https://doi.org/10.1029/2008JD010061, 2008]

Equation (1): "F" should be bold face, since it is a vector. The comment applies to many occurrences of "F" that follow.

L180: "Epsilon" should be bold face, since it is a vector.

L181: "Considering the probability density function instead of vectors" What does the statement mean? Wouldn't it be more appropriate to highlight the key idea of Bayesian statistics? For example: "Assuming that the true state is a particular draw from a Gaussian statistics and that it can be found through a trade-off to a measurement constraint weighted by the respective uncertainties ..." (or better wording).

Equation (3): Notation "(x_a)" is a bit misleading. It refers to the fact that the derivatives are evaluated at x_a (for the first iteration, if x_a is also the initial guess) while the following parentheses "(x-x_a)" refers to a multiplication. I would recommend using a vertical line with subscript x_a to indicate the linearization point of the derivatives.

L198: The Levenberg-Marquardt parameter "gamma" is required to vanish at the final iteration since otherwise it will induce a (probably undesired) regularization of the problem. In my experience, not all readers are aware of this effect and thus, I would recommend mentioning it clearly.

L199: It would be interesting to quantify the advantage in computational cost when not updating the derivatives. Probably, not updating the derivatives means that convergence is somewhat slower and thus, more iterations are required. So, there is a trade-off between number of iterations and cost per iteration. Could you give rough numbers?

Equation (6): A downside of not updating the Jacobians is that the averaging kernels are calculated with respect to the prior state (not the iterated retrieved state). The prior state is typically “further away” from the true state than the retrieval and thus, it is in the non-linear regime. I think this is a minor issue in general, but it might cause confusion when, for example, evaluating the effects of non-linearity on retrievals. One might consider calculating the Jacobians for the first and the last iteration to get rid of this effect.

L263: For a retrieval method paper, comparisons to raw retrievals of other algorithms are the essential tool. Comparisons to bias-corrected retrievals will not inform on methodological issues, but, to a large extent, such comparisons will mirror the effects of the bias corrections, in particular since some algorithms require substantial bias corrections. So, the (anyway sparse) comparisons to FOCAL could be removed from the manuscript.

Fig. 10. I recommend plotting the residuals and transmittances in the same units (either absolute radiance or relative transmittance) to allow for comparison of the amplitude of the residuals wrt. the spectra.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2020-403, 2020.