

The authors describe an airborne remote sensing instrument for UV/vis/near IR scattered sunlight measurements and the method for retrieving trace gas concentrations for the flight altitude, namely DOAS and the scaling method. Further the uncertainties are described to some extent, and some first application together with a comparison with auxiliary in situ measurements are presented.

I see two main problems with the manuscript and the scaling method: a) the uncertainty on the retrieved trace gas concentrations associated with the assumptions on the profile shape in the calculation of the scaling factor appears to be the dominant uncertainty and it is poorly quantified (partially hidden away in the supplementary and mentioned in section 4.2) and b) information on the radiative transfer in the atmosphere can only be obtained from layers where the scaling gas is actually present (This is hinted at in section 4.1). The authors claim that the scaling method is superior to optimal estimation. However, the advantage of optimal estimation is that it provides a formal framework for assessing the information content of a retrieved quantity. I believe the paper would be much stronger if these issues would be properly discussed upfront. Hence I recommend publication only after major revisions in sections 3.6, 3.7, 4, and 6.

General comments:

P. 13-14: I'm not sure why the authors need such a lengthy mathematical description. In principle it is an intercept theorem:

$$\frac{[X]_j^{theo} SCD_P^{theo}}{[P]_j^{theo} SCD_X^{theo}} = \frac{[X]_j SCD_P}{[P]_j SCD_X}$$

Intuitively, this is easier to understand.

P. 16, l. 28 – P. 17, l. 8: I would suggest using the uncertainty of the model/climatological profiles and do a proper error propagation to estimate the uncertainty. The NO₂ mixing ratio is up to four times larger than the standard case at one point in the probed flight. Is that realistic?

P. 18, l. 6-9: This shouldn't be just discussed as differences between different model parametrisations. Basically what it says it that you need good prior knowledge in order to retrieve a meaningful concentration with the scaling method.

P. 3, l. 33-34: The scaling method can hardly be described as novel method when the authors themselves already quote 6 reference for it.

P.1, l. 4-5: Since the paper is about the scaling method, maybe add a bit more explanation in the abstract what is actually done.

P. 21, l. 15-21: The a priori for optimal estimation is also a trace gas profile. Aerosol and cloud profiles are auxiliary information. The a priori for the scaling method is the profile ratio and aerosols and clouds are still auxiliary information. You have shown that those can be neglected. However, this is just how RT works for airborne measurements and hence this statement is valid for both techniques.

The authors poorly describe the need for the scaling method in the introduction. They list 3 main problems with airborne DOAS measurements on page 3. However, the first two problems are somewhat

convoluted and boil down to the remote sensing technique being ill-constrained. The third problem is about the residual trace gas amount in the background measurement. This seems somewhat inflated in its description (also in section 3.2 where the applied technique is then described). The authors then don't make use of their own arguments to describe the scaling method for providing a different a priori information in comparison to what has been done before.

The authors are not very precise in their referencing (see examples below). Please check carefully throughout the manuscript.

Specific comments:

P.1, l. 11-13: not a complete sentence

P.2, l. 1-2: 'It's remoteness initiated...'. I feel the authors are overstating here. Another reason for aircraft measurements would be to study the full extent of the ozone hole and not only locally. Please clarify or add reference for original statement.

P. 2, l. 8: 'aircraft-borne' sounds clunky.

P. 2, l. 11: How do you obtain information of 'photochemistry of pollutants' from column measurements? Please elaborate.

P. 2, l. 13-15: '... monitor the ground for sources and sinks...': Of the cited references only General et al. (2014) describe an anti-correlation between BrO and NO₂. This can hardly be referred to as monitoring the ground for sinks.

P. 2, l. 16-18: Please clarify this statement. Kritten et al. (2014) describes studies on the photochemistry of NO_x and Kreytz et al (2013) on BrO_x. Weidner et al. (2005) describe merely an instrument and its performance. Kritten et al. (2010) present also mainly the technique and some diurnal variation of NO₂. I quick search didn't show anything on trends. Do you maybe mean diurnal trends?

P. 2, l. 20-21: Baidar et al. (2013) describe measurements from a twin otter aircraft not HIAPER.

P. 2, l. 26: Meaningful instead of tractable? You can always do an inversion, also when you point in the wrong direction.

P. 2, l. 26: 'fed by' poor choice of wording

P. 2, l. 28-29: The stabilising system doesn't give the attitude data, but the attitude system provides this information and relays it to the stabilising system.

P. 2, l. 30-31: Again, you can always do an inversion, but it might not yield meaningful results. Also, absorption is not observed, but spectra. You also want to assign a trace gas concentration to a location in the atmosphere and not the absorption.

P. 2, l. 34: Celestial refers to what exactly?

P. 2, l. 33 – p. 3, l. 1: What you are describing constrains foremost the radiative transfer simulations. The inversion is constrained by an a priori. Please elaborate.

P. 3, l. 2: What gases if not O₄? Please explain.

P. 3, l. 4: 'absorption strength' is rather vague.

P. 3, l. 6: Please add 'for airborne applications' after 'constraining the radiative transfer'

P. 3, l. 8-11. Following the previous sentence, you make it appear as if this is the O₄'s fault. But it should be described as ill-constrained by auxiliary parameters. I guess this is your 'second problem'. Please clarify.

P. 3, l. 11-13: Maybe I misunderstand something here. But if there are no aerosols, you cannot assign this lack to the wrong profile layer. Please elaborate.

P. 3, l. 18: 'Wrongly called Fraunhofer reference spectrum'!! I keep on coming across this phrase in publications. It is an 'in-term' which excludes newcomers to the field of DOAS by confusing them with wrong terminology. Please don't use this or elaborate that it has been used historically, but is not precise.

P. 3, l. 18-19. The Lambert-Beer law calls for a background spectrum.

P. 3, l. 21-22, l. 24, l. 25, l. 29, P. 8, l. 16: 'Fraunhofer spectrum' s.a.

P. 3, l. 26-27: Volkamer et al. use zenith spectra.

P. 3, l. 33-34: None of the theses can be accessed without at least a link.

P. 3, l. 34: What sets your current study apart from Stutz et al and Werner et al.? This information is given on P. 4, l. 14 only. Maybe rearrange.

P. 4, l. 1: s.a.: you're not measuring absorption.

P. 4, l. 2: 'convenient' poor choice of word.

P. 4, l. 10-11: Please add references for models.

P. 4, l. 11-13. 'convenient' poor choice of word. Sarcastically, I could state that it is indeed convenient that you validate what you put in as a priori information.

P. 4, l. 28: 'The' instead of 'Its'.

P. 5, l. 12: Spectrometers including detectors?

P. 5, l. 12: Maybe mention here already that the 6 different spectrometers are for 2 telescope geometries and 3 wavelength ranges.

P. 5, l. 31 and p. 6, l. 13: 'onto the lid' implies they are outside that container. Do you mean 'on the underside'?

P. 6, l. 1: I think this is the first time the detectors are mentioned.

P. 6, l. 10: 'subsequently' and 'prior to the flight' is not sufficiently explained.

P. 6, l. 24: 'subset of parameters' has a slightly negative connotation. I ask myself here, what you possibly might have neglected to characterize and why.

P. 6, l. 26: 'fields of view'

P. 7, l. 4, l. 16, and l. 17: play is more commonly referred to as backlash of a gear.

P. 7, l. 21: 'arguably' poor choice of word.

P. 7, l. 33: is below 0.2 deg acceptable?

P. 8, l. 4: The term dSCDs is mainly used for scattered sunlight DOAS and not for active DOAS techniques. So your statement is not fully correct.

P. 8, l. 8-9: And what do you do for the IR?

P. 10, l. 5: 'can be determined'?

P. 10, l. 15: 'their' refers to what?

P. 10, l. 22: RT was used in the previous section.

P. 10, l. 25: And what is the non-standard run?

P. 10, l. 26, p. 11, l. 3: Please rephrase 'fed'.

P. 10, l. 27: EA?

P. 10, l. 30: 'celestial'? 'et cetera'?

P. 11, l. 6: Reference for FAIRO?

P. 12, l. 22: 'in the lower troposphere'?

P. 13, l. 7: Reference to Table 2 doesn't help here if you don't state which are the target gases and which the scaling gases.

P. 13, l. 8-9: 'potential'?

P. 13, eq. 6: why don't you define a B_i here then? This is unnecessarily complicated. See above.

P. 14, l. 12: 'are obtained from Eq. (1)': I think you should mention that they are obtained from a DOAS fit and then Eq. (1).

P. 14, l. 15-24. This is very complicated. You need to read the text, the caption and the figure at the same time to get all the information to understand what is going on. Curtin is a confusing term, so are the random factors. I would need to get a calculator out to assign a specific colour to a number. Why did the authors not choose a non-fractional scaling factor for the units? Where does the uncertainty in α_R come from at this point and how is the uncertainty in the SCD ratio calculated? Please state here where the uncertainties will be described in the manuscript.

P. 14, l. 26: Please define 'compact relationship'

P. 14, l. 26-27: This basically only confirms that the remote sensing aircraft measurements are mainly sensitive to the concentrations at the flight altitude which was shown before (e.g., Baidar et al., 2013; Bruns et al., 2006). The citations I assume refer to the McArtim code? Again, I cannot access these without a link.

P. 15, l. 19-21: Also p and T will have horizontal and vertical gradients and those will depend on the aircraft altitude. The same is valid for using ozone as scaling gas. Please discuss.

P. 15, l. 22: Why do the authors not use proper error propagation here?

P. 15, l. 24-26: Please explain why you are using percentage values for the errors. These numbers won't be applicable for small SCDs.

P. 15, l. 29 – P. 16, l. 10: The supplementary material only shows results for UV and not visible as stated in the manuscript. I have problems understanding Figure 4 in the supplements considering the limited information provided: what are the frequency distributions? Is the altitude the aircraft altitude?

P. 16, l. 8: Then why do you use formaldehyde as a representative case here?

P. 16, l. 10-17: In the supplementary Figure 5a), is the in situ data filtered or smoothed?

P. 17, l. 16: 'validate' is a strong word in the context provided by this section.

P. 17, l. 20: 'agree reasonable well': That is impossible to assess from the figure.

P. 17, l. 20-22: Please don't use absorption here. You are referring to the different absorber concentration profiles. These two sentences are rather misleading as they are right now.

P. 18, l. 17: 'curtains' s.a.

P. 18, l. 27: This is quite an abrupt transition to BrO!

P. 18, l. 29: Isn't Figure 9 during flight section C?

P. 19, l. 9: That could still mean that both models are wrong.

P. 19, l. 20-21: Where do these detection limits come from?

P. 19, l. 21: Maybe replace 'degrade' by 'decrease' or 'is lost'.

P. 20, l. 7: Remove 'eventually'

P. 20, l. 9-10: Why do you introduce FT, PBL, and LMS abbreviations here?

P. 20, l. 10-11: How are they compatible? The previously mentioned studies Fitzenberger et al., and Prados-Roman et al. are both from the Arctic and they do show elevated levels of BrO in the FT. Please clarify your statement.

P. 20, l. 20: Maybe not name GLORIA at this point now otherwise I would ask for more explanations.

P. 20, l. 24-25: 'skylight radiances'? Is the instrument radiometrically calibrated?

P. 21, l. 9-11: See comment above, you basically get as output, what you provide as input. However, both models could still be wrong.

P. 21, l. 14: Where do 3.5 km and 15 km come from all of a sudden?

P. 21, l. 14: 'It can be argued': Please rephrase. 'major'?

Table 2: Formatting seems a bit messed up. Maybe add horizontal lines or larger gaps between the lines for the different trace gases.

Table 4: Is not referred to in the text.

Figure 2: Please add the concentrations to the figure or caption. 'Tob' in caption.

Supplementary Figure 7: and captions in the supplementary don't explain the red line in panel b.

Kritten et al. (2010) has wrong title. It's the AMTD title which was then renamed for the AMT version.