Hüneke et al., The novel HALO mini-DOAS instrument: Inferring trace gas concentrations from air-borne UV/visible limb spectroscopy under all skies using the scaling method

Answers to Reviewer 1

The reviewer's comments are in black, answers are in red.

The manuscript describes the mini-DOAS instrument that has already been operated in a number of science campaigns onboard the German High Altitude and Long-range research aircraft (HALO) together with the so-called "scaling method" to infer trace gas mixing ratios from the mini-DOAS limb measurements. The paper is generally well written and provides an important reference for the mini-DOAS instrument and in particular its data analysis. I recommend publication in AMT after consideration of the following – mostly minor – comments.

Specific comments:

As the instrument has already been flown on a number of science campaigns and this paper gives only a few selected examples, I suggest to include a table listing deployments of the mini-DOAS instrument on HALO campaigns so far.

A new table listing all HALO campaigns in which the mini-DOAS instrument participated is included in the revised manuscript.

p1.l3: please spell out HALO as "German High Altitude and Long-range research aircraft (HALO)" when used for the first time and move the URL reference from the abstract to the introduction section.

The text is rephrased accordingly.

p2.l1: sentence ends early

The text is rephrased accordingly.

p2.l6: better either give an earlier reference for DOAS as well, or explicitely indicate that Platt and Stutz (2008) is a recent review of DOAS and not the original reference

The text is rephrased accordingly.

p2.l22: "German GV" -> "German High Altitude and Long Range Research Aircraft (HALO), that is based on a Gulfstream G550 jet"

The text is rephrased accordingly.

p2.l35: what exactly is meant by "celestial" here?

‚Celestial' is a standard expression in astronomy and it refers to the position of astronomical objects, here the sun and the earth.

p3.l29: "to" -> "to be"

The text is rephrased accordingly.

p4.l11: what exactly does "latter" here refer to? Only CCMs or CTMs and CCMs?

It refers to the incorporated model predictions. The text is rephrased accordingly.

p5.l18: insert "elevation angles"

The text is rephrased accordingly.

p5.l20: why "skylight" and not simply "light"0? I know its picky, but in general there may be other light sources than just skylight.

The text is rephrased accordingly.

p6.l9: Can you explain briefly why surface temperatures are important here and not the much cooler upper tropospheric / lower stratospheric temperatures at cruise altitude?

In this context the relevant time frame is the duration of the pre-flight preparations of the aircraft, while it is located on the apron. Access to the instrument is not possible and at large ambient T's some of the ice-water already melts affecting the instrument in-flight temperature on longer lasting sorties. The text is rephrased to make this clear.

p7.l10: any reference to BAHAMAS ?

No, unfortunately there is no proper reference for BAHAMAS available.

p9.l1: "Polstracc" -> "POLSTRACC"

The text is rephrased accordingly.

P9.l14: "Acridicon" -> "ACRIDICON"

The text is rephrased accordingly.

p10.l27: What is "EA"? Elevation angle?

EA is the elevation angle, as mentioned in section 2.4.

p11.l6: "Fairo" -> "FAIRO"

The text is rephrased accordingly.

p11.l7: "for measuring" -> "for in-situ measuring"

The text is rephrased accordingly.

p11.l16: not sure if "all 13 flights" is still relevant here, as many more flights on more recent campaign have been performed with FAIRO

The statement refers to the fact that the FAIRO instrument was first deployed on HALO during the ESMVal campaign.

p12.l17: horizontal resolution is not well defined for a Lagrangian model. Please give more detail what this refers to.

The horizontal resolution refers to the average distance between two neighboring air parcels in one model layer (see Vogel et al. (2015) and references therein and McKenna, Daniel S., et al. "A new Chemical Lagrangian Model of the Stratosphere (CLaMS) 1. Formulation of advection and mixing." Journal of Geophysical Research: Atmospheres 107.D16 (2002)).

p12/13: EMAC: Maybe include a few sentences about EMAC's chemistry scheme – in contrast or comparison to what has been mentioned for CLaMS.

In contrast to ClaMS, EMAC contains a very detailed tropospheric chemistry scheme. For the present EMAC model run, the submodel MECCA (Module Efficiently Calculating the Chemistry of the Atmosphere) is used to simulate the chemical kinetics, with the photochemical data taken from the JPL compilation (Sander et al., 2011) including recent updates (Jöckel et al., 2016). The details are presented by Jöckel et al. (2016), Section 3.5 on pages 1160/1161.

The text is added to the manuscript.

p14.l12: Suggestion: Say again explicitly that SCD_X and SCD_P are measured by the mini-DOAS, the alpha are from a model and [P] is measured in-situ. (If that is what you are doing.)

The text is rephrased accordingly.

p20.l11: better be a bit more specific of where no elevated BrO concentrations are observed. I assume this statement does not refer to the Antarctic boundary layer?

This refers to the Antarctic free troposphere. The text is rephrased accordingly.

Caption Fig. 2: Suggestion: Include in the caption date, location and campaign of sample measurement.

The text is rephrased accordingly.

Hüneke et al., The novel HALO mini-DOAS instrument: Inferring trace gas concentrations from air-borne UV/visible limb spectroscopy under all skies using the scaling method

Answers to Reviewer 2

The reviewer's comments are in black, answers are in red.

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.

Our reactions to the reviewer's major concerns & comments:

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.

The reviewer is correct in pointing out the uncertainty resulting from the assumed profile shapes. However, uncertainties of profile shapes taken from chemistry transport or chemistry climate models are hard to quantify. Here it should again be pointed out that when applying the scaling method only relative profile shapes matter but not absolute concentrations. Hence, in order to quantify the sensitivity of the scaling method on the actual profiles, the following sensitivity studies on the scaling factor were performed: (1) including a small scale variability derived from the O3 in-situ measurements, (2) including the uncertainty due to the vertical sampling error of the profile factor (see section 3.7.1), and (3) shifting the profile vertically in an arguably extreme manner to account for systematic errors in the adopted vertical dynamics (i.e. the adopted cooling/heating rate) in the models (see supplement, Figure 7). Furthermore, Stutz et al. (2017) carried out a range of sensitivity studies modifying trace gas profile shapes by altering concentrations overhead and below the aircraft (see their Table 5 and supplement). In fact, they arrived at uncertainties of the scaling factor similar to those obtained in our study.

Nevertheless, not all imaginable variations of the profile shapes can be covered in such sensitivity studies and in particular strong trace gas concentration gradients such as at the polar vortex edge pose problems for the method, as mentioned in the text (section 3.7.2). Besides such specific situations, one can reasonably assume that the inferred trace gas concentrations are conservative with respect to the model predicted profiles, i.e. possible deviations from the model predictions are underestimated. Mainly this conclusion is drawn from the observation that e.g. if the actual trace gas concentration was lower than the model prediction at sampling altitude (with all other altitudes the same), then the alpha factor for that trace gas would probably be overestimated, resulting in an overestimated trace gas concentration (compare Eq. 11).

b) Information on the radiative transfer in the atmosphere can only be obtained from layers where the scaling gas is actually present.

The statement of the reviewer deserves some consideration, as in particular mentioned in section 4.1. In the best case, absorption by the scaling gas and the targeted gas occurs in similar altitude layers. It is shown in Figure 6 of the manuscript that by using an inappropriate scaling gas (in this case O4) the change in radiative properties (in this case the occurrence of a cloud layer at 4-8 km) can have a very strong influence on the retrieved NO2 concentrations above 8 km altitude. Comparably, when using O3 as a scaling gas, the influence of the low level clouds on retrieved NO2 concentrations is greatly diminished.

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 the retrieved quantity.

The statement of the reviewer is at first glance reasonable, however only when assuming the forward model (to describe the observation) used in the optimal estimation reasonably describes the physical reality. While the RT can reasonably be approximated for limb measurements under clear (i.e. mostly Rayleigh scattering) skies, for cloudy skies any forward model lacks information on the spatial distribution, micro-physical and optical properties of the relevant scatterers, and their temporal variation over a single set of measurements. Noteworthy for ascent or descent of an aircraft or for a set of limb measurements at different elevation angles, a single set of measurements to obtain a profile may take tenths of minutes for which the RT has to be precisely known when applying optimal estimation.
Therefore, (many) assumptions need to made in the forward model describing the RT under clear skies, which propagate in an unquantifiable manner into the formal solution. In consequence, the information content formally returned by the optimal estimation algorithm is largely flawed, and not yet really quantified.
Further, as shown in previous studies (Stutz et al., 2017, Werner et al, 2017) the relevant RT is at least 2D (clear skies) and under cloudy skies 3D over domains of hundreds of kilometers in the limb direction for measurements in the upper troposphere and lower stratosphere.
In order to overcome these problems - applying optimal estimation for cloudy sky studies - the scaling method has been developed and intensive sensitivity studies have been performed for a wide range of conditions (Knecht, 2015, Stutz et al., 2017, Werner et al, 2017). Moreover, unlike stated by the reviewer the scaling method also offers information content of the measurements, here expressed in alpha factors. They describe the fraction of total measured absorption due to line of sight absorption (i.e. in field of view of the telescope).

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.

In consideration of the general and specific comments of the reviewer, the draft has been modified and explanations added.

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: […] Intuitively, this is easier to understand.

The reviewer is correct in proposing that the approach is similar to an intercept theorem, however only if the profile shapes of the involved trace gases were the same. The suggestion misses the point that by calculating alpha factors the fraction of the absorption collected from the atmosphere which is located within the instrument's line of sight is calculated. Hence, the sensitivity of the measurement to detect the target gas at flight altitude can be estimated, in some analogy with the averaging kernels obtained in the optimal estimation.

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?

For a discussion of the uncertainty of the modelled profile shapes, see our answer to the reviewer's major comments and concerns above. Evidently large relative errors result when the target gas has small concentrations at flight level as compared to all atmospheric layers from which the absorption of the targeted gas is collected. Exactly this situation occurs for the measurements of NO2 within the polar vortex, and a change in profile shape at the vortex edge can therefore lead to strong deviations.

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.

The comparison shows that the prior knowledge (in this case the relative trace gas profile shapes) does have an impact on the retrieved concentrations, as is the case with any inversion method which uses prior knowledge. The point here is to show that while dynamical and chemical differences between models may render them more or less suitable for our task, the differences in the outcome are not unreasonable large to a degree that they heavily impact the result i.e., beyond the stated errors.

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.

The scaling method is novel in the sense that it has only been used in a few studies so far (Stutz et al., 2017), but not for the interpretation of air-borne limb studies in the heavily aerosol and cloud-loaded troposphere.

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

The manuscript describes a novel air-borne DOAS instrument as well as the scaling method and its characterisation. Considering that an abstract should remain concise and short, the sentence „Here we report on the relevant instrumental details and the novel scaling method used to infer the mixing ratios of UV/vis absorbing trace gases from their absorption measured in limb geometry. The uncertainties of the scaling method are assessed in more detail than before for sample measurements of NO2 and BrO." may reflect well the contents of the manuscript.

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 phrasing concerning „a priori" versus „parameters constraining radiative transfer" is corrected in the text. However, we strongly disagree with the reviewer's comment that 'aerosols and clouds can be neglected' in the forward modelling of spectroscopic

measurements when applying 'both techniques'. The reverse is true since any inversion tries to minimize the square distances of measured vs modelled parameter(s) (called cost function), and hence the model needs to reflect somehow the physical reality in order for any inversion to make sense (e.g. Rodgers, 2000, Tarantola, Inverse Problem Theory and Methods for Model Parameter Estimation, Society for Industrial and Applied Mathematics (SIAM), 2005). Evidently since our technique relies on the analysis of light, the forward (RT) model needs to account for all processes affecting the propagation of light, i.e. the presence of aerosol and clouds in the terrestrial atmosphere.

Since often remote sensing measurements in the UV/vis spectral range suffer from a lack of information in the spatial distribution of the aerosols and clouds and their optical properties, other means to constrain the inversion have been developed in the past, for example, by using the measured relative intensities (e.g. Prado-Roman et al., 2011), or the slant column densities of O4 (e.g. Bruns et al., 2006), or any other suitable proxy.

In particular, it was recently shown (Stutz et al. (2017) and others) that O4 is an inadequate constraint on the radiative transfer when inverting for high-altitude limb measurements using optimal estimation. Evidently the same is true when using O4 as a constraint for the RT applying the scaling method (see our Figure 6). If there were other measured parameters available (being less sensitive to aerosol and clouds than O4 or relative radiances) to constrain the radiative transfer in optimal estimation, we would probably agree with the reviewer's comment. However, we are not aware of any suitable atmospheric parameters of that kind.

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.

In the introduction two major problems with air-borne DOAS measurements are addressed, i.e. (1) how to determine the absorption in the background (Fraunhofer) spectrum and (2) how to deal with largely unconstrained radiative transfer when aerosol and clouds are present. In order to bring more structure into the text we switched the paragraphs on page 2, line 30, and page 3, line 17, and rephrased them.

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

Thank you for noting. We accordingly changed the citations at P. 2, l. 16-18.

Specific comments:

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

We changed the text accordingly.

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.

We changed the text accordingly.

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

All occurrences of ‚aircraft-borne' are replaced with ‚air-borne'.

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

Column measurements are often used to infer photochemical parameters, such as the life time of NO2 (e.g. Beirle et al., Megacity Emissions and Lifetimes of Nitrogen Oxides Probed from Space, Science, Vol. 333, Issue 6050, pp. 1737-1739 DOI: 10.1126/science.1207824, 2011.) Here, we added the example of HONO measurements by the CARIBIC instrument (Heue et al., CARIBIC DOAS observations of nitrous acid and formaldehyde in a large convective cloud, ACP, https://www.atmos-chem-phys.net/14/6621/2014/, DOI: 10.5194/acp-14-6621-2014, 2014).

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

However, keeping the notation 'sinks' is justified here, given that there are many studies on sinks of trace gases inferred from remote sensing measurements. Here we just add one example: Beirle et al., Megacity Emissions and Lifetimes of Nitrogen Oxides Probed from Space, Science, Vol. 333, Issue 6050, pp. 1737-1739 DOI: 10.1126/science.1207824, 2011.

P. 2, l. 16-18: Please clarify this statement. Kritten et al. (2014) describes studies on the photochemistry of NOx and Kreycy et al (2013) on BrOx. 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?

Not really. Kreycy et al (2013) deals with the coupling of NO2 and BrO. The sentence is accordingly rephrased for a more accurate citation. Since the context of trends would need further elaboration, this part is dropped from the sentence.

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

We changed the text accordingly.

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

We changed the text accordingly.

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

We changed the text accordingly.

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.

We changed the text accordingly.

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.

We changed the text accordingly.

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

The word celestial is a standard expression in astronomy and it refers to the position of astronomical objects, here the sun and the earth.

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.

The constrains/boundary conditions and assumptions are described foremost because they largely determine the actual RT. We changed the text accordingly.

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

In particular, here those gases are meant which have a sizeable effect on the radiative transfer in the present context (i.e. are optically thick, or tau >1) .

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

We changed the text accordingly by using the notation ‚slant column density'.

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

The suggested phrase is added at the beginning of the sentence.

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.

The text merely describes the sensitivities when using O4 as a parameter to constrain radiative transfer, and points out the insufficiency of using it as RT proxy for limb measurement from high-flying, fast-moving aircraft platforms, which is exactly what the reviewer suggests. The paragraph is rephrased to make clear what is meant by the 'first' and 'second' problem.

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.

If there were no aerosols and clouds in the atmosphere, then they would not need to be considered and the optimal estimation would work out just fine. However, if the scattering properties of the atmosphere changed during the measurements, the inferred O4 slant column density in limb would change too with changing cloud cover underneath even when flying at constant altitude. In consequence, an apparent aerosol layer located at higher altitudes would be retrieved.

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.

The sentence is rephrased accordingly and the notation ‚background spectrum' is used throughout the text.

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

We changed the text accordingly by rephrasing the sentence.

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

The sentence is rephrased accordingly.

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

The sentence is rephrased accordingly.

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

Appropriate links are added in the bibliography.

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.

We changed the text in the abstract accordingly.

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

We changed the text accordingly.

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

We changed the text accordingly.

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

A reference to section 3.5 is added, where the models are described.

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.

We accordingly changed the text to make more clear what is meant. In fact, in the past we collected many examples where the CTM predictions and our measurements as well as the complementary measurements do not agree at all (e.g. on NO/NO2 ratio, or the presence of tropospheric BrO, IO, and CH2O ….), hence any incorrect ‘a priori’ information drawn from the model is not carried over to the final result. Another good example on the power of the scaling method and its comparably weak reliance on any ‘a priori’ information can be found in the publication of Ye et al., (2016) (Rapid cycling of reactive nitrogen in the marine boundary layer, Nature, 532, 489–491, (28 April 2016).

P. 4, l. 28: ‘The’ instead of ‘Its’.

We changed the text accordingly.

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

Modern optical spectrometers usually contain detectors.

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

The first sentence of section 2.1 is rephrased accordingly.

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

In the first instance, it is the lower side, in the second, it is the top of the container. We changed the text accordingly.

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

The detectors are now explicitly mentioned in the beginning of section 2.2.

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

We changed the text accordingly.

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.

The notation “subset” refers to those parameters employed in the DOAS retrieval. We changed the text accordingly.

P. 6, l. 26: ‘fields of view’

We changed the text accordingly.

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

The sentences have been rephrased accordingly in the text.

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

The sentence has been rephrased to reflect that the given numbers are typical examples.

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

Certainly yes, since due to (multiple) scattering the RT in UV/vis spectral range already degrades the effective field of view to larger angles. This can for example be learned when inspecting the width of averaging kernels for similarly stabilized instruments (Baidar at al., 2013), or from the intensive RT exercises we and others performed in the recent past (cf. Raecke, 2013, http://www.iup.uni-heidelberg.de/institut/forschung/groups/atmosphere/stratosphere/publications/pdf/MScThesis_Rasmus_Raecke.pdf, Knecht, 2015 http://www.iup.uni-heidelberg.de/institut/forschung/groups/atmosphere/stratosphere/publications/pdf/MA_thesis_Matthias_final.pdf)

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.

The sentence is rephrased accordingly.

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

Here are two examples for studies carried out using nearIR measurements: 1. Wolf et al., Potential of remote sensing of cirrus optical thickness by airborne spectral radiance measurements at different sideward viewing angles, Atmos. Chem. Phys., 17, 4283-4303, doi:10.5194/acp-17-4283-2017, 2017. And 2. PhD thesis of Scalone, Retrieval of Cirrus Optical Properties in the near-IR spectral range within the NASA ATTREX Project, PhD Thesis, University of Heidelberg, Germany, 2017 (doi:10.11588/heidok.00023004, http://www.ub.uni-heidelberg.de/archiv/23004).

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

We changed the text accordingly.

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

The word 'their' refers to the measured spectra of the flights where optimal estimation cannot be applied. A clarification is added to the text.

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

We changed the text accordingly.

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

The sentence is rephrased and for all simulations the settings are mentioned in the text (from line 25 on).

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

We changed the text accordingly.

P. 10, l. 27: EA?

EA is the elevation angle, as mentioned in section 2.4.

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

The word celestial is standard in astronomy and hence used for the position of the sun and earth. We changed the text accordingly.

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

The reference for FAIRO has been moved to the second sentence of the paragraph.

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

We changed the text accordingly.

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.

We changed the text accordingly (now Table 3).

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

We removed potential from the text.

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

See answer to general comments 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).

We changed the text accordingly.

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. Curtain 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 alpha_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.

In present atmospheric science the notion curtain is a widely used notion for a (CTM) simulated (atmospheric) parameter as a function of time/horizontal coordinate and altitude. There is nothing random about it. However, we include a definition of the notion 'curtain' upon its first appearance in the text (page 11, line 4). Nice examples of curtains of potential vorticity and trace gases can also be found in recent publications of Jurkat et al., (2017) Figure 1 (Depletion of chlorine and nitrogen oxide reservoir species in the 2 lower Antarctic polar vortex measured from aircraft, Geophys. Res. Lett., 44, doi:10.1002/2017GL073270, 2017), Werner et al., (2017) (Figure 2). The fractional numbers just appear in order to project all gases onto a common colour scale. In fact, the application of scaling method does not even require consideration of the individual multiplier because for a specific gas a common multiplier cancels out (see eq. 4 and 5). The overall uncertainty is calculated via equation (13). SCD uncertainties are discussed on page 15, line 23 and uncertainties of the alpha factors are discussed on page 15, from line 27 on. A sentence has been added to point the reader to the respective section.

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

Figure 4 itself provides a good impression of 'a fairly compact relationship' (cited from the text) as compared to data which do not show any correlation.

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.

Sorry. We added the link to the bibliography. In fact, it was recently shown by Stutz et al., (2017) (Figure 9). However, for a newly deployed instrument it appears prudent to show that it is indeed behaving the way it is supposed to. The sentence has been rephrased to reflect this.

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.

The basic idea of the method presumes that the in-situ measured concentration of the scaling gas (be it O3 concentration or O4 calculated from p and T) is representative of the average concentration of that gas in the atmospheric volume sampled by the instrument's telescopes. The alpha factor ratios are calculated for this volume. Hence, small scale variabilities in the concentration of the scaling gas can lead to an error, which is incorporated in the error estimate for the alpha factor ratio (see Table 5). Larger horizontal gradients in trace gas concentrations, e.g. when flying through the polar vortex edge, need to be carefully considered when interpreting the derived trace gas concentrations. We accordingly added a remark in section 3.7.2. The error due to the vertical gradient of the scaling gas is considered (1) by looking at the vertical sampling uncertainty of the telescopes (section 3.7.1, paragraph (c)) and (2) estimated by shifting trace gas profiles vertically to correct for model errors in vertical ascent/descent rates.

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

As written and indicated in text, a Gaussian error propagation is applied.

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.

Uncertainties in percentages are given in order to make them comparable to the other factors influencing the overall error (compare Table 5). "Small" SCDs do not occur, since the reference-SCDs of limb background spectra measured in the troposphere and lowermost stratosphere always include at least the column of stratospheric O3 and NO2, e.g, NO2 reference-SCDs are $5*10^{15}$ … $1*10^{16}$ molec./cm^2, O3 reference-SCDs are approx. $1*10^{19}$ molec/cm^2, and O4 reference-SCDs are approx. $3*10^{43}$ molec^2/cm^5.

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?

The RT simulations were carried out in the UV (343 nm) as well as in the visible (477 nm). For details see Knecht (2015). The caption of Figure 4 in the supplement is rephrased to make clear what is meant.

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

The calculations were performed in support of the interpretation of the measurement taken within the framework of the ACRIDICON campaign (Wendisch et al., 2016, The ACRIDICON-CHUVA campaign to study tropical deep convective clouds and precipitation using the new German research aircraft HALO, Bull. Am. Meteorol. Soc., 97, 10, 1885-1908, doi:10.1175/BAMS-D-14-00255, 2016.) which took place in the Amazon basin in fall 2014. There, strong convection lead to cumulus and cumulonimbus clouds which are largely varied in vertical and horizontal extent, providing an extreme test case to determine the sensitivity of the method as a function of cloud cover .

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

The in-situ data are smoothed in an interval ranging from 100 s before to 100 s after each spectroscopic measurement. A detailed description of the smoothing is added to the text of the supplement.

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

We changed the text accordingly.

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

Clarification is added to the text (differences are below 35 ppt).

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.

We exchanged the word ‘absorption’ with the word ‘concentration’ and rephrased the sentences.

P. 18, l. 17: ‘curtains’ s.a.

See above.

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

The beginning of the subsection is rephrased accordingly.

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

We changed the text accordingly.

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

For a discussion of model uncertainties see comment above.

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

The detection limits are estimated based on the uncertainties at very low mixing ratios. These are indicated in Figure 9, panels c and e. For [NO2] = 10 ppt and [BrO] =2 ppt respectively, the inferred mixing ratios are at least two times larger than the uncertainty.

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

We changed the text accordingly.

P. 20, l. 7: Remove ‘eventually’

We changed the text accordingly.

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

We changed the text accordingly.

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.

For the present DOAS retrieval we estimate our detection limit for BrO is 2 ppt (see comment above) and do not derive elevated BrO in the FT. BrO mixing ratios larger than 2 ppt were not shown either by Fitzenberger et al. or Prados-Roman et al. in the FT above 3.5 km (the minimum altitude of the measurements reported here).

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

The sentence is rephrased accordingly in the text.

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

The instrument itself is not radiometrically calibrated, as mentioned in section 3.3. However, it can be cross-calibrated by comparison with a radiometrically calibrated instrument (see Wolf et al., 2017, their Figure 8).

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.

For the model uncertainties see the discussion above. Here it is necessary to recall that the result of the scaling method is not determined by the absolute concentrations (as predicted by the CTMs) but only by the relative profile shapes of the involved gases.

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

The phrase is dropped because it is not essential here. For altitudes covered by the present flight see Figure 3.

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

We changed the text accordingly.

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

Horizontal lines are added for clarification (now Table 3).

Table 4: Is not referred to in the text.

Table 4 (now Table 5) is referred to in the beginning of section 3.7.1.

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

The slant column densities of the targeted trace gases are added and the text corrected for spelling.

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

The reference to the red line is added to the caption.

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

The error in the bibliography is corrected.

The novel HALO mini-DOAS instrument: Inferring trace gas concentrations from air-borne UV/visible limb spectroscopy under all skies using the scaling method

Tilman Hüneke[1], Oliver-Alex Aderhold[1], Jannik Bounin[1], Marcel Dorf[1,2], Eric Gentry[1,3], Katja Grossmann[1,4], Jens-Uwe Grooß[5], Peter Hoor[6], Patrick Jöckel[7], Mareike Kenntner[1,7], Marvin Knapp[1], Matthias Knecht[1], Dominique Lörks[1], Sabrina Ludmann[1,8], Sigrun Matthes[7], Rasmus Raecke[1], Marcel Reichert[1], Jannis Weimar[1,9], Bodo Werner[1], Andreas Zahn[10], Helmut Ziereis[7], and Klaus Pfeilsticker[1]

[1]Institut für Umweltphysik, University of Heidelberg, Heidelberg, Germany
[2]Max-Planck-Institute for Chemistry, Mainz, Germany
[3]now with Department of Astronomy and Astrophysics, University of California Santa Cruz, Santa Cruz, California, USA
[4]now with Joint Institute For Regional Earth System Science and Engineering (JIFRESSE), University of California Los Angeles, Los Angeles, California, USA
[5]Forschungszentrum Jülich, Institute of Energy and Climate Research - Stratosphere (IEK-7), Jülich, Germany
[6]Institut für Physik der Atmosphäre, University of Mainz, Mainz, Germany
[7]Deutsches Zentrum für Luft- und Raumfahrt, Institut für die Physik der Atmosphäre, Oberpfaffenhofen, Germany
[8]now with Ifeu - Institut für Energie- und Umweltforschung Heidelberg GmbH, Heidelberg, Germany
[9]now with Physikalisches Institut, University of Heidelberg, Heidelberg, Germany
[10]Karlsruhe Institute of Technology (KIT), Institute for Meteorology and Climate Research, Karlsruhe, Germany

Correspondence to: Tilman Hüneke
(Tilman.Hueneke@iup.uni-heidelberg.de)

Abstract. We report on a novel 6 channel optical spectrometer (further on called mini-DOAS instrument) for ~~aircraft-borne~~ air-borne nadir and limb measurements of atmospheric trace gases, liquid and solid water, and spectral radiances in the UV/vis and nearIR spectral ranges. The spectrometer was developed for measurements from aboard the ~~HALO () research aircraft~~ German High Altitude and Long-range research aircraft (HALO) during dedicated research missions. Here we report on the relevant instrumental details and the novel scaling method used to infer the mixing ratios of UV/vis absorbing trace gases from their absorption measured in limb geometry. The uncertainties of the scaling method are assessed ~~for~~ in more detail than before for sample measurements of NO_2 and BrO~~measurements~~. Some first results are reported along with complementary measurements and comparisons with model predictions for a selected HALO research flight from Cape Town to Antarctica, which was performed during the research mission ESMVal on 13 September 2012.

1 Introduction

In the past three decades ~~aircraft-borne~~ air-borne UV/vis spectroscopy measurements developed into a powerful tool to study the photochemistry and radiative properties of the atmosphere. ~~Based~~ The approach is based on the pioneering work of Noxon (1975) and later Noxon et al. (1979) to exploit ground-based spectroscopic observations of the zenith scattered skylight to mon-

itor stratospheric NO_2 (and later O_3, BrO and OClO, see below)~~.~~. The technique was further improved after the discovery of the ozone hole in ~~1985 and the need to unravel its formation mechanism. Its remoteness initiated~~ 1985. Within the framework of ozone hole research, zenith sky UV/vis measurements ~~to be~~ were not only performed from the ground (e.g., Solomon et al., 1987a) but also from research aircrafts. ~~Accordingly, optical~~ Optical spectrometers were deployed on the NASA DC-8 during Airborne Arctic Stratospheric Expedition (AASE) in 1989 (e.g., Wahner et al., 1990a; Schiller et al., 1990; Wahner et al., 1990b) and later (1992 -1995) on the German Transall (e.g., Brandtjen et al., 1994; Pfeilsticker and Platt, 1994). The spectroscopic analysis of the measured skylight spectra for the detection of O_3, NO_2, BrO, and OClO was based on Differential Optical Absorption Spectroscopy (DOAS) ~~(Platt and Stutz, 2008)~~(for a recent overview see Platt and Stutz, 2008), and assisting radiative transfer (RT) calculations allowed to estimate the integrated overhead (or total) column density of the targeted gases (Solomon et al., 1987b).

McElroy et al. (1999) were the first to exploit ~~aircraft-borne~~ air-borne nadir scattered skylight measurements to study plumes of BrO in the lower troposphere during arctic spring. Later ~~aircraft-borne~~ air-borne multi-axis DOAS measurements by Bruns et al. (2004, 2006) over Europe and on major air traffic corridors by Dix et al. (2009) within the CARIBIC project (http://www.caribic-atmospheric.com/) were used to gain information on the distribution and photochemistry of pollutants and their products within the troposphere (e.g., Heue et al., 2014).

Meanwhile, more versatile DOAS-based 2-D imaging nadir techniques have become available to monitor the ground for sources and sinks of UV/visible/nearIR absorbing radicals, pollutants and their products and green-house gases ~~(e.g., Heue et al., 2008; Geri~~ (e.g., Heue et al., 2008; Gerilowski et al., 2011; Beirle et al., 2011; Merlaud et al., 2012; General et al., 2014).

Air-borne UV/vis measurements in limb geometry started with the balloon-borne study of Weidner et al. (2005) which aimed at studies of the photochemistry, budgets and ~~trends~~ variation of the NO_x (Kritten et al., 2010, 2014) and BrO_x (Kreycy et al., 2013) families in the stratosphere~~(e.g., Weidner et al., 2005; Kritten et al., 2010; Kreycy et al., 2013; Kritten et al., 2014)~~. The air-borne limb measurements of scattered skylight continued with the aircraft studies of Prados-Roman et al. (2011) made from aboard the DLR Falcon, and more recently from the American High-performance Instrumented Airborne Platform for Environmental Research (~~GV~~ HIAPER) aircraft ~~(Baidar et al., 2013; Volkamer et al., 2015)~~(Volkamer et al., 2015), the NSF/NCAR C-130 (Gratz et al., 2015; Ye et al., 2016), and the NASA Global Hawk (Stutz et al., 2017; Werner et al., 2017), and those reported here from ~~the German GV aircraft HALO (for first results see Wendisch et al., 2016; Voigt et al., 2016).~~HALO, an aircraft based on a Gulfstream G550 jet (http://www.halo.dlr.de/). For first results from measurement campaigns involving the HALO mini-DOAS instrument, the reader is referred to e.g. Wendisch et al. (2016); Voigt et al. (2016); Wolf et al. (2017); Jurkat et al. (2017). Table 1 lists all deployments of the HALO mini-DOAS instrument in recent years.

One common facet of all these air-borne UV/vis limb measurements is the need for a stable observation geometry (or pointing) of the telescopes (required are a few tenth of a degree), in order to render the underlying mathematical inversion problem for trace gas retrievals ~~tractable~~ meaningful (Rodgers, 2000). Therefore, all modern air-borne UV/vis spectrometers ~~are fed by skylight collected~~ collect skylight from actively controlled telescopes to compensate for the movements (i.e., the roll and pitch angle) of the air-borne measurement platform. Most conveniently the attitude data to control the telescope's pointing

are provided by the aircraft's inertial navigation system (INS) or by custom-built ~~stabilising~~ attitude systems (e.g., Baidar et al., 2013).

~~Potentially the largest problem in the interpretation of UV/vis limb measurements involves the inherent inversion problem when assigning the observed absorption (or inferred slant column density)~~ Air-borne DOAS limb measurements however come with two major difficulties:

The first results from the necessity to know the amount of absorption of the targeted species in the background spectrum. In skylight DOAS all measurements are referred to a background spectrum recorded with the same instrument, since the measured atmospheric absorptions are much smaller (optical densities of atmospheric absorbers typically range between 10^{-4} to ~~the~~ 10^{-2}) than those due to the Fraunhofer lines of the sun's photosphere. Different strategies are available to determine the absorption in the background spectrum, depending on the available observation geometries and target gas. Most easy to deal with are gases with little or negligible amounts located overhead the aircraft (e.g., CH_2O, $C_2H_2O_2$, $HONO$, often IO, $OClO$ at daytime, et cetera ...) because their absorption in the background spectrum is then small or even negligible. It is far more complicated to determine the amount of absorption in the background spectrum of gases with considerable (and often spatially and temporally varying) amounts located overhead the aircraft (e.g., O_3, O_4, NO_2, BrO, ...). Here, direct sun or zenith sky observations are helpful (e.g., Volkamer et al., 2015; Stutz et al., 2017), but for fast moving aircrafts the overhead column density may change too rapidly to carry out Langley-type regressions of the measured absorption as a function of air mass (cf. Gurlit et al., 2005; Dorf et al., 2008). Therefore, the amount of absorption in the background spectrum and its contribution to the total absorption needs to be minimized (e.g., by referring all measurements to low solar zenith angle observations at high altitude) such that the remaining absorption eventually can be calculated using model predictions (see below).

The second major difficulty comes from the necessity to constrain the RT in the atmosphere. The latter is needed for the interpretation of UV/vis limb measurements when assigning concentrations to the different locations in the atmosphere, which is often carried out using inversion techniques such as optimal estimation (Rodgers, 2000). Unfortunately, in a heavily aerosol loaded or even cloudy atmosphere, light paths (or light path distributions) are not well-defined due to multiple scattering of collected skylight. Therefore, the inversion problem becomes almost intractable when the ~~radiative transfer~~ RT forward model is not constrained by other means than the aircraft and telescope attitude, celestial geometry of the sun and earth, and atmospheric pressure and temperature data. Additional data on the micro-physical properties and spatial distribution of aerosols and cloud particles are required~~to constrain the inversion~~. These are often taken from in situ aerosol measurements, lidar or radar observations, model predictions of the spatial distribution of the measured gases, and observations or predictions of the cloud cover, et cetera. ~~In addition~~Because these parameters are usually not known sufficiently well for applications with fast-moving aircraft platforms, the employed retrieval strategies often rely on constraining the ~~radiative transfer by the absorption strength~~ RT by the slant column density of simultaneously measured absorption bands of the collisional complex $O_2{-}O_2$ (in the following briefly called O_4) and/or relative radiances (e.g., Bruns et al., 2006; Prados-Roman et al., 2011; Baidar et al., 2013).

~~Constraining the radiative transfer~~ For air-borne applications, constraining the RT by O_4 ~~however~~ comes with some limitations. First of all, the absorption of O_4 is $\propto [O_2]^2$, thus skylight is much more efficiently absorbed in the lower parts of the

troposphere than in the upper troposphere or stratosphere. A (a priori unknown) fraction of the UV/vis light collected in limb geometry measurements in the middle and upper troposphere or lower stratosphere may be back-scattered from lower parts of the atmosphere (Oikarinen, 2002). A changing ground albedo or cloud cover at low levels may thus modulate ~~(and mimic)~~ the measured limb ~~absorption strength~~ slant column density of O_4 higher up in the atmosphere. Therefore, ~~the~~ changing scattering properties of the troposphere - even of those parts which are not being directly sampled by the telescope's field of view (FOV) - may mimic the presence (or lack) of aerosols and cloud particles at flight altitude (Stutz et al., 2017). If a significant fraction of the targeted gas is located off the telescope's ~~field of view~~FOV, assigning proper amounts of the measured gas to the correct locations in the atmosphere thus becomes ambiguous, or even impossible. In consequence, until the recent past, the retrievals of UV/vis limb measurements had been restricted to clear or almost clear sky observations.

~~A third problem of air-borne DOAS measurements in the UV/vis spectral ranges addresses the need for knowing the amount of absorption of the targeted species in the background spectrum, often called Fraunhofer reference spectrum. In skylight DOAS referring all measurements to a background spectrum appears to be necessary since the measured atmospheric absorptions are much smaller (optical densities of atmospheric absorbers typically range between to) than those due to the Fraunhofer lines of the sun's photosphere. Different strategies are available to determine the absorption in the Fraunhofer spectrum, depending on the available observation geometries and target gas. Most easy to deal with are gases with little or negligible amounts located overhead the aircraft (e.g., , , , often , at daytime, et cetera ...) because their absorption in the Fraunhofer spectrum is then small or even negligible. It is far more complicated to determine the amount of absorption in the Fraunhofer spectrum of gases with considerable (and often spatially and temporally varying) amounts located overhead the aircraft (e.g., , , , , ...). Here, direct sun observations are helpful (e.g., Volkamer et al., 2015; Stutz et al., 2017), but for fast moving aircrafts the overhead column density may change too rapidly to carry out Langley-type regressions of the measured absorption as a function of air mass (cf. Gurlit et al., 2005; Dorf et al., 2008). Therefore, the amount of absorption in the Fraunhofer spectrum and its contribution to the total absorption needs to be minimized (e.g., by referring all measurements to low solar zenith angle observations at high altitude) such that the remaining absorption eventually can be calculated using model predictions (see below).~~

In order to render the interpretation of air-borne UV/vis limb measurements more tractable for all kind of skies, in particular for measurements in partly cloudy skies, we recently developed the so-called scaling method (Raecke, 2013; Großmann, 2014; Werner, 2015; Hüneke, 2016; Stutz et al., 2017; Werner et al., 2017). The scaling method makes use of the concentration of a scaling gas, either in situ measured (e.g. O_3) or calculated (e.g. O_4), which is used together with the ~~simultaneously limb measured absorptions~~ slant column densities from simultaneously measured spectra of the scaling gas (further on denoted P) and the targeted gases (further on denoted X), preferentially monitored in the same wavelength region. The latter appears to be ~~convenient~~ advantageous in order to ~~eliminate~~ minimise any wavelength dependence of the atmospheric Rayleigh and Mie scattering (see Stutz et al. (2017) and ~~their~~ its supplement, and below). The in situ measured concentration and the remotely observed absorption of the scaling gas P can then be used to infer an effective light path length (or distribution) common for the gases P and X (see section 3 below). The underlying assumption is a horizontally constant trace gas concentration along the line of sight equal to the in situ measured concentration. One draw-back of the scaling method comes from its (moderate) sensitivity towards the relative vertical profiles shapes (but not absolute concentrations) of the involved trace gases.

The sensitivity can best be dealt with by using a scaling gas P with a similar profile shape to that of the target gas X. The relative profile shapes of both gases can then be taken from either in situ measurements performed during dives of the aerial vehicle, any a ~~prior~~ priori knowledge, and/or from chemistry transport models (CTMs, e.g. ~~,~~ CLaMS, SLIMCAT, ...) or chemistry climate models (CCMs, e.g. ~~, EMAC). The latter is very convenient~~ EMAC; for information on the models see section 3.5). Incorporating model predictions is very straightforward since the limb measurements are often used to validate the ~~predictions of the respective CTMs~~ framework of predictions together with the other complementary measurements performed on board the respective research aircrafts.

The present study describes the novel UV/vis/nearIR HALO mini-DOAS instrument and explores the scaling method in more detail together with its uncertainties and potential errors.

The paper is structured as follows. In section 2 the instrument is described and characterized. Details of the employed methods are provided in section 3. These include the spectral retrieval, ~~radiative transfer~~ RT calculation, complementary measurements, CTM and CCM modelling and a description of the scaling method and its uncertainties. Section 4 describes sensitivity studies of the retrieval method by comparing inferred $[NO_2]$ using different CTM and CCM trace gas profile predictions and different scaling gases. Finally, our results for inferred $[NO_2]$ and $[BrO]$ are inter-compared with complementary measurements and model predictions for a HALO flight from Cape Town to Antarctica during austral spring 2012 (section 5). Section 6 concludes the study.

2 Instrument Description

The novel mini-DOAS instrument builds on the heritage of similar instruments assembled by our research group and collaborating partners for deployments on aircraft (e.g., the DLR Falcon, Geophysica, NASA Global Hawk, NSF/NCAR C-130) and high flying balloons (LPMA/DOAS and MIPAS/TELIS/mini-DOAS payload) observations (Ferlemann et al., 2000; Weidner et al., 2005; Kritten et al., 2010; Prados-Roman et al., 2011; Kreycy et al., 2013; Gratz et al., 2015; Ye et al., 2016; Stutz et al., 2017; Werner et al., 2017).

~~Its~~ The major design criteria for air-borne measurements are a small weight (several to tens of kg), a small power consumption (200 W), multiple channels of moderate spectral resolution (i.e., ranging from several tenth of nm in UV to several nm in nearIR) for UV/vis/nearIR analysis of the skylight received from nadir and simultaneously in scanning limb direction, a stable optical imaging, and finally an easy to operate instrument, either by on board operators (e.g., on HALO) or fully automated for deployments on unmanned aircrafts, such as the NASA Global Hawk (Stutz et al., 2017). On HALO the mini-DOAS instrument is installed in the unpressurized so-called 'boiler room' located in the rear of the HALO aircraft, which is not accessible during the flight. While this position favors the aircraft's balance and weight distribution and provides more versatile options to assemble more maintenance-prone instruments within the cabin, it comes with the handicap of strongly changing ambient conditions to operate the instrument (i.e., boiler room temperatures may change from -30°C during polar missions to +50°C in tropical missions, and the ambient pressures may change between 1000 mbar at the ground and 150 mbar at cruise altitude), which are prohibitive for operating stable optical instrumentation. Therefore, we follow the proven concept of our air-borne

DOAS instrumentation, where the optical spectrometers are kept at vacuum pressures and temperature ~~stabilized~~ stabilised at 0°C by immersing the whole spectrometer container into a water-ice vessel (Weidner et al., 2005). The latter also comes with the advantages of minimising the time ($\propto$ 2h) to get the instrument flight-ready and larger auxiliary instrumentation (a cooler etc.) is not necessary in the field.

The mini-DOAS instrument consists of three major parts (Figure 1): (a) an aperture plate, from which three nadir and three limb scanning telescopes collect skylight and which is mounted into the aircraft fuselage, (b) a spectrometer unit, which houses six cooled and evacuated grating spectrometers, and (c) a control unit to automatically operate the instrument and support communication with the aircraft data network.

2.1 The aperture plate and telescope

The aperture plate accommodates ~~two telescopes each~~ six telescopes in total for measurements in the UV ~~, visible and nearIR spectral ranges, one of each for nadir and limb~~ (2x), visible (2x) and nearIR (2x) spectral ranges. One set of UV/vis/nearIR-telescopes is used for limb and the other for nadir observations. It is mounted into an existing aperture opening ($28 \times 20.5 \times 9$ cm^3) of the HALO aircraft fuselage and has a weight of about 4 kg. The three limb telescopes point to the starboard side of the aircraft, perpendicular to the aircraft fuselage axis, and are moveable to attain elevation angles (EA) from $+3°$ to $-93°$ relative to the horizon, in steps of less than $0.005°$. During the flight they are commanded to compensate for the changing roll angle of the aircraft (see below), while the three nadir telescopes are held fixed. The six telescopes have diameters of 1.2 cm each, and six silica fiber bundles conduct the collected ~~skylight~~ light from the telescopes to the spectrometers. At the spectrometer end, the fibers are linearly arranged and placed at the entrance slits of the spectrometers. At the telescope end, the fibers are linearly arranged as well positioned in the focal point of the telescope lenses, forming ~~field of views (FOVs)~~ fields of view of 3.15° in the horizontal 0.38° in the vertical for the UV and visible telescopes, and 1.68° in the horizontal and 0.76° in the vertical for the nearIR telescopes (for the other details see Table 2). Finally, an industrial miniature camera is attached to the telescope aperture plate and oriented towards the sky's limb for monitoring of the investigated sky area simultaneously with the spectroscopic measurements.

2.2 Spectrometer unit

The six grating spectrometers are assembled in a Czerny-Turner configuration with the specifications given in Table 2. Back-thinned silicon CCDs with 2048 channels are employed for detection in the UV and visible wavelength range, while the nearIR-spectrometers use InGaAs photodiode arrays with 512 channels. In order to clearly identify each spectrometer and the corresponding telescope, they are labeled by the wavelength range and numbered 1 through 6. Spectrometers UV1, VIS3, and NIR5 (odd numbers) are then used in nadir viewing geometry, and spectrometers UV2, VIS4 and NIR6 (even numbers) are used in limb viewing geometry. All spectrometers are mounted ~~onto the~~ on the lower side of the lid of a vacuum tight container. The spectrometer container lid also accommodates vacuum tight connectors and feed-throughs for the fiber bundles and the connection to the detector electronics. Prior to each mission the vacuum tight spectrometer container is evacuated to some 10^{-5} mbar (leakage rate 2×10^{-5} mbar $\cdot$ l/s) to keep the spectrometer and detectors clean from contamination and the optical imaging stable. The

whole vacuum tight spectrometer container is immersed into a vessel filled with 7 l of water/ice, in order to stabilise the spectrometer and detector temperatures at around 0 °C. The whole spectrometer unit is further insulated using a combination of silica vacuum insulation panels (thermal conductivity of $0.008\ \mathrm{W/(m \cdot K)}$) and a more flexible Polyvenylidenfluorid (PVDF) foam (thermal conductivity of $0.037\ \mathrm{W/(m \cdot K)}$). Prior to a flight, the water ice vessel is filled with ~~approx.~~ 4 kg of ice and 3 l of cooled water, providing a latent heat of melting of ~~-~~1300 kJ. Ambient conditions before the flight determine the amount of latent heat necessary to cool down the instrument and the heat flow into it on the ground. The instrument is cooled down usually between 1.5 and 2.5 hours before take-off due to pre-flight aircraft procedures. In consequence the holding time of the water ice against melting as well as the quality of stabilising the instrument temperature is somewhat variable, but typically ranges between 6 to 8 hours in a flight. When operating under arctic conditions, i.e. with an already cooled instrument prior to flight preparations, constant temperatures are maintained for 10 hours or more, showing that average heat flows during operation are well below 36 W. In a worst case scenario, i.e. in very hot and humid ambient conditions in the tropics (e.g., in Manaus/Brazil in fall 2014, or the Maldives in August 2015), the instrument has to be ~~subsequently cooled~~ cooled additionally by adding ice and removing liquid water directly prior to the flight. Under these conditions, the average heat flow during flight preparation and measurement flight is around 80 W, and therefore in the present configuration the instrument is limited to 3 – 4 hours of stable temperatures ($\Delta T \leq \pm$ 1 °C). Therefore, after having made some experience with the instrument's heat budget, three Peltier elements were additionally mounted on top of the spectrometer container lid.

2.3 Control unit

The power supply, the read-out electronics for the six spectrometers, the controllers for the telescope motion, the control board for the Peltier elements, house-keeping electronics as well as a single board personal computer for instrument control and data storage and communication with the operator in the aircraft cabin is integrated into two removable electronic boxes, mounted above the spectrometer unit (yellow boxes in Figure 1). The measurements and control processes including read-out of the aircraft attitude data and the motion control of the three limb telescopes is controlled by a LabView software running on the single board computer. Finally the whole instrument is mounted on a custom-built rack of $45 \times 47 \times 54\ \mathrm{cm}^3$. The total weight of the instrument is 57 kg, including the water/ice, and it consumes 100-200 W of 28 Volt DC power provided by the aircraft, depending on the power consumption of the Peltier elements.

2.4 Pre-flight test measurements

Prior to each mission, the instrument is optically and electronically characterized in the laboratory~~for a subset of parameters~~. This characterization includes recording of the dark currents and offset voltage of the CCD detectors, recording of line shapes and the optical dispersion, recording of trace gas absorption spectra, measurements of the telescope's ~~field of views~~fields of view, and alignment of telescopes to the major aircraft axis (roll angle).

Dark current and offset voltage: Dark current and offset voltages of the CCD detectors are recorded prior to each flight for post-flight data processing (Platt and Stutz, 2008).

Slit function: The spectrometer slit function and wavelength dispersion are monitored in the laboratory and in the field prior to each flight using HgNe and Kr emission lamps (see Table 2). Moreover, since test measurements in the laboratory show that the slit functions are sensitive to the spectrometer's temperature, their T-dependence is extensively studied and monitored in the laboratory. For example it is found that the width of the slit function is most sensitive at low temperatures with a sensitivity of 0.005 nm/K (0.04 channels/K). However, due to the thermal stability of the instrument, a temperature sensitive slit function does not need to be taken into account for most spectral retrievals.

The effective ~~field of view~~ FOV ($\mathrm{FOV_{eff}}$) of the telescopes is made up of three contributions, which are (a) the optical FOV of the telescope ($\mathrm{FOV_{opt}}$), (b) the lag time between aircraft movement and telescope attitude correction (Δ_{attit}) and (c) the ~~play~~ backlash of the telescope gear (Δ_{gear}). These are discussed in the following paragraphs.

$\mathrm{FOV_{opt}}$ (a): The optical FOV of the telescopes is measured in the laboratory in advance of the deployment to any mission. $\mathrm{FOV_{opt}}$ is listed in Table 2. The vertical $\mathrm{FOV_{opt}}$ in the UV/vis is $\approx 0.38°$.

Telescope attitude control (b): In order to maintain the targeted elevation angle (EA) of the telescopes relative to the horizon during flight, the changing roll angle of the moving aircraft has to be corrected for. The aircraft's attitude data is received from the aircraft sensor data system (BAsic HALO Measurement And Sensor system, or in brief BAHAMAS) aboard the HALO aircraft at a frequency of 10 Hz and a time delay < 1 ms via an Ethernet UDP broadcast. Due to the continuous movement of the aircraft and the time delay between data transmission and actual motor movement, a small difference between the targeted and the actual telescope angle can thus be expected. Tests involving a continuous and arbitrary sampling of the aircraft roll angle and the telescope position yields a mismatch of both angles with a standard deviation of $\Delta_{\mathrm{attit}} \approx 0.17°...0.18°$ (Fig. 1 in the supplement).

Telescope gear (c): In addition, the pointing precision is limited by the ~~play~~ backlash of the telescope's gear (Δ_{gear}). Telescope gear ~~play~~ backlash ($\Delta_{\mathrm{gear}} \approx 0.05°$) is determined by the shift of the recorded radiance maximum when the telescope's FOV is measured by scanning in opposite directions.

Gaussian summation of contributions (a) ... (c) gives a $\mathrm{FOV_{eff}}$ for e.g. the VIS4 telescope ranging between 0.54°(during mission ML-Cirrus (Voigt et al., 2016)) and 0.64°(during the TACTS/ESMVal ~~missions (e.g. Müller et al., 2016))for the VIS4 telescope, for which the tests are carried out. Arguably it is of the same size for the other limb telescopes.~~ mission (e.g. Müller et al., 2016)).

Telescope alignment to the major aircraft axis: After integration of the instrument into the aircraft, the telescope angle with respect to the aircraft is calibrated by placing a Ne gas lamp in 15 m distance and at the same height as the telescopes in the line of sight of the telescopes. The lamp is modified so that light is only emitted through a narrow ($\sim 5\ \mathrm{mm}$) slit. Scanning over the lamp again gives the ~~field of view of~~ FOV of the telescope, whose maximum is used to determine the angle that represents a horizontal line of sight with respect to the horizon. Under the assumption of a 2 cm uncertainty in the height of the lamp relative to the aperture plate (1 cm at each side), the angle uncertainty is 0.076°. When the aircraft is grounded, the aircraft roll angle given by the aircraft attitude data has a standard deviation of 0.2°. Accordingly, the systematic error in telescope alignment is $\Delta_{\mathrm{align}} < 0.3°$.

The systematic misalignment (Δ_{align}) can be tested independently by observation of the radiance 'knee', i.e. the apparent maximum in the relative radiances received from a set of ~~elevation angles~~ EAs in limb direction, which is wavelength dependent (see Figure 5 in Deutschmann et al. (2011) and Figure 5 in Weidner et al. (2005)). Figure 2 in the supplement shows measured and modelled relative radiances in the UV and visible wavelength ranges, indicating a systematic misalignment below 0.2°. This accuracy is sufficiently narrow due to the widening of the effective FOV due to light scattering (as indicated by RT calculations, see e.g. Raecke, 2013; Knecht, 2015) and it is comparable to other stabilised air-borne DOAS instruments (e.g. Baidar et al., 2013, Figure 3).

3 Methods

3.1 DOAS retrieval

The spectral retrieval is based on the DOAS method (Platt and Stutz, 2008) ~~. The~~ and it is applied to measurements in the UV and visible wavelength ranges. The evaluation of measurements in the near infrared is carried out using other approaches (Wolf et al., 2017; Scalone, 2017), which are not in the scope of this study. The primary product of the DOAS spectral retrieval in scattered sunlight applications are so-called differential slant column densities (dSCDs) given in molecules per cm^2 (Platt and Stutz, 2008), i.e., the amount of absorption measured in a foreground versus background ~~(Fraunhofer)~~ spectrum. Since the details of the spectral retrieval and its uncertainties have been described in previous studies (Harder et al., 1998; Aliwell et al., 2002; Weidner et al., 2005; Dorf et al., 2006; Butz et al., 2006; Kritten et al., 2010; Stutz et al., 2017), here only those details are discussed which depart from our previous work. Table 3 provides a brief summary of the different DOAS settings and typical dSCD errors. Table 4 lists the absorption cross sections used in the analysis together with their uncertainties as stated in the literature. In all spectral retrievals a polynomial of degree 2 is included to compensated for broad-band extinction features in the ~~radiative transfer~~ RT of the atmosphere, together with a ~~Fraunhofer reference~~background spectrum, a Ring spectrum and an additional Ring spectrum multiplied by λ^4 as suggested by Wagner et al. (2009). The trace gas cross section spectra are calculated by convolving the literature absorption cross sections listed in Table 4 with the measured dispersion and a Gaussian line-shape describing the Hg line at 404 nm (UV) or the Kr line at 450 nm (vis). Inaccuracies in wavelength calibration due to small changes in the instrument's optics and errors in the wavelength calibration of the fitted spectra are accounted for during the spectral retrieval. All trace gas cross sections are linked together and the package of trace gas cross sections is allowed to shift against the ~~Fraunhofer reference~~ background spectrum and the Ring spectra which are linked together. Typical spectral shifts for both groups of spectra are well below 1 detector pixel.

3.1.1 Spectral retrieval of O_3, BrO, OClO, CH_2O, and O_4 in the UV spectral range

Five different spectral windows are analyzed for the absorption of O_3, BrO, OClO, CH_2O, and O_4 in the UV wavelength region (Tables 3 and 4). All five intervals are different but show significant overlap (Table 3).

O_3 is retrieved in the 335 – 362 nm wavelength region of the Huggins band in order to achieve a larger spectral overlap with the other targeted gases in the UV spectral range which is found necessary in support of the scaling method (see section 3.6). Here O_3, BrO, NO_2, and O_4 references are included in the spectral retrieval (Table 3). The average error in the inferred O_3-dSCD is 6.4×10^{16} molec/cm^2 for the UV spectral range. It is noteworthy that the spectral retrieval for O_3 could be improved by using the stronger ozone absorptions bands of the Huggins band occurring towards the lower wavelength end of the UV spectrometer (310 nm), but then spectral overlap with the other gases as well as the much stronger absorption would negatively infer with the quality of the O_3-scaling method.

O_4 is retrieved in a spectral window ranging from 350 – 370 nm in order to allow fitting of the collisional band $^1\Sigma_g^+$ + $^1\Sigma_g^+(\nu{=}1)$ (at 360.5 nm) (Table 3).

The BrO analysis window covers 342 – 362 nm, the vibrational transitions ~~3, 4, 5, and 6~~3 ← 0, 4 ← 0, 5 ← 0, and 6 ← 0 of the $A^2\Pi_{3/2} \longleftarrow X^2\Pi_{3/2}$ electronic transition. Reference spectra of O_3 for 223 K and 293 K (the latter orthogonalised to the 223 K reference spectrum) are included in the spectral retrieval together with reference spectra of NO_2, CH_2O, and O_4 (for the other parameters see Table 3. Figure 2 (bottom left) shows an example for the retrieval of BrO from a limb spectrum collected in the lowermost arctic stratosphere during the ~~Polstracc~~ POLSTRACC mission (http://www.polstracc.kit.edu) on January 31, 2016. Here the BrO-dSCD equals $(5.8 \pm 0.3) \times 10^{14}$ molec/cm^2.

OClO is retrieved in the 353 – 392 nm spectral range, i.e. of the vibrational bend and stretch transitions of the $A^2A_2 \longleftarrow X^2B_1$ electronic transition. The spectral fit includes references spectra of O_3 at 223 K and 293 K (the latter orthogonalised to the 223 K reference spectrum) as well as reference spectra of NO_2 and O_4. Figure 2 (bottom right) shows an OClO retrieval from the ~~Polstracc~~ POLSTRACC flight on January 31, 2016. In this case the OClO-dSCD is $(5.7 \pm 0.2) \times 10^{14}$ molec/cm^2.

CH_2O is retrieved in a spectral window ranging from 323 to 357 nm, i.e. the rovibrational bands of the $\tilde{A}^1A_2 \longleftarrow \tilde{X}^1A_1$ electronic transitions. The spectral window is chosen in order to distinguish the signature from other trace gas absorptions in this wavelength range, particularly of O_3, BrO, and HONO. The spectral retrieval includes absorption cross sections of O_3 at 223 K and 293 K (the latter orthogonalised to the 223 K spectrum), spectra of NO_2 and CH_2O taken at 293 K (since the bulk of CH_2O is expected to be present in the lower troposphere) as well as of O_4, HONO, and BrO. Figure 2 (top left) shows a sample CH_2O retrieval of a limb spectrum recorded during the HALO research flight above the Amazonian rain forest on Sept. 16, 2014 performed within the framework of the ~~Acridicon~~ ACRIDICON mission (Wendisch et al., 2016). In this case, the CH_2O-dSCD amounts to $(1.28 \pm 0.05) \times 10^{17}$ molec/cm^2.

3.1.2 Spectral retrieval of O_3, O_4, NO_2, H_2O, IO, and $C_2H_2O_2$ in the visible spectral range

The main species measured in the visible spectral range are O_3, O_4, NO_2, and H_2O and if sufficiently present IO, and $C_2H_2O_2$. Here the focus is put on the spectral retrieval of O_3, O_4, and NO_2, since the former two gases are used for the scaling method and the later complements the measurements of NO and total NO_y by the AENEAS instrument (see section 3.4.2) on board HALO. The spectral retrieval of IO, $C_2H_2O_2$ and water vapor is not discussed further in this manuscript.

Ozone is analyzed in the 450 – 500 nm wavelength band of the Chappius absorption band. The center of both fitting window is thus shifted by 20 nm relative to NO_2. In the spectral retrieval, absorption cross sections of NO_2 at 223 K, together with

O_4 and water vapor (Table 3) are included. The average error in the inferred O_3-dSCD is 4×10^{17} molec/cm^2 in the visible spectral range.

The $^1\Sigma_g^+ + {}^1\Delta_g$ absorption of O_4 at 477.3 nm is analyzed in the 460 – 490 nm wavelength band with the same combination of reference spectra as those used in the O_3 retrieval (Table 3). For O_4 the average retrieval error is 5.6×10^{41} molec^2/cm^5.

NO_2 is thus analyzed in a relatively wide spectral window ranging from 424 – 490 nm of the sub-bands of the electronic transition $^2B_1 \longleftarrow {}^2A_1$ thus supporting both small dSCD errors while maintaining a stability of the least squares fit involved in the spectral retrieval. Reference spectra of O_3 at 223 K and 293 K (the latter orthogonalised to the 223 K spectrum), O_4 and water vapor are included in the retrieval (Table 3). Figure 2 (top right) shows an example of a spectral retrieval of NO_2 with a dSCD of $(2.17 \pm 0.05) \times 10^{16}$ molec/cm^2 for a limb spectrum taken within the framework of the ESMVal mission close to Antarctica on 13 September 2012. The simultaneous detection of O_3 and O_4 is also evident in this spectral retrieval.

3.2 Determination of the amount SCD_{ref}

In order to obtain the total slant column density (SCD), which is needed to solve the inversion problem, the amount of absorption SCD_{ref} contained in the ~~Fraunhofer reference~~ background spectrum needs to be determined and added to the measured dSCD, i.e.

$$SCD = dSCD + SCD_{ref}, \quad (1)$$

where SCD_{ref} ~~is~~ can be determined using (a) the so-called Langley method (i.e., a regression of dSCD as a function of total air mass), (b) inferred from a priori assumptions (for example for photo-labile species like OClO SCD_{ref} = 0 can be reasonably assumed for high sun), or (c) simulated if the light paths (i.e. the optical state of the atmosphere) and the concentration field of the species are well-known. As the mini-DOAS instrument is installed in the bottom of the aircraft fuselage, a direct sun light spectrum cannot be recorded, which prevents the use of method (a). Instead in most cases when methods (a) and (b) are not feasible, SCD_{ref} needs to be determined from the known RT and concentration field of the respective trace gas. For this purpose flight sections with clear sky conditions are selected and a non-linear retrieval constrained by measured relative radiances and/or O_4 optical densities is carried out in order to infer the aerosol extinction (e.g. Prados-Roman et al., 2011; Stutz et al., 2017). The inferred aerosol profile and the a priori trace gas concentration profiles known for example from chemical modelling (section 3.5) are then used in the RT modelling (see section 3.3) to simulate SCD_{ref}. In order to obtain SCD_{ref} for flights where non-linear aerosol profile retrievals are not suitable, ~~their spectra~~ the measured spectra of these flights are fitted against a reference spectrum for which SCD_{ref} has previously been determined. The dSCD-offset relative to the yet-undetermined reference spectrum is then used to calculate the missing SCD_{ref}. The uncertainty of SCD_{ref} is considerably decreased when the retrieval is referred to clear sky measurements. The largest dependencies of SCD_{ref} are the aerosol optical depth at flight altitude, the trace gas concentration at flight altitude, and the overhead column. Typical SCD_{ref} errors are of the order of 10 – 20%.

3.3 Radiative transfer modelling

The ~~radiative transfer~~ RT is simulated in 2D (and in selected cases in 3D, see supplement Figure 3) using version 3.5 of the Monte Carlo ~~radiative transfer (RT)~~ RT model McArtim (Deutschmann et al., 2011). The model's input is chosen according to the on board measured atmospheric temperatures and pressures, including climatological aerosol profiles from SAGE II (https://eosweb.larc.nasa.gov/project/sage2/sage2_v7_table) and Calipso (https://eosweb.larc.nasa.gov/project/calipso/cal_lid_l3_apro_cloudfree-standard-V3-00). ~~In the standard run, the~~ The ground albedo is set to 0.15 over sea and 0.3 over land, accounting for the surface reflectance and broken cloud cover. The RT model is further ~~fed~~ provided with the actual geolocation of the HALO aircraft, solar zenith and azimuth angles as encountered during each measurement, the telescopes azimuth and EAs, as well as the ~~field of view (FOV)~~ FOV of the mini-DOAS telescopes. Stutz et al. (2017) show in their Figure 5 one example of simulated measurements for limb observations at about 18 km altitude. The simulations demonstrate that the Earth's sphericity, the correct treatment of atmospheric refraction, cloud cover, ground albedo, ~~celestial data, wavelength et cetera~~ observation geometry, and wavelength dependency of scattering effects are relevant in the context of the interpretation of UV/vis/nearIR limb measurements performed within the lower and middle atmosphere (Deutschmann et al., 2011). Even though the HALO mini-DOAS spectrometers are not radiometrically calibrated on an absolute scale, past comparison exercises with independently measured and McArtim simulated limb radiance provided confidence on the quality of the RT simulations (see Fig. 5 and Fig. 6 in Deutschmann et al. (2011), Fig. 2 in Kreycy et al. (2013), and Wolf et al. (2017)).

For the forward simulations of the trace gas absorptions measured in limb direction, the RT model is ~~further fed with~~ run using simulated trace gas curtains along the flight track (for details see section 3.5 and Figure 3, panels a and b). The term "curtain" is furtheron used to describe the magnitude of a (simulated) atmospheric parameter as a function of time/horizontal coordinate and altitude.

3.4 Additional measurements

3.4.1 ~~Fairo~~FAIRO

FAIRO is a new, light-weight (14.5 kg) and accurate 2-sensor device for in situ measuring O_3. It combines two techniques, i.e. (a) a UV photometer that measures the light absorption of O_3 at a wavelength of λ = 250 - 260 nm emitted by a UV-LED and (b) a chemiluminescence detector that monitors the chemiluminescence generated by O_3 on the surface of an organic dye adsorbed on dry silica gel (Zahn et al., 2012). Both techniques are simultaneously applied in order to combine the high measurement accuracy of UV photometry with the high measurement frequency of chemiluminescence detection. The UV photometer shows a 1-σ precision of 0.08 ppb at a measurement frequency of 0.25 Hz (and a pressure of 1 bar) and an accuracy of 1.5% (determined by the uncertainty of the O_3 cross section). The chemiluminescence detector shows a precision of 0.05 ppb at a measurement frequency of 12.5 Hz~~(Zahn et al., 2012)~~. In post-processing the chemiluminescence detector data is calibrated using the UV photometer data. FAIRO was first deployed on HALO during the TACTS/ESMVal mission (July to September 2012); its performance was excellent during all 13 flights.

3.4.2 AENEAS

NO and NO_y measurements on board HALO are performed using a two-channel chemiluminescence detector (AENEAS - Atmospheric nitrogen oxide measurement system) in combination with a catalytic conversion technique (Ziereis et al., 2000; Stratmann et al., 2016). A commercial two-channel chemiluminescence detector (ECO PHYSICS, Switzerland) is modified for use on board of research aircrafts. The chemiluminescence technique is widely used for the detection of atmospheric NO and relies on the emission of light in the near infrared following the reaction of NO with O_3 (e.g. Drummond et al., 1985). Heated gold tubes in combination with CO or H_2 as reducing agent are frequently used to convert all species of the odd nitrogen family (NO_2, HNO_2, HNO_3, HO_2NO_2, N_2O_5, PAN, ...) into NO (e.g. Bollinger et al., 1983; Fahey et al., 1985) that is subsequently detected by chemiluminescence. The conversion efficiency of the gold converter is quantified using gas phase titration of NO and O_3 before and after each flight with a conversion efficiency of typically more than 98%. The statistical detection limit is 7 pmol/mol for the NO measurements and 8 pmol/mol for the NO_y measurements for an integration time of 1 s. The overall uncertainty for the NO and NO_y measurements is 8% (6.5%) for volume mixing ratios of 0.5 nmol/mol (1 nmol/mol).

3.4.3 TRIHOP

The TRIHOP instrument is a three channel Quantum Cascade Laser Infrared Absorption spectrometer capable of the subsequent measurement of CO, CO_2, CH_4, and N_2O (Schiller et al., 2008; Müller et al., 2016). The instrument applies Quantum Cascade Laser Absorption Spectroscopy (QCLAS) in the mid-infrared with a multipass absorption cell (type White), which is kept at a constant pressure of p = 30 hPa and has a path length of 64 m and a volume of 2.7 L. During TACTS/ESMVal the instrument is in situ calibrated approx. every 30 min during the flights against a secondary standard of compressed ambient air. The mixing ratios of the secondary standard are determined before and after the mission in the laboratory against National Oceanic and Atmospheric Administration (NOAA) standards. Therefore, the in-flight calibrations allow to identify and correct slow instrumental drifts in the post-flight data evaluation. The integration time for each species is 1.5 s at a duty cycle of 8 s, which finally limits the temporal resolution of the measurements. During TACTS/ESMVal TRIHOP CH_4 (N_2O) data achieved a 2-σ precision of 10 (1.1) ppbv and stability of the instrument of 15 (2.2) ppbv, respectively, before applying the post flight data correction. The total uncertainty relative to the working standard of 18 (2.5) ppbv can be regarded as an upper limit.

3.5 Chemistry transport and chemistry climate models

The output of the CTM CLaMS and the CCM EMAC are used in the present study. They differ in a number of ways, in particular in their representation of dynamical features of the atmosphere and the used chemistry schemes. The models are introduced in the following and their differences are highlighted later in sections 3.7.2 and 4.2 in the context of the scaling method.

CLaMS is a Lagrangian CTM system developed at Forschungszentrum Jülich, Germany. The specific model setup is described in detail by Vogel et al. (2015). It is driven by horizontal winds from ERA-Interim reanalysis (Dee et al., 2011) provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). The horizontal resolution is 100 km and

the simulation period ranges from May 2012 until October 2012. It is initialized using satellite data from AURA-MLS and ACE-FTS as well as tracer-tracer-correlations. For further details of the model simulation, see Vogel et al. (2015) and references therein. Due to its Lagrangian design, the model is especially good at representing trace gradients (e.g. the extratropical tropopause or the polar vortex edge). It should be noted that the present ClaMS simulation is not optimized in particular to reproduce photochemical processes in the lower troposphere. Therefore, the employed chemistry setup does only contain reactions of importance within the stratosphere (Grooß et al., 2014) and it does neither contain sources of larger hydro-carbon compounds (e.g. VOCs and NMHCs) nor any interactions of the chemical compounds with clouds.

The ECHAM/MESSy Atmospheric Chemistry (EMAC, http://www.messy-interface.org/) model is a numerical chemistry and climate simulation system that includes sub-models describing processes in the troposphere and middle atmosphere and their interaction with oceans, land and human influences (Jöckel et al., 2010). It uses the second version of the Modular Earth Submodel System (MESSy2) to link multi-institutional computer codes. The core atmospheric model is the 5th generation European Centre Hamburg general circulation model (ECHAM5, Roeckner et al., 2006). Here, we analyse data of the RC1SD-base-10a simulation (Jöckel et al., 2016) sampled along the aircraft flight track with the submodel S4D (Jöckel et al., 2010). The time resolution is the model time step length, i.e., 12 minutes for the applied model resolution. For the RC1SD-base-10a simulation, EMAC has been nudged towards ERA-Interim reanalysis data (Dee et al., 2011) to reproduce the "observed" synoptic situation in the model (for details see Jöckel et al., 2016). The model is applied in the T42L90MA-resolution, i.e. with a spherical truncation of T42 (corresponding to a quadratic Gaussian grid of approx. 2.8 by 2.8 degrees in latitude and longitude) with 90 vertical hybrid pressure levels up to 0.01 hPa. In contrast to CLaMS, EMAC contains a very detailed tropospheric chemistry scheme. The submodel MECCA (Module Efficiently Calculating the Chemistry of the Atmosphere, Sander et al., 2011a) is used to simulate the chemical kinetics, with the photochemical data taken from the JPL compilation (Sander et al., 2011b), including recent updates (Jöckel et al., 2016, Section 3.5).

3.6 The scaling method

The scaling method makes use of the information on the relevant ~~radiative transfer~~ RT gained from a simultaneously in situ and remotely (line-of-sight) measured scaling gas P and the remotely measured absorption of the target gas X to infer the absolute concentration $[X]$ (Raecke, 2013; Großmann, 2014; Werner et al., 2017; Stutz et al., 2017). Ideally, the absorption bands of X (e.g. NO_2) and P (~~Table 3~~O_3, O_4) are close to each other in order to diminish the influence of wavelength dependent Rayleigh and Mie scattering on the results ~~. The potential~~ (Table 3). The advantages of the scaling method over optimal estimation come from largely removing uncertainties in ~~radiative transfer~~ RT due to aerosols and clouds.

Mathematically, the method evolves along the following lines. The total measured SCD (= dSCD + SCD_{ref}) (eq. 1) can be split into slant column densities ($[X]_i \cdot B_{X_i} \cdot z_i$) of individual atmospheric layers i of thickness z_i with concentrations $[X]_i$ and so called box air mass factors (BoxAMFs) B_{X_i} for the targeted gas X (here BrO and NO_2) and the scaling gas P (here O_3 and O_4), i.e.

$$\mathrm{SCD}_X = \sum_i [X]_i \cdot B_{X_i} \cdot z_i \tag{2}$$

$$\mathrm{SCD}_P = \sum_i [P]_i \cdot B_{P_i} \cdot z_i, \tag{3}$$

For the atmospheric layer of interest j, i.e. the altitude range around aircraft altitude where the limb line of sight penetrates through and most of the absorption is picked up, the concentrations for both gases can be expressed as

$$[X]_j = \frac{\mathrm{SCD}_X - \sum_{i \neq j} [X]_i \cdot B_{X_i} \cdot z_i}{B_{X_j} \cdot z_j} \tag{4}$$

$$[P]_j = \frac{\mathrm{SCD}_P - \sum_{i \neq j} [P]_i \cdot B_{P_i} \cdot z_i}{B_{P_j} \cdot z_j}. \tag{5}$$

By noting that for weak absorbers (i.e. those with optical densities much smaller than unity), the BoxAMFs B_{X_j} and B_{P_j} are the same for both gases X and P when measured in the same wavelength range, the ratio of equations 4 and 5 yields:

$$\frac{[X]_j}{[P]_j} = \left(\frac{\mathrm{SCD}_X - \sum_{i \neq j} [X]_i \cdot B_{X_i} \cdot z_i}{\mathrm{SCD}_P - \sum_{i \neq j} [P]_i \cdot B_{P_i} \cdot z_i} \right) \tag{6}$$

Further, by defining so called α-factors (α_X, and α_P), which describe the fraction of the absorption in layer j relative to the total atmospheric absorption for both gases, i.e.

$$\alpha_{X_j} = \frac{\mathrm{SCD}_X - \sum_{i \neq j} [X]_i \cdot B_{X_i} \cdot z_i}{\mathrm{SCD}_X} \tag{7}$$

$$= \frac{[X]_j \cdot B_{X_j} \cdot z_j}{\sum_i [X]_i \cdot B_{X_i} \cdot z_i} \tag{8}$$

and

$$\alpha_{P_j} = \frac{\mathrm{SCD}_P - \sum_{i \neq j} [P]_i \cdot B_{P_i} \cdot z_i}{\mathrm{SCD}_P} \tag{9}$$

$$= \frac{[P]_j \cdot B_{P_j} \cdot z_j}{\sum_i [P]_i \cdot B_{P_i} \cdot z_i}, \tag{10}$$

the main equation of the scaling method can be written as

$$[X]_j = \frac{\alpha_{X_j}}{\alpha_{P_j}} \cdot \frac{\mathrm{SCD}_X}{\mathrm{SCD}_P} \cdot [P]_j \tag{11}$$

$$= \alpha_R \cdot \mathrm{SCD}_R \cdot [P]_j. \tag{12}$$

Here $[P]_j$ is the in situ measured concentration of the scaling gas (e.g., O_3, O_4, ...), but averaged over the time of spectrum integration, and SCD_X, and SCD_P are obtained from HALO mini-DOAS measurements using a DOAS fit and then Eq. (1). α_R and SCD_R are the ratios of the α-factors (obtained from RT model simulations) and the SCDs, respectively. Equations (8) and (10) are solved using the calculated BoxAMFs B_{X_i} and B_{P_i} of atmospheric layer i (RT model described in section 3.3) and the concentrations $[X]_i$ and $[P]_i$ from CTM/CCM predictions (section 3.5).

Figure 3 displays the major ingredients going into the scaling method. It shows CLaMS simulated curtains of concentrations of O_3 (panel a) and NO_2 (panel b), simulated BoxAMFs (panel c), and α-factors for O_3 and NO_2 and their ratio ($\alpha_R = \frac{\alpha_{NO_2}}{\alpha_{O_3}}$) together with its uncertainty (panel d) for the HALO flight from Cape Town to Antarctica and back on 13 September 2012. Measured SCDs and their ratio are shown in panel (e) and the retrieved NO_2 mixing ratio in panel (f). The uncertainties displayed here are discussed in section 3.7.

This flight is chosen to demonstrate the key features of the method and its sensitivity to various parameters. For this flight, leading from the southern subtropics/mid-latitudes into Antarctica in spring, it is expected that (a) the overhead (stratospheric) O_3 and NO_2 concentrations largely vary in space and time and (b) at flight altitude the concentration of both gases is low and (c) in particular NO_2 exhibits strong concentration gradients near the tropopause and between air outside and inside the polar vortex, thus providing a critical case to test the scaling method. For this flight the RT modeled α-factors range from 0.03 to 0.4 for O_3 and 0.02 to 0.3 for NO_2, and α_R ranges from 0.05 to 0.9.

Even though the α-factors are comparably small and largely varying in space and time, the comparison of in situ measured and remotely sensed O_3 indicates a fairly compact relation (Figure 4), similar to other air-borne limb measurements (e.g. Bruns et al., 2004; Baidar et al., 2013; Stutz et al., 2017, Figure 9). Together with RT simulations (Raecke, 2013; Knecht, 2015) this provides confidence in the retrieval of flight level trace gas concentrations from ~~limb~~ UV/vis spectroscopy.

Evidently, the scaling and target gas are not detected at exactly the same wavelength but rather in overlapping wavelength bands. The λ-dependence of α_R is investigated in separate sensitivity simulations. For that purpose α-factors are calculated for the lower and upper wavelength end of the spectral retrieval for each gas. In agreement with Stutz et al. (2017), it is found that α_R may only change by as much as a few percent in our applications. Thus, the error is negligible as compared to the other errors discussed in the following section.

3.7 Errors of the scaling method

The errors and uncertainties of the scaling method fall into the categories random (presumably Gaussian distributed) errors and systematic errors. The sources and magnitudes of both are discussed in the following.

3.7.1 Random errors of the scaling method

The random errors and sensitivities of the scaling method towards all input parameters are addressed by inspecting the Gaussian error propagation of Eq. (12). The uncertainty $\Delta[X]_j$ is calculated from

$$\Delta[X]_j = \left[\left(\frac{\Delta\alpha_R}{\alpha_R}\right)^2 + \left(\frac{\Delta\mathrm{SCD}_R}{\mathrm{SCD}_R}\right)^2 + \left(\frac{\Delta[P]_j}{[P]_j}\right)^2\right]^{0.5} \cdot [X]_j \quad (13)$$

In the following we discuss the different contributions to $\Delta[X]_j$ in Eq. (13). The magnitudes of the contributions are summarised in Table 5.

$\Delta[P]_j$: When using in situ measured O_3 as scaling gas, the uncertainty $\Delta[P]_j$ is given by the uncertainty of the O_3 measurements (~~Fairo~~FAIRO, section 3.4.1). For the comparison of in situ with limb measured O_3 the low frequency (0.25 Hz) precision is obviously most relevant, since the light paths in limb direction average over extended air masses and thus in situ measured O_3 needs to be averaged. At 1 bar the stated O_3 error by ~~Fairo~~ FAIRO is $\leq 1\%$ for $[O_3]$ = 40 ppb. However, in this context more relevant are errors due to horizontal and vertical gradients in the $[O_3]$ which are considered below (see subsection (b) in the paragraph on $\Delta\alpha_R$).

When using O_4 as scaling gas, the altitude and temperature dependent O_4 concentration (in terms of $\mathrm{molec}^2/\mathrm{cm}^6$) can easily be calculated with an uncertainty of $\leq \pm 1\%$ (Greenblatt et al., 1990; Pfeilsticker et al., 2001; Thalman and Volkamer, 2013).

$\left(\frac{\Delta\mathrm{SCD}_R}{\mathrm{SCD}_R}\right)^2 = \left(\frac{\Delta\mathrm{SCD}_P}{\mathrm{SCD}_P}\right)^2 + \left(\frac{\Delta\mathrm{SCD}_X}{\mathrm{SCD}_X}\right)^2$: The $\Delta\mathrm{SCD}_P$ and $\Delta\mathrm{SCD}_X$ errors each have two contributions, i.e. the dSCD errors due to the DOAS retrieval (section 3.1) and the error in determining $\mathrm{SCD}_{\mathrm{ref}}$ (section 3.2), which are added in quadrature. The dSCD error comprises the error of the spectral retrieval and the error of the trace gas cross section. Typical dSCD errors are mentioned in section 3.1 and are often of the order of few percent. Depending on the species, the $\mathrm{SCD}_{\mathrm{ref}}$ errors range from 1% to 20%, but they are typically 10% (see section 3.2).

$\Delta\alpha_R$: The major contribution to the overall error $\Delta[X]_j$ may come from random errors in calculating α_R. In the following their uncertainties (ordered into contributions (a), (b) and (c), see below) are subsequently addressed.

(a) Error due to scattering by aerosols and clouds: The influence of aerosols and clouds on α_R is studied from simulations of UV/vis limb measurements in a surrogate cloud field (Figures 3 and 4 in the supplement, and Knecht (2015)). Atmospheric parameters (temperature, pressure, and cloud cover) typical for the rainy season over the Amazon (e.g. Wendisch et al. (2016) and references therein) are assumed for the simulations, because such a scenario may represent the most severe disturbance of the radiative field in the UV/vis spectral range. The configuration of the cloud field is described in the supplement (Figure 3). For the cloudy sky, α_R is narrowly distributed within a range of typically $\Delta\alpha_R \leq \pm 5\%$ around the clear sky case with some outliers within an interval of $\Delta\alpha_R \leq \pm 15\%$ (Figure 4 in the supplement). A notable finding is that α_R follows the assumed concentration ratio of the target gas and scaling gas, however by a somewhat damped amplitude, i.e. within an interval of 0.6 $\leq \alpha_R \leq$ 1.8, whereas the concentration ratio ranges between 0.2 and 1.7. In conclusion the scaling method thus largely removes the uncertainties in the concentration retrieval due the complexity of the ~~radiative transfer~~ RT in the UV/vis spectral range for

a cloudy atmosphere~~, but the~~. The modelled α_R ~~largely~~ depend on the relative profile shapes of the target gas and scaling gas, but not on absolute concentrations. Overall this finding is in agreement with the recent findings of Stutz et al. (2017).

(b) Uncertainties in α_R due to small scale variability not covered by the CTM is addressed by a comparison of CLaMS simulated and ~~Fairo~~ FAIRO measured O_3 (Figure 5 in the supplement). For the HALO flight from Cape Town to Antarctica on 13 September 2012 CLaMS tends to systematically over-predict measured O_3 by up to 400 ppb, most likely due to errors in the vertical advection of the air masses in the sub-polar atmosphere. The impact of such a systematic error on the O_3 scaling is discussed below (see section 3.7.2). Moreover, the difference of measured minus simulated $[O_3]$ clusters around several peaks with typical widths of $\Delta[O_3] \approx 40$ ppb, indicative of the sub-grid variability of $[O_3]$ not captured in the CLaMS simulations. Including the sub-grid variability in the α-factor calculation results in $\Delta\alpha_R < 0.1$ and a typical $\Delta\alpha_R \approx 0.05$. The same comparison for the retrieved NO_2 results in a typical sub-grid variability of 10 ppt and a similar $\Delta\alpha_R$ as for ozone.

(c) The telescope FOV precision and pointing accuracy (sect. 2.4) results in a rectangular window of about 500 m in height (at the location of maximum contribution to the radiance) from which the skylight is received. This is of the order of the vertical resolution of most CTMs and CCMs. It is therefore coherent to consider an uncertainty of $\pm$ 500 m of the altitude where the vertical profile is sampled. In order to test how this uncertainty propagates into $\Delta\alpha_R$ all simulated trace gas profiles are artificially shifted by 500 m upwards and downwards and the largest and lowest α_R are then used as uncertainty boundaries for each measurement geometry.

During most flight sections, $\Delta\alpha_R$ is dominated in equal parts by the uncertainty due to Mie scattering and sub-grid variability. However, if the vertical gradient of the involved trace gases is strong around flight altitude (e.g. at 08:00 – 09:00 UTC in Figure 3), the vertical sampling uncertainty is the dominating effect (Figure 6 in the supplement). The resulting uncertainties are typically $\Delta\alpha_R \approx 10\%$ ... 20% for O_3 and NO_2, and in rare cases of large vertical gradients up to $\Delta\alpha_R \approx 50\%$.

3.7.2 Potential systematic errors of the scaling method

In our study a priori information on the profile shapes is either taken from CTM/CCM modelling, or in the case of O_4 from calculations. It is thus necessary to consider how uncertainties in the predicted profile shapes propagate into the inferred concentrations at flight level.

~~Systematic~~ Since a measure of the uncertainty of modelled trace gas profile shapes is not readily available, systematic errors of α_R are investigated by modifying the involved trace gas concentration profiles in two distinct ways: By (a) changing the concentration of the scaling gas to match the in situ measured concentration while keeping the concentration of the target gas at flight altitude fixed, and (b) by shifting the CTM/CCM predicted concentration profiles of the scaling and target gas vertically in such a way that predicted N_2O concentrations at flight altitude agree with in situ measurements (Figure 7 in the supplement). It is found that errors (or biases) larger than the random error may occur if (a) the scaling gas concentration at flight altitude is significantly mispredicted by the models while the target gas concentration is not (or vice versa) or if (b) the CTM/CCM does not capture a strong vertical ascent/descent of air masses in a region with strong (and different) vertical concentration gradients of scaling and target gas. Both of these aspects need to be considered in the interpretation of measurements derived via the scaling method. For example, comparing predicted and measured concentrations of tropospheric tracers such as CH_4 and N_2O

may give confidence in the representation of ascent/descent processes near the tropopause and thus justify confidence in the predicted trace gas profile shapes. A systematic error can also occur if the investigated trace gases exhibit strong horizontal gradients inside the volume sampled by the telescope's FOV, e.g. at the edge of a tropopause fold or the polar vortex. Thus, some caution is necessary when interpreting measurements close (tens of km) to such gradients.

4 Sensitivity studies

Sensitivity studies regarding the employed scaling gas and the employed CTM/CCM are carried out for the ESMVal flight on 13 September 2012 leading from Cape Town southwards to 65°S and back. The lower edge of the Antarctic polar vortex was penetrated during the flight between approximately 08:00 and 13:00 UTC, i.e. south of 49°S. More information on the flight, in particular the transport of dehydrated air masses form the Antarctic vortex into the upper and middle troposphere can be found in the publication of Rolf et al. (2015).

4.1 Intercomparison of scaling with O_3 and O_4

We compare ~~and validate~~ the inferred $[NO_2]$ for the HALO flight on 13 September 2012, using O_3 and O_4 as scaling gases, respectively. Figure 5 shows calculated α_R (panel a) and inferred $[NO_2]$ (panel b) using either O_3 (red symbols, further on denoted as $[NO_2]_{O_3}$) or O_4 (blue symbols, denoted as $[NO_2]_{O_4}$) as the scaling gas assuming clear skies (continuous lines) or a cloud layer (circles, description in the following paragraph) in the RT calculations. ~~While the~~ The retrieved $[NO_2]_{O_3}$ and $[NO_2]_{O_4}$ agree reasonably well before 13:00 UTC, ~~they differ~~ exhibiting differences below 35 ppt and often as low as 10 ppt. The differences after 13:00 UTC ~~. The difference comes~~ come from the different ~~sensitivity~~ sensitivities of O_3 and O_4 ~~absorption~~ measurements towards the optical state (e.g., cloud cover) of the atmosphere. While ~~absorption due to~~ the concentration of O_3 as well as NO_2 is largest in the stratosphere and usually smaller in the lower troposphere, it is the opposite for ~~the absorption of~~ O_4. Therefore, the shielding effect of lower and mid-level aerosols and clouds is expected to matter most for the limb detection of O_4 in the upper troposphere, but less for O_3 and NO_2.

The shielding effect of low and mid-level aerosols and clouds is investigated by additional RT calculations considering an uniform cloud cover (optical thickness τ = 20) located at 4 – 8 km altitude. The resulting α_R and inferred $[NO_2]$ are indicated as circles in panels a and b of Figure 5. Evidently including the cloud cover reduces α_R in O_4 scaling but does not significantly change α_R in O_3 scaling. Most striking is the influence of (broken) clouds on the O_4 scaling as evidenced by the large reduction in the calculated α_R for measurements prior to 8:00 UTC and after 13:00 UTC. Some proxy information on the cloud cover below the aircraft can be inferred from the colour index calculated from backscattered radiances at 600 nm / 430 nm received by the nadir VIS3 channel (panel c in Figure 5). Unlike for the time period between 9:00 and 12:30 UTC, when a more or less uniform cloud layer prevailed below the aircraft, the broken cloud cover past 13:00 UTC caused inferred $[NO_2]_{O_4}$ to become rather variable. In contrast $[NO_2]_{O_3}$ is much less variable and closely follows the ClaMS/EMAC predicted $[NO_2]$, except for the period between 13:00 and 13:40 UTC. Here the inclusion of a cloud cover in the RT model causes $[NO_2]_{O_4}$ to converge towards $[NO_2]_{O_3}$.

Figure 6 shows the differences in inferred $[NO_2]_{O_3}$ and $[NO_2]_{O_4}$ profiles, assuming clear and cloudy skies. Evidently inferred $[NO_2]_{O_3}$ is much less sensitive to the cloud cover than $[NO_2]_{O_4}$. The small differences (mostly < 5%) at higher altitudes for inferred $[NO_2]_{O_3}$ provide confidence in the $[NO_2]_{O_3}$ retrieval for the upper troposphere and lower stratosphere. In contrast, $[NO_2]_{O_4}$ is strongly dependent on assumptions regarding the cloud cover. These results are in agreement with those reported by Stutz et al. (2017). It is worth noting that within the Antarctic troposphere $[NO_2]$ is found to be rather low (< 20 ppt), and hence the systematic difference in the inferred $[NO_2]$ (up to 50% for the $[NO_2]_{O_3}$ and up to 80% for the $[NO_2]_{O_4}$) indicate the detection limit of the DOAS limb technique for NO_2.

In conclusion the profile shape dependence of the scaling method thus mandates to carefully choose the scaling gas, i.e. O_3 appears more appropriate as a scaling gas for the detection of gases of low tropospheric and large stratospheric abundance when probed from an aircraft flying in the middle and upper troposphere and lowermost stratosphere (e.g., such as NO_2, BrO) while O_4 appears to be more suited for gases of large concentrations in the lower troposphere (e.g., such as CH_2O, $C_2H_2O_2$, IO, and in polluted environments HONO and NO_2) when probed from low flying air-borne vehicles.

4.2 EMAC versus CLaMS profile predictions

Next, the sensitivity of inferred $[NO_2]$ and $[BrO]$ as a function of ~~predicted and~~ the predicted trace gas curtains is investigated. ~~mixing~~ Mixing ratios are retrieved using trace gas curtains predicted by CLaMS (Figure 3) and EMAC (Figure 7).

The retrieved NO_2 mixing ratios agree within the random errors during most flight sections (Figure 8, panel b). However, some differences between the models have an impact on retrieval results, such as the higher spatial and temporal resolution of the CLaMS model. For example, a local maximum in $[NO_2]$ is predicted by CLaMS between 13:00 and 13:30 UTC but not by EMAC (Figures 3 and 7, respective panel f, and Figure 9, panel c). The retrieved $[NO_2]$ using predicted O_3 from CLaMS (further on denoted $[NO_2]_{O_3,CLaMS}$) is $[NO_2]_{O_3,CLaMS} \approx 0.18 \pm 0.02$ ppb, while $[NO_2]_{O_3,EMAC} \approx 0.12 \pm 0.02$ ppb. Compared with the retrieved $[NO_2]$ for this period, the CLaMS prediction appears to be overestimated, while the EMAC prediction appears to be underestimated. Thus, model predictions with spatial resolutions comparable to the measurements (ca. 6 km horizontally) are desirable when applying the scaling method.

~~While~~ In the case of BrO, good agreement is reached ~~for~~ in the range of 2 – 5 ppt in the extratropical lowermost stratosphere (flight sections A and E), but the difference between predicted $[BrO]_{O_3,CLaMS}$ and $[BrO]_{O_3,EMAC}$ are more substantial throughout flight sections B, C, and D (Figure 8, panel a, and Figure 9, panel e). Two reasons for these differences can be identified. First, there is a discrepancy in predicted tropospheric BrO concentrations between the models, which leads to a difference in calculated α_{BrO} at all altitudes. Below 9 km altitude, CLaMS predicts 3 – 5 ppt, while EMAC predicts concentrations close to zero (Figure 9, panel b, dashed and dotted lines). This discrepancy is probably due to missing tropospheric sinks in the CLaMS model (sect. 3.5). Hence, the EMAC-predicted $[BrO]$ profile is expected to be more realistic. Secondly, while the extent of the polar vortex is predicted roughly in the same manner, the treatment of subsidence and methane degradation differs between the models. This can be observed by comparing measured and predicted methane mixing ratios in flight sections B and D (Figure 8, panel c, and Figure 9, panel g). For both flight sections measurements indicate air mass ages up to 4.5 years in combination with strong dehydration (Rolf et al., 2015) and denitrification ~~(?)~~(Jurkat et al., 2017). However, the

subsidence of O_3 appears to be overestimated in the CLaMS model, since the vertical profile of measured O_3 concentrations is more accurately represented by EMAC (Figure 9, panel a).

In conclusion, differences in relative profile shapes predicted by the employed models and their spatial and temporal resolution influence the retrieval results of the scaling method. These differences are particularly large, if fundamental properties of the atmosphere, e.g. the presence of BrO in the troposphere or the subsidence in the polar vortex, are treated differently by the models. In most cases, inferred mixing ratios agree whatever model predictions (CLaMS vs. EMAC) are taken.

5 Sample results and discussion

Finally, we discuss the mini-DOAS observations from the flight on 13 September 2012 in the context of complementary measurements and model predictions (Figures 8 and 9). Beside the mini-DOAS measurements of O_3, NO_2, and BrO complementary instrumentation provided information on the following gases: O_3 from the ~~Fairo~~ FAIRO instrument, NO and total NO_y from the AENEAS instrument, and CO and CH_4 from the TRIHOP instrument (section 3.4). These measurements are further compared with the predictions of CLaMS and EMAC, which support the interpretation with respect to the atmospheric dynamics and photochemistry. Most notable is the joint detection of NO, NO_2, and total NO_y (and of BrO) in a remote location, such as in the Antarctic troposphere and lowermost stratosphere, since such measurements are infrequent or to date not existing. Overall, mixing ratios of BrO and NO_2 are inferred for the whole flight with a time resolution of 30 s and a resulting spatial resolution of 6 km, although ~~radiative transfer~~ RT implies further averaging along the line of sight (perpendicular to flight direction) of 200 km and along flight direction of 10 km. ~~A conservative estimate for the detection limits at~~ The detection limits for the measurements shown here are estimated by inspecting the uncertainty at very low mixing ratios~~is~~, e.g. during the dive (Figure 9, panels c and e). For $[BrO] = 2$ ppt and $[NO_2] = 10$ ppt ~~for and , respectively~~respectively, the inferred mixing ratios are at least two times larger than the uncertainty. Measurements of CH_4, which is well mixed in the troposphere and ~~degrades~~ is lost in the stratosphere, provide a measure of stratospheric age of the air. Accordingly, the flight is subdivided into five flight sections A ... E (Figure 8, panel c) in order to distinguish data recorded in the midlatitude lowermost stratosphere (flight sections A and E), polar winter vortex air (flight sections B and D) and the polar troposphere (flight section C). In September 2012 the tropospheric CH_4 mixing ratio at Cape Grim, Tasmania was 1778 ppb (http://www.csiro.au/greenhouse-gases/).

Inferred BrO mixing ratios are around 4 ppt / 7 ppt in flight section B and 6 ppt / 8 ppt in flight section D, based on retrievals using CLaMS / EMAC in the scaling method, respectively (Panel a of Figure 8; differences between both retrievals are discussed above in section 4.2). These concentrations are on the higher end of comparable BrO measurements in the same altitude range (12 – 13 km) reported in the literature (Harder et al., 1998; Dorf et al., 2006; Hendrick et al., 2007; Werner et al., 2017), which could be caused by the subsidence of stratospheric air from higher altitudes discussed above. Panel e of Figure 9 shows the vertical BrO profile retrieved from the ascent of the dive at 65°S. The retrieved $[BrO]_{O_3,EMAC}$ and $[BrO]_{O_3,CLaMS}$ are both below the detection limit of 2 ppt in the altitude range below 9.5 km, even when using RT calculations based on CLaMS, which predicts 3 ppt BrO in the troposphere. Hence, below 9.5 km altitude BrO could not be detected above the detection limit. The amount and distribution of halogen oxides such as BrO (panel e) in the troposphere is a matter of current

debate (Harder et al., 1998; Fitzenberger et al., 2000; Van Roozendael et al., 2002; Saiz-Lopez and von Glasow, 2012; Volkamer et al., 2015; Wang et al., 2015; Schmidt et al., 2016; Sherwen et al., 2016; Werner et al., 2017) and is of significant scientific interest due to its potential influence on tropospheric ozone chemistry (von Glasow et al., 2004) and thus radiative forcing (Sherwen et al., 2017). Reported tropospheric background profiles at polar latitudes include those by Fitzenberger et al. (2000), who derive tropospheric BrO profiles above Kiruna (Sweden) from balloon measurements and conclude that tropospheric [BrO] amounting to 0.4 – 2.3 ppt ~~eventually~~ was present, assuming a uniform distribution within the troposphere. Prados-Roman et al. (2011) use air-borne DOAS measurements based in Spitzbergen to derive a BrO mixing ratio profile in Arctic spring with 15 ppt in the planetary boundary layer~~(PBL)~~, 1.5 ppt in the free troposphere~~(FT)~~, and up to 6 ppt at 10 km in the lowermost stratosphere~~(LMS)~~. The measurements derived in the present study are compatible with these previously inferred background profiles and do not show elevated BrO concentrations in the Antarctic ~~spring~~ free troposphere in September 2012.

Retrieved NO_2 (Figure 8, panel b) exhibits similar features as the independently measured NO. In polar vortex air (flight sections B and D), $[NO_2]$ is mostly between 5 and 20 ppt, i.e. near or below the detection limit of 10 ppt, similar to the in situ measured [NO]. Such small amounts of NO_x limit the deactivation of active chlorine, i.e. the formation of $ClONO_2$, and thus prolong ozone destruction in the polar winter vortex air. Interestingly, enhanced NO_2 together with increased NO are detected in the free troposphere (9 – 13 km altitude) during the dive (Figure 9, panels b and c). The largest [NO] of 60 – 80 ppt is measured around 10 – 12 km, while the largest $[NO_2]$ of 30 – 40 ppt is inferred around 11 – 13 km. Increased NO_2 concentrations are also predicted by the CLaMS model (Figure 3, panel b), indicating in-mixing of tropospheric air from more NO_x rich ~~latitudes~~air masses. Rolf et al. (2015) also infer in-mixing of moister mid-latitude air into the bottom of the polar vortex, albeit not at the same time during the flight, since the ~~GLORIA instrument~~ used instrument (GLORIA) was switched off during the dive. At altitudes below 9 and 10 km, respectively, retrieved [NO] and $[NO_2]$ are near their detection limits of 7 ppt and 10 ppt, indicating very pristine air.

6 Conclusions

We describe a novel six channel optical spectrometer for ~~aircraft-borne~~ air-borne limb and nadir detection of UV/visible/nearIR absorbing gases (e.g., O_3, O_4, NO_2, HONO, CH_2O, $C_2H_2O_2$, BrO, IO, OClO, and others), liquid and solid water as well as of skylight radiances. Further, features of a novel retrieval method (called scaling method) are discussed which go beyond those recently reported by Stutz et al. (2017) and Werner et al. (2017). Here we demonstrate how absolute concentrations of the UV/vis absorbing gases can be inferred from limb measurements in the troposphere and lower stratosphere under all (clear and cloudy) skies. The novel scaling methods largely avoids ambiguities in the necessary mathematical inversion to interpret the limb measurements which are introduced by the complexity of ~~radiative transfer~~ RT in the UV/vis spectral range for cloudy and heavily aerosol-loaded atmospheres. Instead, the relative profile shapes of the scaling gas and the target gas are the main a priori information used in the scaling method. Thus, uncertainties in the trace gas retrieval are primarily due to uncertainties in the relative profile shapes, which can be minimized when the retrieval uses a priori profile shapes for example from CTM/CCM predictions, calculations (e.g. O_4), or otherwise available measurements. The present study examines the resulting random and

systematic errors of trace gas concentrations retrieved via the scaling method. The random error is estimated to be 10 – 20 % for most measurement conditions, dominated in equal parts by uncertainties due to Mie extinction and small scale variabilities of the concentrations of the involved trace gases. The random error is comparatively large close to strong vertical or horizontal trace gas concentration gradients. Systematic biases can occur when trace gas profile shapes are strongly misrepresented by model predictions. Thus, comparing independent trace gas measurements of e.g. tropospheric or stratospheric tracers with model predictions is essential in the interpretation of retrieval results. For limb measurements in the upper troposphere and lower stratosphere the comparison of both scaling gases indicates a sensitivity of O_4 scaling for low clouds, while the O_3 scaling is insensitive. This is consistent with the expectation that a scaling gas with similar profile shape as compared to the target gas is best suited for the method. The comparison of retrievals involving a CTM (CLaMS) and a CCM (EMAC) reveals that results are in agreement within the random error, as long as the fundamental properties of the atmosphere are represented in a similar way (e.g. presence or absence of a trace gas in the troposphere). Further, the comparison indicates that CTM/CCM curtains with spatial resolutions close to those of the measurements are desirable.

The present study shows the applicability of the scaling method to HALO mini-DOAS measurements of NO_2 and BrO ~~at altitudes between 3.5 and 15 km~~ under all sky conditions. ~~It can be argued that the~~ The scaling method replaces the ~~major a priori~~ constraint on RT used in the traditional optimal estimation (~~i.e.~~e.g. O_4 to infer the aerosol and cloud profile) by a ~~different a priori (i.e. the~~ scaling factor to estimate effective light path lengths (derived from relative trace gas profile ~~shape~~shapes and RT modelling). The latter ~~a priori~~ approach has the advantages that (a) ~~it is~~ trace gas concentration profiles are more homogeneous in space and time on the scales relevant for ~~the~~ air-borne DOAS measurements than cloud patterns, and (b) ~~it~~ the former can be predicted more reliably by modern CTMs/CCMs as compared to the presence of aerosols and clouds. Thus, the scaling method provides a novel and reliable means for inferring trace gas concentrations from air-borne UV/vis limb measurements. The significantly decreased dependency on aerosol and cloud properties increases the ability to make use of already recorded data and decidedly widens the applicability of air-borne UV/vis limb spectroscopy as a means of investigating atmospheric photochemistry.

Acknowledgements. This study was funded through the Deutsche ForschungsGemeinschaft, DFG (HALO-SPP 1294 and grants PF-384/7-1, PF-384/7-2, PF384/9-1, PF384/9-2, and PF-384/16-1). Additional funding from EU-SHIVA (FP7-ENV-2007-1-226224) is highly acknowledged. The authors gratefully acknowledge the computing time granted on the supercomputer JURECA at Jülich Supercomputing Centre (JSC) under the VSR project ID JICG11. The EMAC simulations have been performed at the German Climate Computing Centre (DKRZ) through support from the Bundesministerium für Bildung und Forschung (BMBF). DKRZ and its scientific steering committee are gratefully acknowledged for providing the HPC and data archiving resources for this consortial project ESCiMo (Earth System Chemistry integrated Modelling). We thank the Deutsches Zentrum für Luft- und Raumfahrt (DLR) for the support to get the instrument certificated and the DLR Flugexperimente Team at Oberpfaffenhofen, in particular Heinrich Brockstieger, Frank Probst, Martina Hierle, and Andrea Hausold, for the support given during the TACTS/ESMVal, NarVal, ~~Cirrus, Acridicon~~ML-Cirrus, ACRIDICON, OMO, and ~~Polstracc~~ POLSTRACC missions.

References

Aliwell, S. R., Van Roozendael, M., Johnston, P. V., Richter, A., Wagner, T., Arlander, D. W., Burrows, J. P., Fish, D. J., Jones, R. L., Tørnkvist, K. K., Lambert, J.-C., Pfeilsticker, K., and Pundt, I.: Analysis for BrO in zenith-sky spectra: An intercomparison exercise for analysis improvement, Journal of Geophysical Research: Atmospheres, 107, ACH 10–1–ACH 10–20, doi:10.1029/2001JD000329, http://dx.doi.org/10.1029/2001JD000329, 2002.

Baidar, S., Oetjen, H., Coburn, S., Dix, B., Ortega, I., Sinreich, R., and Volkamer, R.: The CU Airborne MAX-DOAS instrument: vertical profiling of aerosol extinction and trace gases, Atmospheric Measurement Techniques, 6, 719–739, doi:10.5194/amt-6-719-2013, http://www.atmos-meas-tech.net/6/719/2013/, 2013.

Beirle, S., Boersma, K. F., Platt, U., Lawrence, M. G., and Wagner, T.: Megacity Emissions and Lifetimes of Nitrogen Oxides Probed from Space, Science, 333, 1737–1739, doi:10.1126/science.1207824, http://science.sciencemag.org/content/333/6050/1737, 2011.

Bogumil, K., Orphal, J., Homann, T., Voigt, S., Spietz, P., Fleischmann, O., Vogel, A., Hartmann, M., Kromminga, H., Bovensmann, H., et al.: Measurements of molecular absorption spectra with the SCIAMACHY pre-flight model: instrument characterization and reference data for atmospheric remote-sensing in the 230–2380 nm region, Journal of Photochemistry and Photobiology A: Chemistry, 157, 167–184, 2003.

Bollinger, M. J., Sievers, R. E., Fahey, D. W., and Fehsenfeld, F. C.: Conversion of nitrogen dioxide, nitric acid, and n-propyl nitrate to nitric oxide by a gold-catalyzed reduction with carbon monoxide, Analytical Chemistry, 55, 1980–1986, 1983.

Brandtjen, R., Klüpfel, T., Perner, D., and Knudsen, B. M.: Airborne measurements during the European Arctic Stratospheric Ozone Experiment: Observation of OClO, Geophysical Research Letters, 21, 1363–1366, doi:10.1029/93GL01868, http://dx.doi.org/10.1029/93GL01868, 1994.

Bruns, M., Buehler, S. A., Burrows, J. P., Heue, K.-P., Platt, U., Pundt, I., Richter, A., Rozanov, A., Wagner, T., and Wang, P.: Retrieval of profile information from airborne multiaxis UV-visible skylight absorption measurements., Applied Optics, 43, 4415 – 4426, doi:10.1364/AO.43.004415, 2004.

Bruns, M., Buehler, S. A., Burrows, J. P., Richter, A., Rozanov, A., Wang, P., Heue, K. P., Platt, U., Pundt, I., and Wagner, T.: NO_2 profile retrieval using airborne multi axis UV-visible skylight absorption measurements over central Europe, Atmos. Chem. Phys., 6, 3049–3058, doi:10.5194/acpd-6-493-2006, 2006.

Butz, A., Bösch, H., Camy-Peyret, C., Chipperfield, M. P., Dorf, M., Dufour, G., Grunow, K., P., J., Kühl, S., Payan, S., Pepin, I., Pukite, J., Rozanov, A., von Savigny, C., Sioris, C., Wagner, T., Weidner, F., and Pfeilsticker, K.: Inter-comparison of Stratospheric O_3 and NO_2 abundances retrieved from balloon-borne direct sun observations and Envisat/SCIAMACHY limb measurements, Atmos. Chem. Phys., 6, 1293–1314, 2006.

Chance, K. and Orphal, J.: Revised ultraviolet absorption cross sections of H 2 CO for the HITRAN database, Journal of Quantitative Spectroscopy and Radiative Transfer, 112, 1509–1510, 2011.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, doi:10.1002/qj.828, http://dx.doi.org/10.1002/qj.828, 2011.

Deutschmann, T., Beirle, S., Frieß, U., Grzegorski, M., Kern, C., Kritten, L., Platt, U., Pukite, J., Wagner, T., Werner, B., and Pfeilsticker, K.: The Monte Carlo Atmospheric Radiative Transfer Model McArtim: Introduction and Validation of Jacobians and 3D Features, Journal of Quantitative Spectroscopy and Radiative Transfer, 112, 1119–1137, 2011.

Dix, B., Brenninkmeijer, C. A. M., Frieß, U., Wagner, T., and Platt, U.: Airborne multi-axis DOAS measurements of atmospheric trace gases on CARIBIC long-distance flights, Atmospheric Measurement Techniques, 2, 639–652, doi:10.5194/amt-2-639-2009, http://www.atmos-meas-tech.net/2/639/2009/, 2009.

Dorf, M., Bösch, H., Butz, A., Camy-Peyret, C., Chipperfield, M. P., Engel, A., Goutail, F., Grunow, K., Hendrick, F., Hrechanyy, S., Naujokat, B., Pommereau, J.-P., Van Roozendael, M., Sioris, C., Stroh, F., Weidner, F., and Pfeilsticker, K.: Balloon-borne stratospheric BrO measurements: comparison with Envisat/SCIAMACHY BrO limb profiles, Atmospheric Chemistry and Physics, 6, 2483–2501, doi:10.5194/acp-6-2483-2006, http://www.atmos-chem-phys.net/6/2483/2006/, 2006.

Dorf, M., Butz, A., Camy-Peyret, C., Chipperfield, M. P., Kritten, L., and Pfeilsticker, K.: Bromine in the tropical troposphere and stratosphere as derived from balloon-borne BrO observations, Atmospheric Chemistry and Physics, 8, 7265–7271, doi:10.5194/acp-8-7265-2008, http://www.atmos-chem-phys.net/8/7265/2008/, 2008.

Drummond, J. W., Volz, A., and Ehhalt, D. H.: An optimized chemiluminescence detector for tropospheric NO measurements, Journal of Atmospheric Chemistry, 2, 287–306, doi:10.1007/BF00051078, http://dx.doi.org/10.1007/BF00051078, 1985.

Fahey, D., Eubank, C., Hübler, G., and Fehsenfeld, F.: Evaluation of a catalytic reduction technique for the measurement of total reactive odd-nitrogen NO y in the atmosphere, Journal of Atmospheric Chemistry, 3, 435–468, 1985.

Ferlemann, F., Bauer, N., Fitzenberger, R., Harder, H., Osterkamp, H., Perner, D., Platt, U., Scheider, M., Vradelis, P., and Pfeilsticker, K.: Differential optical absorption spectroscopy instrument for stratospheric balloon-borne trace gas studies, Appl. Opt., 39, 2377–2386, 2000.

Fitzenberger, R., Bösch, H., Camy-Peyret, C., Chpperfield, M., Harder, H., Platt, U., Pyle, J., Wagner, T., and Pfeilsticker, K.: First profile measurement of tropospheric BrO, Geophys. Res. Lett., 27, 2921–2924, 2000.

Fleischmann, O. C., Hartmann, M., Burrows, J. P., and Orphal, J.: New ultraviolet absorption cross-sections of BrO at atmospheric temperatures measured by time-windowing Fourier transform spectroscopy, Journal of Photochemistry and Photobiology A: Chemistry, 168, 117–132, 2004.

General, S., Pöhler, D., Sihler, H., Bobrowski, N., Frieß, U., Zielcke, J., Horbanski, M., Shepson, P. B., Stirm, B. H., Simpson, W. R., Weber, K., Fischer, C., and Platt, U.: The Heidelberg Airborne Imaging DOAS Instrument (HAIDI) – a novel imaging DOAS device for 2-D and 3-D imaging of trace gases and aerosols, Atmospheric Measurement Techniques, 7, 3459–3485, doi:10.5194/amt-7-3459-2014, http://www.atmos-meas-tech.net/7/3459/2014/, 2014.

Gerilowski, K., Tretner, A., Krings, T., Buchwitz, M., Bertagnolio, P. P., Belemezov, F., Erzinger, J., Burrows, J. P., and Bovensmann, H.: MAMAP – a new spectrometer system for column-averaged methane and carbon dioxide observations from aircraft: instrument description and performance analysis, Atmospheric Measurement Techniques, 4, 215–243, doi:10.5194/amt-4-215-2011, http://www.atmos-meas-tech.net/4/215/2011/, 2011.

Gorshelev, V., Serdyuchenko, A., Weber, M., Chehade, W., and Burrows, J.: High spectral resolution ozone absorption cross-sections–Part 1: Measurements, data analysis and comparison with previous measurements around 293 K, Atmospheric Measurement Techniques, 7, 609–624, 2014.

Gratz, L. E., Ambrose, J. L., Jaffe, D. A., Shah, V., Jaeglé, L., Stutz, J., Festa, J., Spolaor, M., Tsai, C., Selin, N. E., Song, S., Zhou, X., Weinheimer, A. J., Knapp, D. J., Montzka, D. D., Flocke, F. M., Campos, T. L., Apel, E., Hornbrook, R., Blake, N. J., Hall, S., Tyndall,

G. S., Reeves, M., Stechman, D., and Stell, M.: Oxidation of mercury by bromine in the subtropical Pacific free troposphere, Geophysical Research Letters, 42, 10,494–10,502, doi:10.1002/2015GL066645, http://dx.doi.org/10.1002/2015GL066645, 2015GL066645, 2015.

Greenblatt, G. D., Orlando, J. J., Burkholder, J. B., and Ravishankara, A. R.: Absorption measurements of oxygen between 330 and 1140 nm, J. Geophys. Res., 95, 18 577–18 582, 1990.

Grooß, J.-U., Engel, I., Borrmann, S., Frey, W., Günther, G., Hoyle, C. R., Kivi, R., Luo, B. P., Molleker, S., Peter, T., Pitts, M. C., Schlager, H., Stiller, G., Vömel, H., Walker, K. A., and Müller, R.: Nitric acid trihydrate nucleation and denitrification in the Arctic stratosphere, Atmospheric Chemistry and Physics, 14, 1055–1073, doi:10.5194/acp-14-1055-2014, http://www.atmos-chem-phys.net/14/1055/2014/, 2014.

Großmann, K.: Aircraft-borne DOAS limb observations of UV/visible absorbing trace gas species over Borneo: Implications for the photochemistry of iodine, volatile organic oxide degradation, and lightning-produced radicals, Phd, University of Heidelberg, Heidelberg, Germany, doi:10.11588/heidok.00017874, http://www.ub.uni-heidelberg.de/archiv/17874, 2014.

Gurlit, W., Bösch, H., Bovensmann, H., Burrows, J., Butz, A., Camy-Peyret, C., Dorf, M., Gerilowski, K., Lindner, A., Noel, S., Platt, U., Weidner, F., and Pfeilsticker, K.: The UV-A and visible solar irradiance spectrum: Inter-comparison of absolutely calibrated, spectrally medium resolved solar irradiance spectra from balloon- and satellite-borne measurements, Atmos. Chem. Phys., 5, 1879–1890, 2005.

Harder, H., Camy-Peyret, C., Ferlemann, F., Fitzenberger, R., Hawat, T., Osterkamp, H., Schneider, M., Perner, D., Platt, U., Vradelis, P., and Pfeilsticker, K.: Stratospheric BrO profiles measured at different latitudes and seasons: atmospheric observations, Geophys. Res. Lett., 25, 3843–3846, 1998.

Hendrick, F., Van Roozendael, M., Chipperfield, M. P., Dorf, M., Goutail, F., Yang, X., Fayt, C., Hermans, C., Pfeilsticker, K., Pommereau, J.-P., Pyle, J. A., Theys, N., and De Mazière, M.: Retrieval of stratospheric and tropospheric BrO profiles and columns using ground-based zenith-sky DOAS observations at Harestua, 60^o N, Atmospheric Chemistry and Physics, 7, 4869–4885, doi:10.5194/acp-7-4869-2007, http://www.atmos-chem-phys.net/7/4869/2007/, 2007.

Heue, K.-P., Wagner, T., Broccardo, S. P., Walter, D., Piketh, S. J., Ross, K. E., Beirle, S., and Platt, U.: Direct observation of two dimensional trace gas distributions with an airborne Imaging DOAS instrument, Atmospheric Chemistry and Physics, 8, 6707–6717, doi:10.5194/acp-8-6707-2008, http://www.atmos-chem-phys.net/8/6707/2008/, 2008.

Heue, K.-P., Riede, H., Walter, D., Brenninkmeijer, C. A. M., Wagner, T., Frieß, U., Platt, U., Zahn, A., Stratmann, G., and Ziereis, H.: CARIBIC DOAS observations of nitrous acid and formaldehyde in a large convective cloud, Atmospheric Chemistry and Physics, 14, 6621–6642, doi:10.5194/acp-14-6621-2014, https://www.atmos-chem-phys.net/14/6621/2014/, 2014.

Hüneke, T.: The scaling method applied to HALO measurements: Inferring absolute trace gas concentrations from airborne limb spectroscopy under all sky conditions, Phd, University of Heidelberg, Heidelberg, Germany, http://www.ub.uni-heidelberg.de/archiv/22573, 2016.

Jöckel, P., Kerkweg, A., Pozzer, A., Sander, R., Tost, H., Riede, H., Baumgaertner, A., Gromov, S., and Kern, B.: Development cycle 2 of the modular earth submodel system (MESSy2), Geoscientific Model Development, 3, 717–752, 2010.

Jöckel, P., Tost, H., Pozzer, A., Kunze, M., Kirner, O., Brenninkmeijer, C. A. M., Brinkop, S., Cai, D. S., Dyroff, C., Eckstein, J., Frank, F., Garny, H., Gottschaldt, K.-D., Graf, P., Grewe, V., Kerkweg, A., Kern, B., Matthes, S., Mertens, M., Meul, S., Neumaier, M., Nützel, M., Oberländer-Hayn, S., Ruhnke, R., Runde, T., Sander, R., Scharffe, D., and Zahn, A.: Earth System Chemistry integrated Modelling (ESCiMo) with the Modular Earth Submodel System (MESSy) version 2.51, Geoscientific Model Development, 9, 1153–1200, doi:10.5194/gmd-9-1153-2016, http://www.geosci-model-dev.net/9/1153/2016/, 2016.

Jurkat, T., Voigt, C., Kaufmann, S., Zahn, A., Sprenger, M., Hoor, P., Bozem, H., Müller, S., Dörnbrack, A., Schlager, H., Bönisch, H., and Engel, A.: A quantitative analysis of stratospheric HCl, HNO_3, and O_3 in the tropopause region near the subtropical jet, Geophysical Research Letters, 41, 3315–3321, doi:10.1002/2013GL059159, http://dx.doi.org/10.1002/2013GL059159, 2014.

Jurkat, T., Voigt, C., Kaufmann, S., Grooß, J.-U., Ziereis, H., Dörnbrack, A., Hoor, P., Bozem, H., Engel, A., Bönisch, H., Keber, T., Hüneke, T., Pfeilsticker, K., Zahn, A., Walker, K. A., Boone, C. D., Bernath, P. F., and Schlager, H.: Depletion of ozone and reservoir species of chlorine and nitrogen oxide in the lower Antarctic polar vortex measured from aircraft, Geophysical Research Letters, pp. n/a–n/a, doi:10.1002/2017GL073270, http://dx.doi.org/10.1002/2017GL073270, 2017GL073270, 2017.

Klepp, C., Ament, F., Bakan, S., Hirsch, L., and Stevens, B.: The NARVAL Campaign Report, 2014.

Knecht, M.: Simulation of radiative field modification due to tropical clouds, Master's thesis, Institut für Umweltphysik, Universität of Heidelberg, Heidelberg, Germany, http://www.iup.uni-heidelberg.de/institut/forschung/groups/atmosphere/stratosphere/publications/pdf/MA_thesis_Matthias_final.pdf, 2015.

Kreycy, S., Camy-Peyret, C., Chipperfield, M. P., Dorf, M., Feng, W., Hossaini, R., Kritten, L., Werner, B., and Pfeilsticker, K.: Atmospheric test of the $\mathrm{J(BrONO_2)/k_{BrO+NO_2}}$ ratio: implications for total stratospheric Br_y and bromine-mediated ozone loss, Atmospheric Chemistry and Physics, 13, 6263–6274, doi:10.5194/acp-13-6263-2013, http://www.atmos-chem-phys.net/13/6263/2013/, 2013.

Kritten, L., Butz, A., Dorf, M., Deutschmann, T., Kühl, S., Prados-Roman, C., Pukite, J., Rozanov, A., Schofield, R., and Pfeilsticker, K.: Time dependent profile retrieval of UV/vis absorbing radicals from balloon-borne limb measurements – a case study on NO_2 and O_3, Atmospheric Measurement Techniques, 3, 933–946, doi:10.5194/amt-3-933-2010, https://www.atmos-meas-tech.net/3/933/2010/, 2010.

Kritten, L., Butz, A., Chipperfield, M. P., Dorf, M., Dhomse, S., Hossaini, R., Oelhaf, H., Prados-Roman, C., Wetzel, G., and Pfeilsticker, K.: Constraining the $\mathrm{N_2O_5}$ UV absorption cross section from spectroscopic trace gas measurements in the tropical mid-stratosphere, Atmospheric Chemistry and Physics, 14, 9555–9566, doi:10.5194/acp-14-9555-2014, http://www.atmos-chem-phys.net/14/9555/2014/, 2014.

Kromminga, H., Orphal, J., Spietz, P., Voigt, S., and Burrows, J.: New measurements of OClO absorption cross-sections in the 325–435 nm region and their temperature dependence between 213 and 293 K, Journal of Photochemistry and Photobiology A: Chemistry, 157, 149–160, 2003.

McElroy, C. T., McLinden, C. A., and McConnell, J. C.: Evidence for bromine monoxide in the free troposphere during the Arctic polar sunrises, Nature 397, 338-341, 397, 338–341, 1999.

Merlaud, A., Van Roozendael, M., van Gent, J., Fayt, C., Maes, J., Toledo-Fuentes, X., Ronveaux, O., and De Mazière, M.: DOAS measurements of NO_2 from an ultralight aircraft during the Earth Challenge expedition, Atmospheric Measurement Techniques, 5, 2057–2068, doi:10.5194/amt-5-2057-2012, http://www.atmos-meas-tech.net/5/2057/2012/, 2012.

Müller, S., Hoor, P., Bozem, H., Gute, E., Vogel, B., Zahn, A., Bönisch, H., Keber, T., Krämer, M., Rolf, C., Riese, M., Schlager, H., and Engel, A.: Impact of the Asian monsoon on the extratropical lower stratosphere: trace gas observations during TACTS over Europe 2012, Atmospheric Chemistry and Physics, 16, 10 573–10 589, doi:10.5194/acp-16-10573-2016, http://www.atmos-chem-phys.net/16/10573/2016/, 2016.

Noxon, J. F.: Nitrogen dioxide in the stratosphere and troposphere measured by ground-based absorption spectroscopy, Science, 189, 547 – 549, doi:10.1126/science.189.4202.547, 1975.

Noxon, J. F., Whipple, E. C., and Hyde, R. S.: Stratospheric NO2: 1. Observational method and behavior at mid-latitude, Journal of Geophysical Research: Oceans, 84, 5047–5065, doi:10.1029/JC084iC08p05047, http://dx.doi.org/10.1029/JC084iC08p05047, 1979.

Oikarinen, L.: Effect of surface albedo variations on UV-visible limb-scattering measurements of the atmosphere, Journal of Geophysical Research: Atmospheres, 107, ACH 13–1–ACH 13–15, doi:10.1029/2001JD001492, http://dx.doi.org/10.1029/2001JD001492, 4404, 2002.

Pfeilsticker, K. and Platt, U.: Airborne measurements during the Arctic Stratospheric Experiment: Observation of O3 and NO2, Geophysical Research Letters, 21, 1375–1378, doi:10.1029/93GL01870, http://dx.doi.org/10.1029/93GL01870, 1994.

Pfeilsticker, K., Bösch, H., Camy-Peyret, C., Fitzenberger, R., Harder, H., and Osterkamp, H.: First atmospheric profile measurements of the atmospheric UV/vis O_4 absorption bands strength: Implications for the spectroscopy and the formation enthalpy of the O_2-O_2 dimer, Geophysical Research Letters, 28, 4595–4598, doi:10.1029/2001GL013734, http://dx.doi.org/10.1029/2001GL013734, 2001.

Platt, U. and Stutz, J.: Differential Optical Absorption Spectroscopy (DOAS), Principle and Applications, ISBN 3-340-21193-4, Springer Verlag, Heidelberg, 2008.

Prados-Roman, C., Butz, A., Deutschmann, T., Dorf, M., Kritten, L., Minikin, A., Platt, U., Schlager, H., Sihler, H., Theys, N., Van Roozendael, M., Wagner, T., and Pfeilsticker, K.: Airborne DOAS limb measurements of tropospheric trace gas profiles: case studies on the profile retrieval of O_4 and BrO, Atmospheric Measurement Techniques, 4, 1241–1260, doi:10.5194/amt-4-1241-2011, http://www.atmos-meas-tech.net/4/1241/2011/, 2011.

Raecke, R.: Atmospheric spectoscopy of trace gases and water vapour in the tropical tropopause layer from the NASA Global Hawk, Master thesis, Institut für Umweltphysik, Universität Heidelberg, http://www.iup.uni-heidelberg.de/institut/forschung/groups/atmosphere/stratosphere/publications/pdf/MScThesis_Rasmus_Raecke.pdf, 2013.

Rodgers, C.: Inverse methods for atmospheric sounding, World Scientific, Singapore, New Jersey, London, Hongkong, 2000.

Roeckner, E., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kornblueh, L., Manzini, E., Schlese, U., and Schulzweida, U.: Sensitivity of Simulated Climate to Horizontal and Vertical Resolution in the ECHAM5 Atmosphere Model, Journal of Climate, 19, 3771–3791, doi:10.1175/JCLI3824.1, http://dx.doi.org/10.1175/JCLI3824.1, 2006.

Rolf, C., Afchine, A., Bozem, H., Buchholz, B., Ebert, V., Guggenmoser, T., Hoor, P., Konopka, P., Kretschmer, E., Müller, S., Schlager, H., Spelten, N., Sumińska-Ebersoldt, O., Ungermann, J., Zahn, A., and Krämer, M.: Transport of Antarctic stratospheric strongly dehydrated air into the troposphere observed during the HALO-ESMVal campaign 2012, Atmospheric Chemistry and Physics, 15, 9143–9158, doi:10.5194/acp-15-9143-2015, http://www.atmos-chem-phys.net/15/9143/2015/, 2015.

Rothman, L. S., Gordon, I. E., Barbe, A., Benner, D. C., Bernath, P. F., Birk, M., Boudon, V., Brown, L. R., Campargue, A., Champion, J.-P., et al.: The HITRAN 2008 molecular spectroscopic database, Journal of Quantitative Spectroscopy and Radiative Transfer, 110, 533–572, 2009.

Saiz-Lopez, A. and von Glasow, R.: Reactive halogen chemistry in the troposphere, Chemical Society Reviews, 41, 6448–6472, 2012.

Sander, R., Baumgaertner, A., Gromov, S., Harder, H., Jöckel, P., Kerkweg, A., Kubistin, D., Regelin, E., Riede, H., Sandu, A., Taraborrelli, D., Tost, H., and Xie, Z.-Q.: The atmospheric chemistry box model CAABA/MECCA-3.0, Geoscientific Model Development, 4, 373–380, doi:10.5194/gmd-4-373-2011, https://www.geosci-model-dev.net/4/373/2011/, 2011a.

Sander, S., Friedl, R. R., Barkern, J., Golden, D., Kurylo, M., Wine, P., Abbat, J., Burkholder, J., Moortgaret, C., Huie, R., and Orkin, R. E.: Chemical kinetics and photochemical data for use in atmospheric studies, Technical Report, NASA/JPL Publication, 17, 2011b.

Scalone, L.: Retrieval of Cirrus Optical Properties in the near-IR spectral range within the NASA ATTREX Project, Phd, University of Heidelberg, Heidelberg, Germany, doi:10.11588/heidok.00023004, http://www.ub.uni-heidelberg.de/archiv/23004, 2017.

Schiller, C., Wahner, A., Platt, U., Dorn, H.-P., Callies, J., and Ehhalt, D. H.: Near UV atmospheric absorption measurements of column abundances during Airborne Arctic Stratospheric Expedition, January – February 1989: 2. OClO observations, Geophysical Research Letters, 17, 501–504, doi:10.1029/GL017i004p00501, http://dx.doi.org/10.1029/GL017i004p00501, 1990.

Schiller, C., Bozem, H., Gurk, C., Parchatka, U., Königstedt, R., Harris, G., Lelieveld, J., and Fischer, H.: Applications of quantum cascade lasers for sensitive trace gas measurements of CO, CH4, N2O and HCHO, Applied Physics B, 92, 419–430, doi:10.1007/s00340-008-3125-0, http://dx.doi.org/10.1007/s00340-008-3125-0, 2008.

Schmidt, J. A., Jacob, D. J., Horowitz, H. M., Hu, L., Sherwen, T., Evans, M. J., Liang, Q., Suleiman, R. M., Oram, D. E., Le Breton, M., Percival, C. J., Wang, S., Dix, B., and Volkamer, R.: Modeling the observed tropospheric BrO background: Importance of multiphase chemistry and implications for ozone, OH, and mercury, Journal of Geophysical Research: Atmospheres, 121, 11,819–11,835, doi:10.1002/2015JD024229, http://dx.doi.org/10.1002/2015JD024229, 2015JD024229, 2016.

Serdyuchenko, A., Gorshelev, V., Weber, M., Chehade, W., and Burrows, J.: High spectral resolution ozone absorption cross-sections–Part 2: Temperature dependence, Atmospheric Measurement Techniques, 7, 625–636, 2014.

Sherwen, T., Schmidt, J. A., Evans, M. J., Carpenter, L. J., Großmann, K., Eastham, S. D., Jacob, D. J., Dix, B., Koenig, T. K., Sinreich, R., Ortega, I., Volkamer, R., Saiz-Lopez, A., Prados-Roman, C., Mahajan, A. S., and Ordóñez, C.: Global impacts of tropospheric halogens (Cl, Br, I) on oxidants and composition in GEOS-Chem, Atmospheric Chemistry and Physics, 16, 12 239–12 271, doi:10.5194/acp-16-12239-2016, http://www.atmos-chem-phys.net/16/12239/2016/, 2016.

Sherwen, T., Evans, M. J., Carpenter, L. J., Schmidt, J. A., and Mickley, L. J.: Halogen chemistry reduces tropospheric O_3 radiative forcing, Atmospheric Chemistry and Physics, 17, 1557–1569, doi:10.5194/acp-17-1557-2017, http://www.atmos-chem-phys.net/17/1557/2017/, 2017.

Solomon, S., Mount, G. H., Sanders, R. W., and Schmeltekopf, A. L.: Visible spectroscopy at McMurdo Station, Antarctica: 2. Observations of OClO, Journal of Geophysical Research: Atmospheres, 92, 8329–8338, doi:10.1029/JD092iD07p08329, http://dx.doi.org/10.1029/JD092iD07p08329, 1987a.

Solomon, S., Schmeltekopf, A. L., and Sanders, R. W.: On the interpretation of zenith sky absorption measurements, Journal of Geophysical Research: Atmospheres, 92, 8311–8319, doi:10.1029/JD092iD07p08311, http://dx.doi.org/10.1029/JD092iD07p08311, 1987b.

Stratmann, G., Ziereis, H., Stock, P., Brenninkmeijer, C., Zahn, A., Rauthe-Schöch, A., Velthoven, P., Schlager, H., and Volz-Thomas, A.: NO and NO_y in the upper troposphere: Nine years of CARIBIC measurements onboard a passenger aircraft, Atmospheric Environment, 133, 93 – 111, doi:http://dx.doi.org/10.1016/j.atmosenv.2016.02.035, http://www.sciencedirect.com/science/article/pii/S1352231016301480, 2016.

Stutz, J., Kim, E., Platt, U., Bruno, P., Perrino, C., and Febo, A.: UV-visible absorption cross sections of nitrous acid, Journal of geophysical research, 105, e14 592, 2000.

Stutz, J., Werner, B., Spolaor, M., Scalone, L., Festa, J., Tsai, C., Cheung, R., Colosimo, S. F., Tricoli, U., Raecke, R., Hossaini, R., Chipperfield, M. P., Feng, W., Gao, R.-S., Hintsa, E. J., Elkins, J. W., Moore, F. L., Daube, B., Pittman, J., Wofsy, S., and Pfeilsticker, K.: A new Differential Optical Absorption Spectroscopy instrument to study atmospheric chemistry from a high-altitude unmanned aircraft, Atmospheric Measurement Techniques, 10, 1017–1042, doi:10.5194/amt-10-1017-2017, http://www.atmos-meas-tech.net/10/1017/2017/, 2017.

Thalman, R. and Volkamer, R.: Temperature dependent absorption cross-sections of O2-O2 collision pairs between 340 and 630 nm and at atmospherically relevant pressure, Phys. Chem. Chem. Phys., 15, 15 371–15 381, doi:10.1039/C3CP50968K, http://dx.doi.org/10.1039/C3CP50968K, 2013.

Van Roozendael, M., Wagner, T., Richter, A., Pundt, I., Arlander, D., Burrows, J., Chipperfield, M. P., Fayt, C., Johnston, P., Lambert, J., Kreher, K., Pfeilsticker, K., Platt, U., Pommereau, J. P., Sinnhuber, B., Tornkvist, K., and Wittrock, F.: Intercomparison of BrO Measurements from ERS-2 GOME, Ground-based and balloon platforms, Adv. Space. Res., 29, 1161 – 1666, 2002.

Vogel, B., Günther, G., Müller, R., Grooß, J.-U., Hoor, P., Krämer, M., Müller, S., Zahn, A., and Riese, M.: Fast transport from Southeast Asia boundary layer sources to northern Europe: rapid uplift in typhoons and eastward eddy shedding of the Asian monsoon anticyclone, Atmospheric Chemistry and Physics, 14, 12 745–12 762, doi:10.5194/acp-14-12745-2014, http://www.atmos-chem-phys.net/14/12745/2014/, 2014.

Vogel, B., Günther, G., Müller, R., Grooß, J.-U., and Riese, M.: Impact of different Asian source regions on the composition of the Asian monsoon anticyclone and of the extratropical lowermost stratosphere, Atmospheric Chemistry and Physics, 15, 13 699–13 716, doi:10.5194/acp-15-13699-2015, http://www.atmos-chem-phys.net/15/13699/2015/, 2015.

Voigt, C., Schumann, U., Minikin, A., Abdelmonem, A., Afchine, A., Borrmann, S., Boettcher, M., Buchholz, B., Bugliaro, L., Costa, A., Curtius, J., Dollner, M., Dörnbrack, A., Dreiling, V., Ebert, V., Ehrlich, A., Fix, A., Forster, L., Frank, F., Fütterer, D., Giez, A., Graf, K., Grooß, J.-U., Groß, S., Heimerl, K., Heinold, B., Hüneke, T., Järvinen, E., Jurkat, T., Kaufmann, S., Kenntner, M., Klingebiel, M., Klimach, T., Kohl, R., Krämer, M., Krisna, T. C., Luebke, A., Mayer, B., Mertes, S., Molleker, S., Petzold, A., Pfeilsticker, K., Port, M., Rapp, M., Reutter, P., Rolf, C., Rose, D., Sauer, D., Schäfler, A., Schlage, R., Schnaiter, M., Schneider, J., Spelten, N., Spichtinger, P., Stock, P., Walser, A., Weigel, R., Weinzierl, B., Wendisch, M., Werner, F., Wernli, H., Wirth, M., Zahn, A., Ziereis, H., and Zöger, M.: ML-CIRRUS - The airborne experiment on natural cirrus and contrail cirrus with the high-altitude long-range research aircraft HALO, Bulletin of the American Meteorological Society, 0, null, doi:10.1175/BAMS-D-15-00213.1, http://dx.doi.org/10.1175/BAMS-D-15-00213.1, 2016.

Volkamer, R., Baidar, S., Campos, T. L., Coburn, S., DiGangi, J. P., Dix, B., Eloranta, E. W., Koenig, T. K., Morley, B., Ortega, I., Pierce, B. R., Reeves, M., Sinreich, R., Wang, S., Zondlo, M. A., and Romashkin, P. A.: Aircraft measurements of BrO, IO, glyoxal, NO_2, H_2O, O_2–O_2 and aerosol extinction profiles in the tropics: comparison with aircraft-/ship-based in situ and lidar measurements, Atmospheric Measurement Techniques, 8, 2121–2148, doi:10.5194/amt-8-2121-2015, http://www.atmos-meas-tech.net/8/2121/2015/, 2015.

von Glasow, R., von Kuhlmann, R., . Lawrence, M. G., , Platt, U., and Crutzen, P. J.: Impact of reactive bromine chemistry in the troposphere, Atmos. Chem. Phys., 4, 2481–2497, 2004.

Wagner, T., Beirle, S., and Deutschmann, T.: Three-dimensional simulation of the Ring effect in observations of scattered sun light using Monte Carlo radiative transfer models, Atm. Meas. Tech, 2, 113–124, 2009.

Wahner, A., Callies, J., Dorn, H.-P., Platt, U., and Schiller, C.: Near UV atmospheric absorption measurements of column abundances during Airborne Arctic Stratospheric Expedition, January – February 1989: 1. Technique and NO2 observations, Geophysical Research Letters, 17, 497–500, doi:10.1029/GL017i004p00497, http://dx.doi.org/10.1029/GL017i004p00497, 1990a.

Wahner, A., Callies, J., Dorn, H.-P., Platt, U., and Schiller, C.: Near UV atmospheric absorption measurements of column abundances during Airborne Arctic Stratospheric Expedition, January – February 1989: 3. BrO observations, Geophysical Research Letters, 17, 517–520, doi:10.1029/GL017i004p00517, http://dx.doi.org/10.1029/GL017i004p00517, 1990b.

Wang, S., Schmidt, J. A., Baidar, S., Coburn, S., Dix, B., Koenig, T. K., Apel, E., Bowdalo, D., Campos, T. L., Eloranta, E., Evans, M. J., DiGangi, J. P., Zondlo, M. A., Gao, R.-S., Haggerty, J. A., Hall, S. R., Hornbrook, R. S., Jacob, D., Morley, B., Pierce, B., Reeves, M., Romashkin, P., ter Schure, A., and Volkamer, R.: Active and widespread halogen chemistry in the tropical and subtropical free troposphere, Proceedings of the National Academy of Sciences, 112, 9281–9286, doi:10.1073/pnas.1505142112, http://www.pnas.org/content/112/30/9281.abstract, 2015.

Weidner, F., Bösch, H., Bovensmann, H., Burrows, J. P., Butz, A., Camy-Peyret, C., Dorf, M., Gerilowski, K., Gurlit, W., Platt, U., von Friedeburg, C., Wagner, T., and Pfeilsticker, K.: Balloon-borne limb profiling of UV/vis skylight radiances, O_3, NO_2, and BrO: technical set-up and validation of the method, Atmospheric Chemistry and Physics, 5, 1409–1422, doi:10.5194/acp-5-1409-2005, http://www.atmos-chem-phys.net/5/1409/2005/, 2005.

Wendisch, M., Pöschl, U., Andreae, M. O., Machado, L. A. T., Albrecht, R., Schlager, H., Rosenfeld, D., Martin, S. T., Abdelmonem, A., Afchine, A., Araùjo, A., Artaxo, P., Aufmhoff, H., Barbosa, H. M. J., Borrmann, S., Braga, R., Buchholz, B., Cecchini, M. A., Costa, A., Curtius, J., Dollner, M., Dorf, M., Dreiling, V., Ebert, V., Ehrlich, A., Ewald, F., Fisch, G., Fix, A., Frank, F., Fütterer, D., Heckl, C., Heidelberg, F., Hüneke, T., Jäkel, E., Järvinen, E., Jurkat, T., Kanter, S., Kästner, U., Kenntner, M., Kesselmeier, J., Klimach, T., Knecht, M., Kohl, R., Kölling, T., Krämer, M., Krüger, M., Krisna, T. C., Lavric, J. V., Longo, K., Mahnke, C., Manzi, A. O., Mayer, B., Mertes, S., Minikin, A., Molleker, S., Münch, S., Nillius, B., Pfeilsticker, K., Pöhlker, C., Roiger, A., Rose, D., Rosenow, D., Sauer, D., Schnaiter, M., Schneider, J., Schulz, C., de Souza, R. A. F., Spanu, A., Stock, P., Vila, D., Voigt, C., Walser, A., Walter, D., Weigel, R., Weinzierl, B., Werner, F., Yamasoe, M. A., Ziereis, H., Zinner, T., and Zöger, M.: The ACRIDICON-CHUVA campaign: Studying tropical deep convective clouds and precipitation over Amazonia using the new German research aircraft HALO, Bulletin of the American Meteorological Society, 0, null, doi:10.1175/BAMS-D-14-00255.1, http://dx.doi.org/10.1175/BAMS-D-14-00255.1, 2016.

Werner, B.: Spectroscopic UV/vis limb measurements from aboard the NASA Global Hawk: Implications for the photochemistry and budget of bromine in the tropical tropopause layer, Phd, University of Heidelberg, Heidelberg, Germany, 2015.

Werner, B., Stutz, J., Spolaor, M., Scalone, L., Raecke, R., Festa, J., Colosimo, S. F., Cheung, R., Tsai, C., Hossaini, R., Chipperfield, M. P., Taverna, G. S., Feng, W., Elkins, J. W., Fahey, D. W., Gao, R.-S., Hintsa, E. J., Thornberry, T. D., Moore, F. L., Navarro, M. A., Atlas, E., Daube, B. C., Pittman, J., Wofsy, S., and Pfeilsticker, K.: Probing the subtropical lowermost stratosphere and the tropical upper troposphere and tropopause layer for inorganic bromine, Atmospheric Chemistry and Physics, 17, 1161–1186, doi:10.5194/acp-17-1161-2017, http://www.atmos-chem-phys.net/17/1161/2017/, 2017.

Wolf, K., Ehrlich, A., Hüneke, T., Pfeilsticker, K., Werner, F., Wirth, M., and Wendisch, M.: Potential of remote sensing of cirrus optical thickness by airborne spectral radiance measurements at different sideward viewing angles, Atmospheric Chemistry and Physics, 17, 4283–4303, doi:10.5194/acp-17-4283-2017, http://www.atmos-chem-phys.net/17/4283/2017/, 2017.

Ye, C., Zhou, X., Pu, D., Stutz Jochen, Festa, J., Spolaor, M., Tsai, C., Cantrell, C., Mauldin, R. L., Campos, T., Weinheimer, A., Hornbrook, R. S., Apel, E. C., Guenther, A., Kaser, L., Yuan, B., Karl, T., Haggerty, J., Hall, S., Ullmann, K., Smith, J. N., Ortega, J., and Knote, C.: Rapid cycling of reactive nitrogen in the marine boundary layer, Nature, 532, 489–491, doi:doi:10.1038/nature17195, 2016.

Zahn, A., Weppner, J., Widmann, H., Schlote-Holubek, K., Burger, B., Kühner, T., and Franke, H.: A fast and precise chemiluminescence ozone detector for eddy flux and airborne application, Atmospheric Measurement Techniques, 5, 363–375, doi:10.5194/amt-5-363-2012, http://www.atmos-meas-tech.net/5/363/2012/, 2012.

Ziereis, H., Schlager, H., Schulte, P., Velthoven, P. v., and Slemr, F.: Distributions of NO, NO x, and NO y in the upper troposphere and lower stratosphere between 28 and 61 N during POLINAT 2, Journal of Geophysical Research: Atmospheres, 105, 3653–3664, 2000.

Table 1. Science mission deployments of the HALO mini-DOAS instrument from 2012 through 2016.

Date	Name		Flights	Hours	References
Aug.-Sept. 2012	TACTS / ESMVal	Transport and Composition in the UT/LS / Earth System Model Validation	11	112	Jurkat et al. (2014, 2017); Vogel et al. (2014); Rol
Dec.-Jan. 2013/14	NARVAL	Next Generation Remote Sensing for Validation Studies	15	120	Klepp et al. (2014)
Mar.-Apr. 2014	ML-Cirrus	Mid-Latitude Cirrus	16	77	Voigt et al. (2016); Wolf et al. (2017)
Aug.-Oct. 2014	ACRIDICON	Aerosol, Cloud, Precipitation, and Radiation Interactions and Dynamics of Convective Cloud Systems	14	96	Wendisch et al. (2016)
July-Aug. 2015	OMO	Oxidation Mechanism Observations	20	116	
Dec.-Mar. 2015/16	POLSTRACC	Polar Stratosphere in a Changing Climate	18	156	

Table 2. Optical specification of the mini-DOAS instrument

Channel name	UV	VIS	NIR
Telescope focal length and f-number		30 mm, f/2.5	
Telescope lense coating	UV-AR	VIS 0°	NIR II
Telescope Schott filter type	BG3	GG400	RG850
Number of fibers and diameter		7×200 μm	2×400 μm
Fibre bundle entrance slit dimension		1652 μm $\times$ 200 μm	884 μm $\times$ 400 μm
$\mathrm{FOV}_{\mathrm{opt}}$ $2 \cdot \gamma$		$3.15° \times 0.38°$	$1.68° \times 0.76°$
Spectrometer entrance slit dimension		1500 μm $\times$ 100 μm	500 μm $\times$ 100 μm
Spectrometer focal length and f-number		60 mm, f/4	
Grating [grooves/mm]	2100	1300	300
Sensor name		S10141-1107S	G9204-512
Sensor type		Si-CCD	InGaAs-PDA
Number of channels on sensor		2048	512
Sensor area per channel		12×1464 μm^2	15×500 μm^2
Full Well Capacity		$2 \cdot 10^5$ e^-	$1.87 \cdot 10^8$ e^-
Quantum Efficiency (i)	0.60	0.85	0.80
Covered wavelength range	310 - 440 nm	420 - 640 nm	1100 - 1680 nm
Resolution (slit function FWHM)	0.47 nm / 6.1 px	1.1 nm / 8.4 px	$\approx$ 10 nm / $\approx$ 11 px

(i) corresponding to the wavelengths of 400 nm (UV), 540 nm (VIS) bzw. 1500 nm (NIR).

Table 3. Details of the spectral analysis of various trace gases

Target gas	λ **(nm)**	**Fitted absorbers**	**Add. Param.**	**Polyn.**	$\sigma(\mathrm{dSCD})$
O_4	350 – 370	1, 2, 3, 5, 7, 9	I_{Ofs}(i), R (ii), $R \cdot \lambda^4$	2	
	460 – 490	1, 2, 4, 6	I_{Ofs}, R, $R \cdot \lambda^4$	2	5×10^{41}
O_3	335 – 362	1, 2, 4, 7, 9	I_{Ofs}, R, $R \cdot \lambda^4$	2	
	450 – 500	1, 2, 4, 6	I_{Ofs}, R, $R \cdot \lambda^4$	2	4×10^{18}
NO_2	407 – 435	1, 2, 3, 4, 6, 10	I_{Ofs}, R, $R \cdot \lambda^4$	2	
	424 – 490	1, 2/3, 4/5, 6	I_{Ofs}, R, $R \cdot \lambda^4$	2	2×10^{15}
H_2O	490 – 520	1, 2, 5, 6	I_{Ofs}, R, $R \cdot \lambda^4$	2	
HCHO	323 – 357	1, 2, 3, 5, 7, 8, 9	I_{Ofs}, R, $R \cdot \lambda^4$	2	7×10^{15}
HONO	337 – 372	1, 2, 3, 4, 7, 8, 9	I_{Ofs}, R, $R \cdot \lambda^4$	2	
BrO	342 – 363	1, 2, 3, 4, 7, 9	I_{Ofs}, R, $R \cdot \lambda^4$	2	2×10^{13}
OClO	353 – 392	1, 2, 3, 4, 10	I_{Ofs}, R, $R \cdot \lambda^4$	2	3×10^{13}

(i) I_{Ofs}: Offset spectrum; (ii) R: Ring spectrum; (iii) $R \cdot \lambda^4$: Ring spectrum multiplied with λ^4.

Table 4. Trace gas absorptions cross sections used for the DOAS retrieval.

No.	Absorber	Temp.	Reference	Uncertainty
1	O_4	293 K	Thalman and Volkamer (2013)	4%
2	O_3	223 K	Gorshelev et al. (2014); Serdyuchenko et al. (2014)	3%
3	O_3	293 K	Gorshelev et al. (2014); Serdyuchenko et al. (2014)	3%
4	NO_2	223 K	Bogumil et al. (2003)	3.4%
5	NO_2	293 K	Bogumil et al. (2003)	3.4%
6	H_2O	273 K	Rothman et al. (2009)	
7	HCHO	293 K	Chance and Orphal (2011)	5%
8	HONO	298 K	Stutz et al. (2000)	5%
9	BrO	223 K	Fleischmann et al. (2004)	10%
10	OClO	213 K	Kromminga et al. (2003)	5%

Table 5. Summary of random errors as discussed in section 3.7.1. The percentages in columns three and four refer to deviations of the parameter in the first column.

Parameter	Cause of the error	Typical value	Maximum value
$\Delta\alpha_R$	RTM noise	3.5%	3.5%
	Mie scattering	10%	15%
	small scale variability	0 – 20%	100%
	vertical sampling	0 – 10%	60%
$\Delta\mathrm{SCD}_R$	DOAS fit error	5%	100%
	cross section	3%	6%
	$\mathrm{SCD}_{\mathrm{ref}}$	5 – 10%	20%
$\Delta[X]$	O_3 measurement	<1%	1%
	O_4 calculation	1%	1%

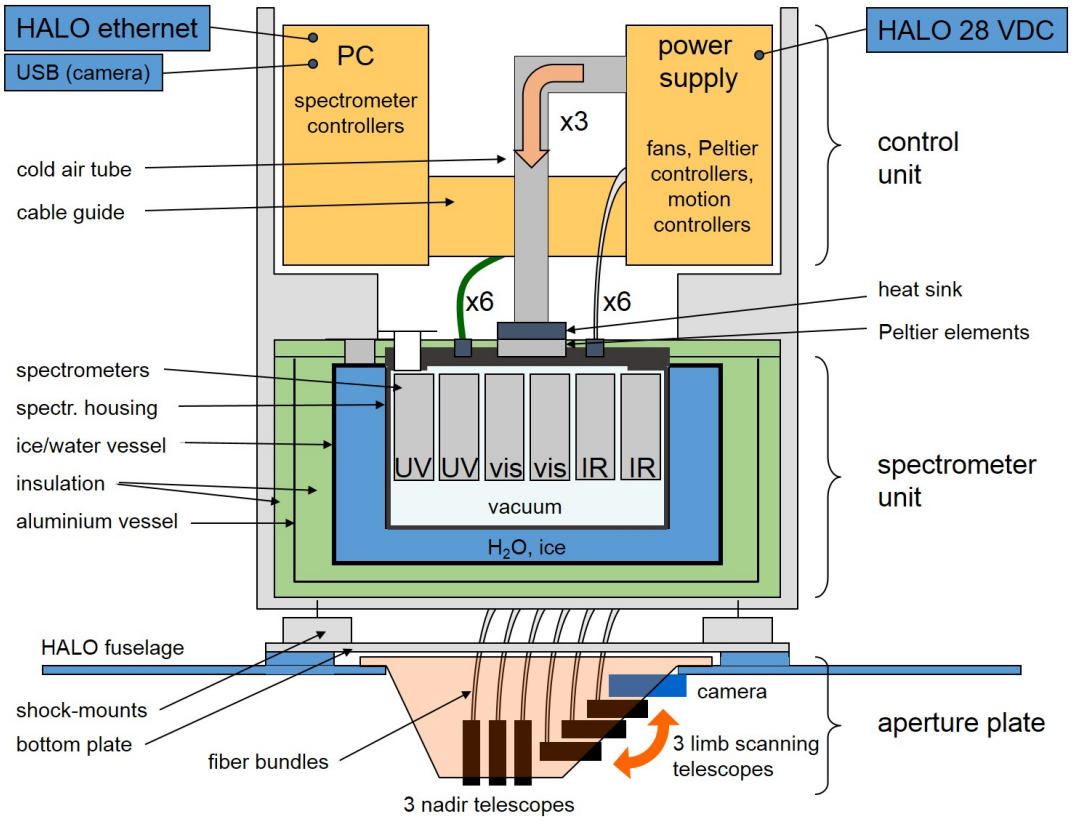

Figure 1. Sketch of the HALO mini-DOAS instrument.

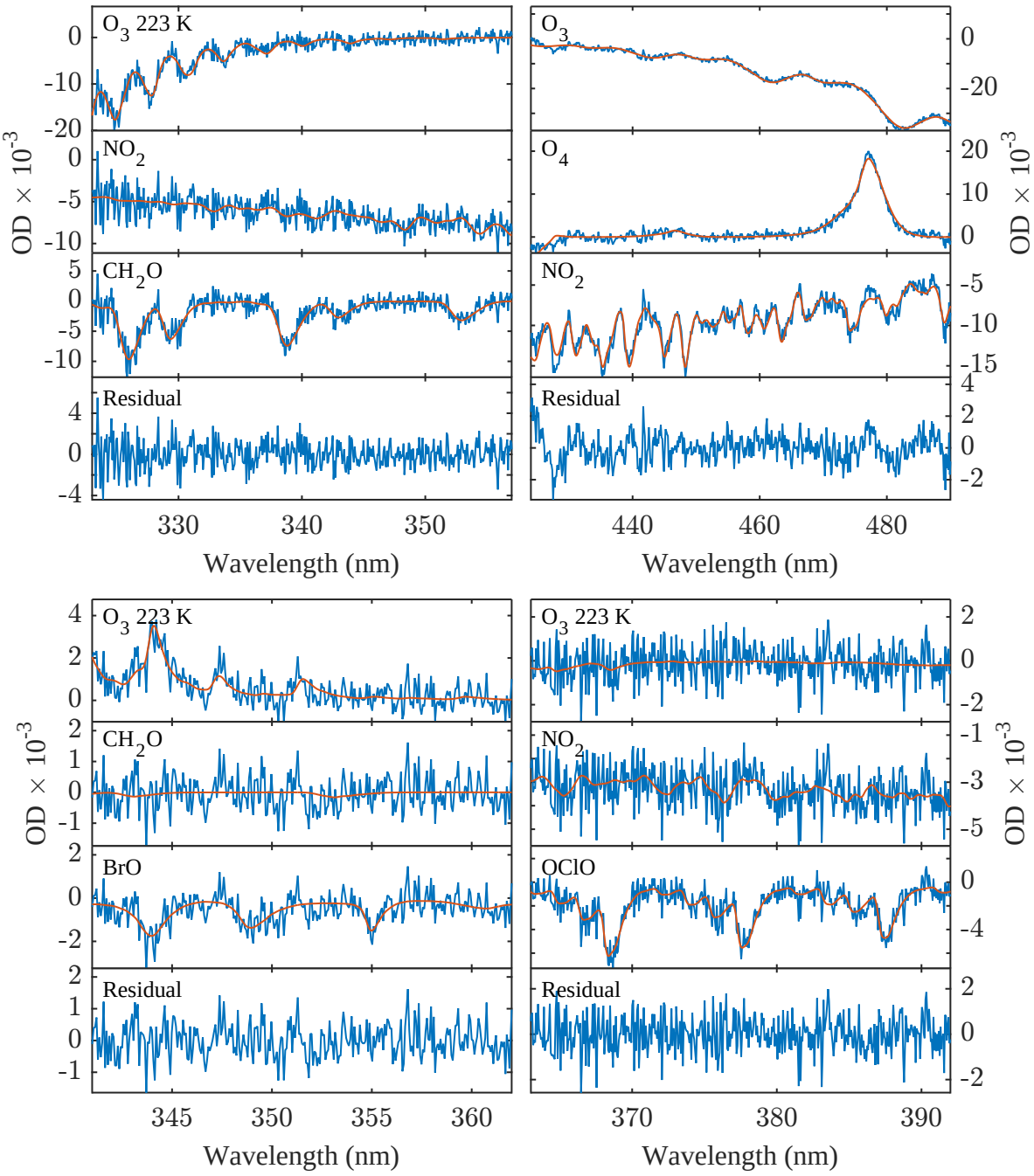

Figure 2. Sample spectral retrievals as described in section 3.1. ~~Tob~~ Top left: CH_2O (1.4×10^{17} molec./cm^2) retrieval in the UV spectral range (ACRIDICON, 16 September 2014). Top right: NO_2 (2×10^{16} molec./cm^2) retrieval in the visible spectral range (ESMVal, 13 September 2012). Bottom left: BrO (4×10^{14} molec./cm^2) retrieval in the UV. Bottom right: OClO (5.6×10^{14} molec./cm^2) in the UV (both POLSTRACC, 31 January 2016). Fitted reference absorption cross sections are shown in red and the residual structures are shown in blue.

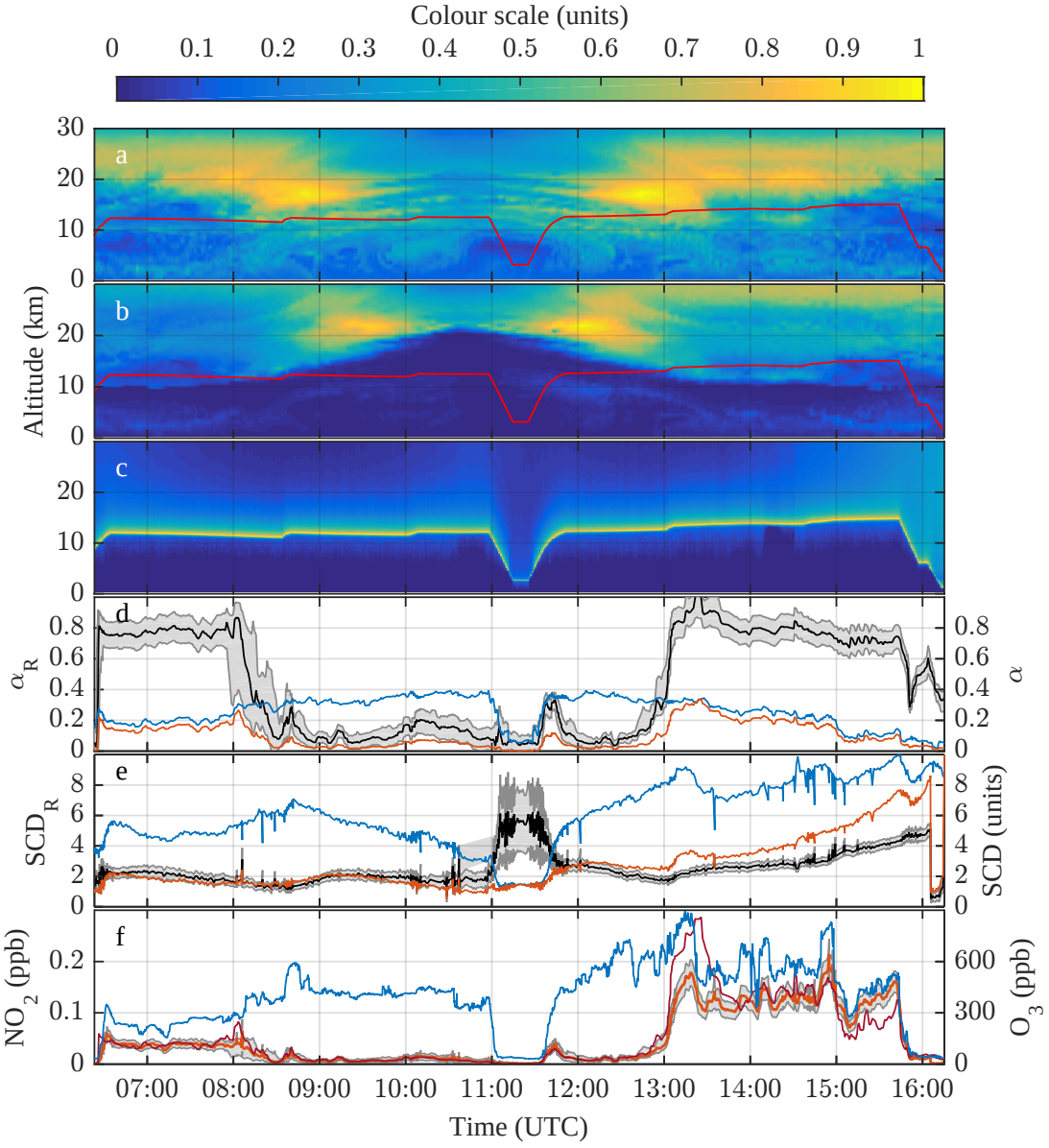

Figure 3. Illustration of NO_2 mixing ratio retrieval for the ESMVal flight on 13 September 2012 using the CTM CLaMS. Panel a: CLaMS predicted $[O_3]$ curtain (colour scale $\times\ 7.9 \cdot 10^{12}\ \mathrm{cm}^{-3}$) and aircraft altitude (red line). Panel b: CLaMS predicted $[NO_2]$ curtain (colour scale $\times\ 2.9 \cdot 10^{9}\ \mathrm{cm}^{-3}$) and aircraft altitude (red line). Panel c: BoxAMFs calculated by the RTM McArtim (colour scale $\times \log(217)$). Panel d: Calculated α_{O_3} (blue) and α_{NO_2} (red) as well as α_R (black line) and its uncertainty range (grey shaded area). Panel e: Retrieved SCD_{O_3} (blue, scale divided by 9.0×10^{20}) and SCD_{NO_2} (red, scale divided by 3.4×10^{17}) as well as SCD_R (black line, scale multiplied by 10^4) and its uncertainty range (grey shaded area). Panel f: Retrieved $[NO_2]_{O_3}$ (light red line) and its uncertainty range (grey shaded area) together with in situ measured O_3 (blue line). The dark red line shows the NO_2 mixing ratios predicted by the CLaMS model.

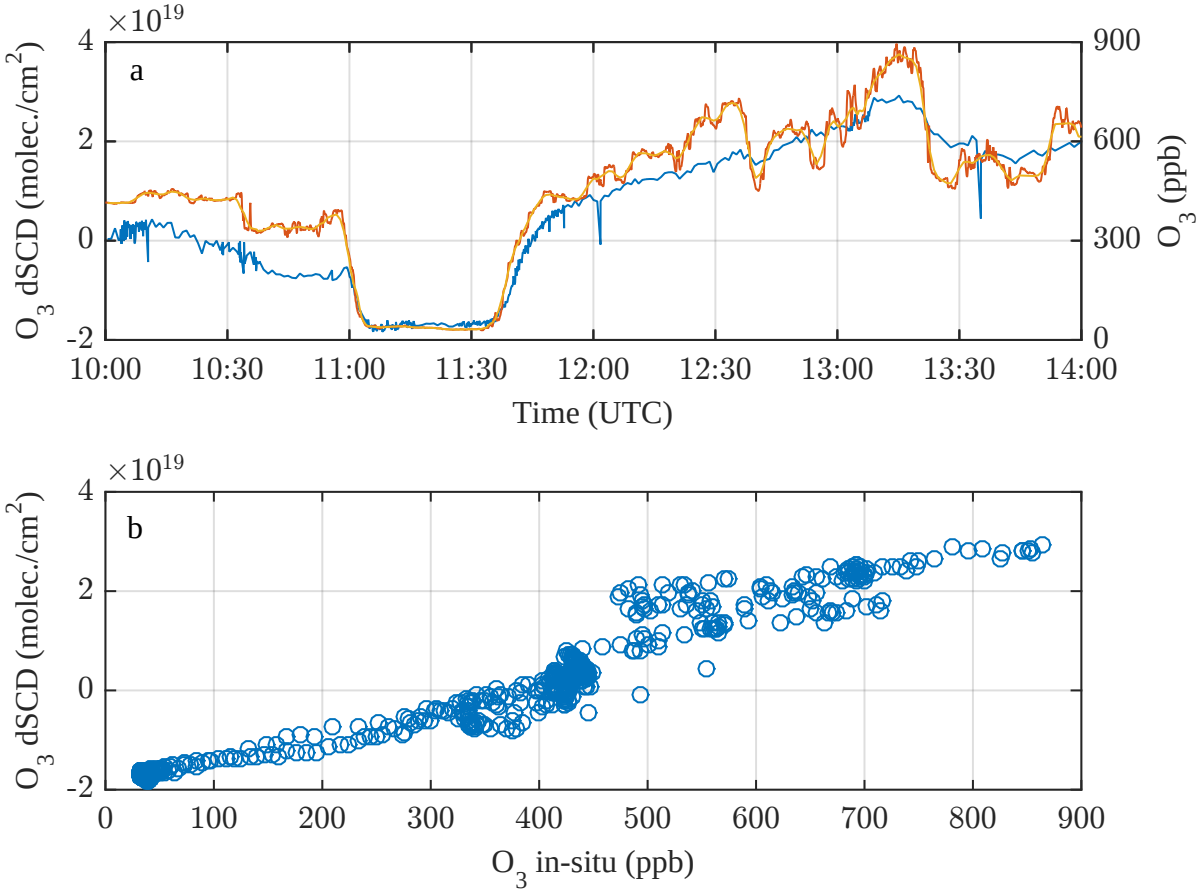

Figure 4. Comparison of in situ measured and remotely sensed O_3. Panel a: Time series of high time resolution (red line) and five minutes running average (orange line) of O_3 measured by the ~~Fairo~~ FAIRO instrument, and remotely sensed O_3 (blue line) for a segment of the HALO flight from Cape Town to Antarctica on 13 September 2012. Panel b: Scatterplot of averaged in situ measurements and remotely sensed O_3 for the flight segment shown in panel a.

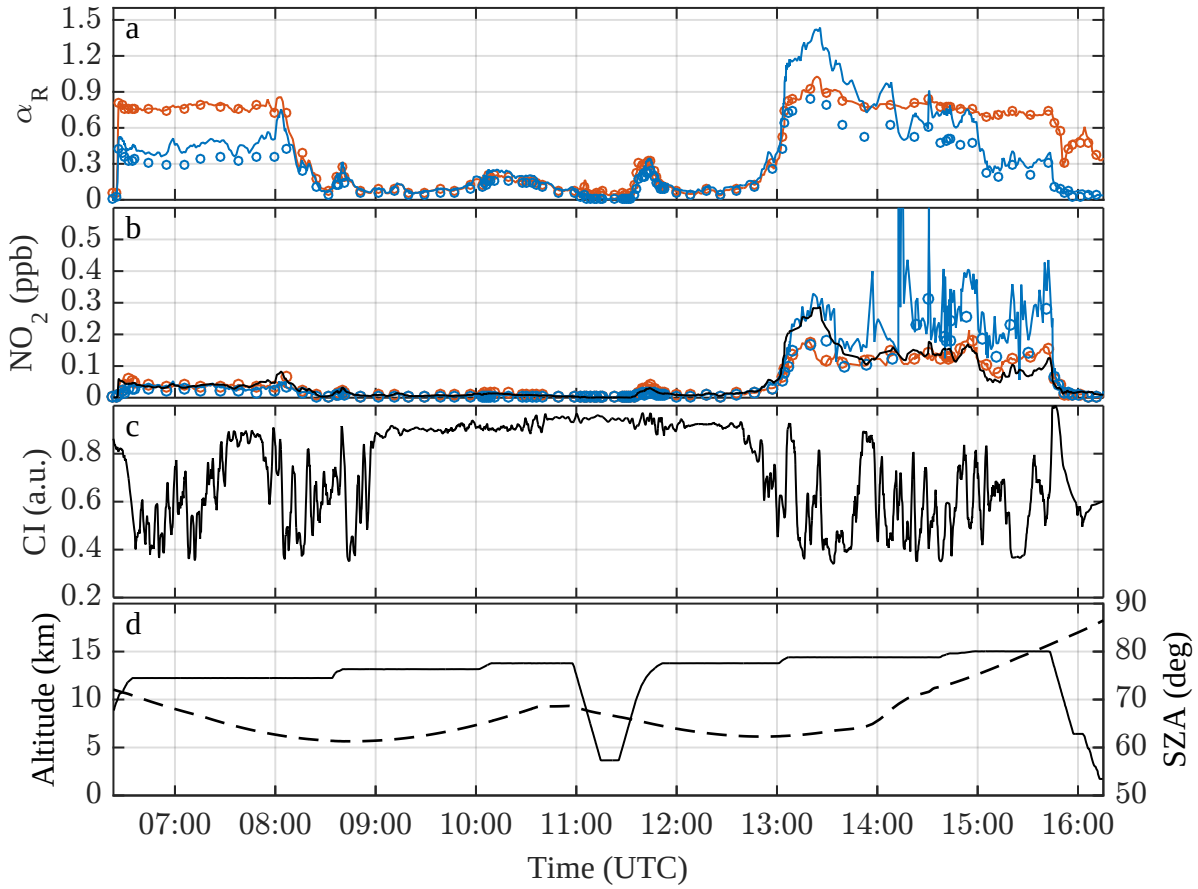

Figure 5. Retrieved $[NO_2]_{O_3}$ (red) and $[NO_2]_{O_4}$ (blue) for the ESMVal research flight on 13 September 2012. Calculations assuming clear skies are displayed as lines, calculations including a cloud layer at 4 – 8 km are displayed as circles. Panel a: Timeseries of calculated α_R. Panel b: Timeseries of inferred $[NO_2]$ together with NO_2 concentrations as predicted by CLaMs (black line). Panel c: Colour index (CI, 600 nm / 430 nm radiances) observed by the VIS3 channel in nadir geometry. A large/small colour index indicates a cloud cover/clear sky below the aircraft, respectively. Panel d: Pressure altitude of HALO (black line) and solar zenith angle (SZA, black dashed line).

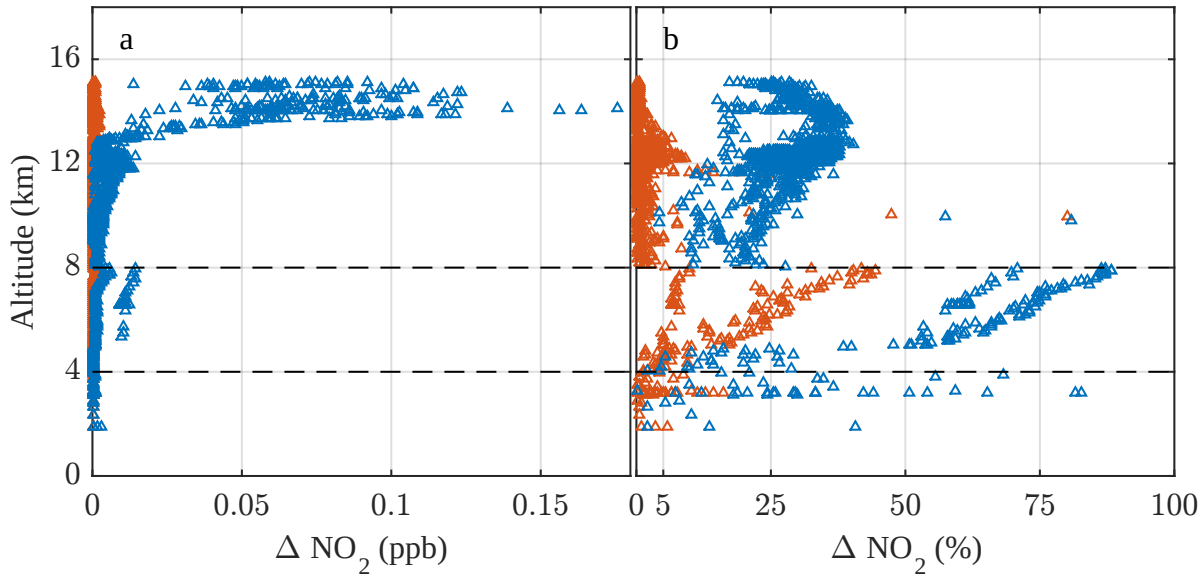

Figure 6. Impact of a cloud layer on retrieved $[NO_2]_{O_3}$ (red) and $[NO_2]_{O_4}$ (blue) for the ESMVal research flight on 13 September 2012. Shown are altitude profiles of the difference $\Delta[NO_2] = |[NO_2]_{clear} - [NO_2]_{clouded}|$ of the clear sky and clouded sky calculations, calculated from the data shown in Figure 5, panel b. The altitude range of the cloud layer as encountered during the dive and implemented in the clouded sky calculations is indicated by dashed lines.

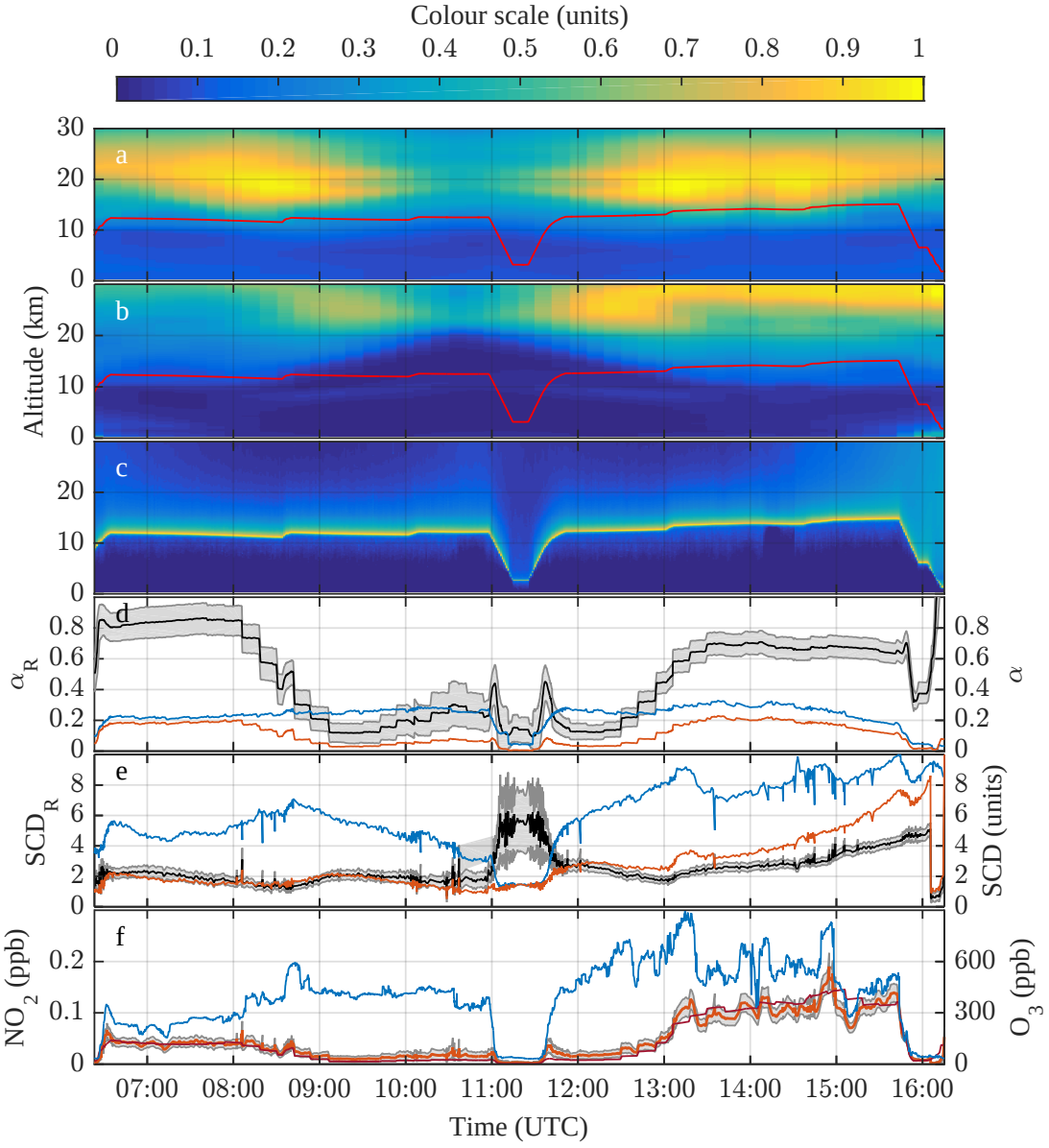

Figure 7. Illustration of NO_2 mixing ratio retrieval for the ESMVal flight on 13 September 2012 using the CCM EMAC. Panel a: EMAC predicted $[O_3]$ curtain (colour scale $\times\ 7.9 \cdot 10^{12}$ cm^{-3}) and aircraft altitude (red line). Panel b: EMAC predicted NO_2 curtain (colour scale $\times\ 2.9 \cdot 10^{9}$ cm^{-3}) and aircraft altitude (red line). Panel c: Box-AMFs calculated by the RTM McArtim (colour scale $\times \log(217)$). Panel d: Calculated α_{O_3} (blue) and α_{NO_2} (red) as well as α_R (black line) and its uncertainty range (grey shaded area). Panel e: Retrieved SCD_{O_3} (blue, scale divided by 9.0×10^{20}) and SCD_{NO_2} (red, scale divided by 3.4×10^{17}) as well as SCD_R (black line, scale multiplied by 10^4) and its uncertainty range (grey shaded area). Panel f: Retrieved $[NO_2]_{O_3}$ (light red line) and its uncertainty range (grey shaded area) together with in situ measured O_3 (blue line). The dark red line shows the NO_2 mixing ratio predicted by the EMAC model.

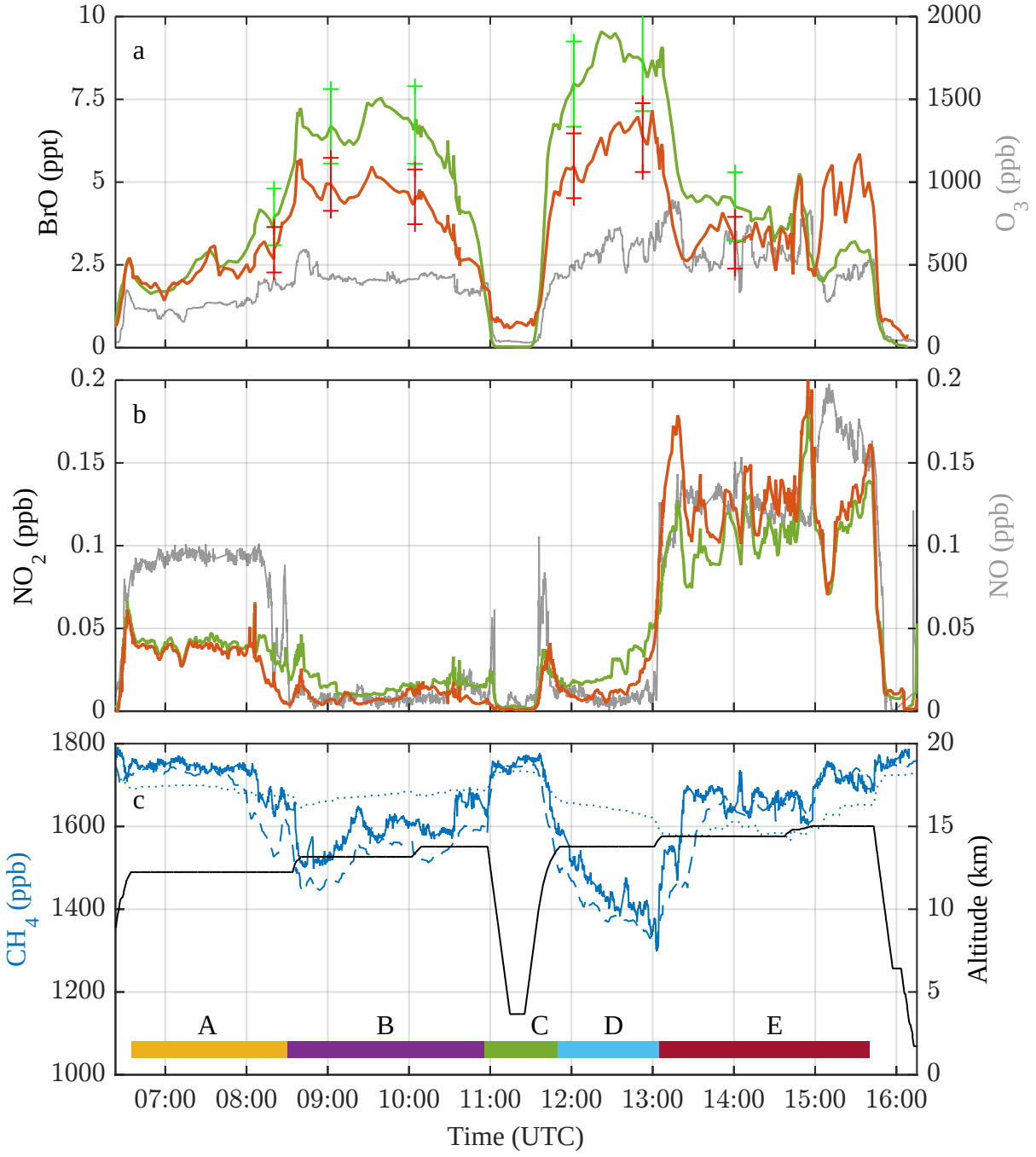

Figure 8. Time series of measured trace gas mixing ratios recorded during the ESMVal research flight on 13 September 2012. Panel a: In situ measured concentration of O_3 (grey) and inferred $[BrO]_{O_3}$ using profile shape predictions by ClaMS (red) and EMAC (green). Panel b: In situ measured NO (grey) and inferred $[NO_2]_{O_3}$ using profile shape predictions by CLaMS (red) and EMAC (green). The uncertainties are discussed in section 3.7.1. Panel c: Pressure altitude of HALO (black) and CH_4 mixing ratios (blue), the latter as derived from in situ measurements by TRIHOP (continuous line), CLaMS prediction (dashed line), EMAC prediction (dotted line). Additionally, flight sections A through E are marked for reference in the text.

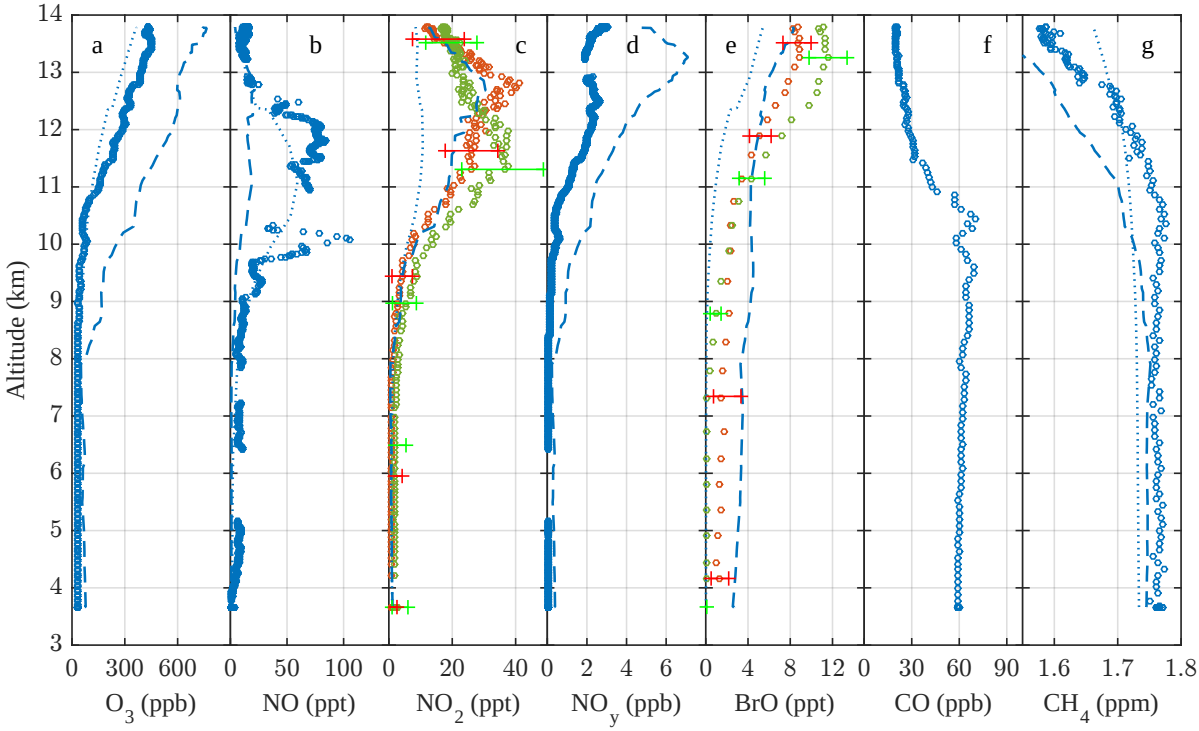

Figure 9. Altitude profiles of trace gas concentrations measured during the ascent (65°S, 21°E) of flight section C of the ESMVal research flight on 13 September 2012. Panel a: Measurements of O_3 (~~Fairo~~FAIRO), panel b: NO (AENEAS), panel c: NO_2 (mini-DOAS), NO_y (AENEAS, panel d), BrO(mini-DOAS instrument, panel e), CO (TRIHOP, panel f), and CH_4 (TRIHOP, panel g) are shown. Panels a, b, d, f, g: In situ measurements are indicated as blue circles, CLaMS predictions as dashed lines, and EMAC preidictions as dotted lines. Panels c, e: Inferred $[NO_2]_{O_3}$ and $[BrO]_{O_3}$ (circles), respectively, with random errors (error bars) using profile shapes predicted by CLaMS (red symbols) and EMAC (green symbols).

Supporting Information for the paper entitled "The novel HALO mini-DOAS instrument: Inferring trace gas concentrations from air-borne UV/visible limb spectroscopy under all skies using the scaling method"

September 15, 2017

Contents:

1 Telescope pointing precision and accuracy

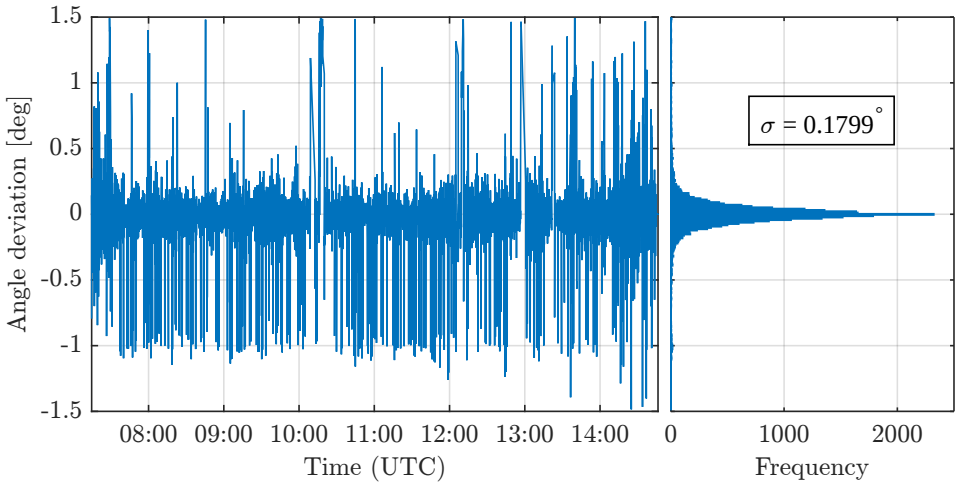

Figure 1: Deviation of the targeted elevation angle - taken from the aircraft's INS system - and the actual elevation angle for the HALO research flight of the TACTS mission on 30 August 2012.

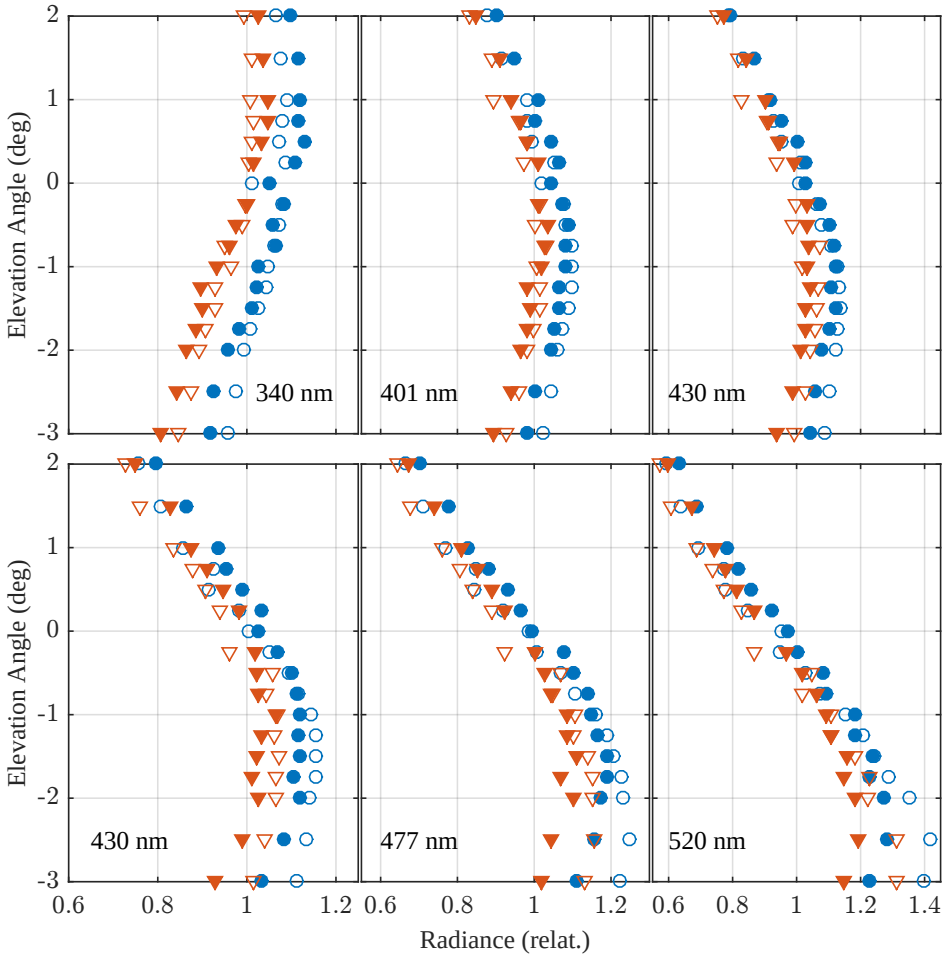

Figure 2: Comparison of modelled (filled symbols) and measured (unfilled symbols) relative radiances for a set of elevation angles between +2° and -3°, recorded by the UV2 (upper panels) and VIS4 (lower panels) channels. Blue and red symbols indicate two subsequently recorded limb scans. Models and measurements are normalised to the mean of the respective dataset. The simulations employ the radiative transfer model (RTM) McArtim and climatological background aerosol profiles from satellite measurements. The measurements were carried out during the Polstracc science mission on 26 February 2016 at 13.7 km altitude, SZA 80° ... 82°, SRAA 67° ... 69°, and at 70°N, 47°W.

2 α_R random errors

This section gives further details on the contributions to the random error $\Delta\alpha_R$, as described in section 3.7.1 of the paper.

2.1 Uncertainty of α_R due to scattering by aerosols and clouds

This section describes the RT model simulations to determine the uncertainty of α_R due to scattering by aerosols and clouds (section 3.7.1, $\Delta\alpha_R$, paragraph (a)).

For the cloud cover, the following is assumed (Figure 3): (a) A deck of marine strato-cumulus with a cloud base at 0.6 km and a cloud top at 1.2 km, a scattering coefficient of 20/km, an asymmetry factor (g) of g = 0.85, (b) Cb clouds with a cloud base at 0.6 km and a cloud top at 12.5 km, a scattering coefficient of 40/km and a g = 0.85 (in the lower part) and g = 0.7 in upper part, and (c) cirrus clouds (Ci) with a cloud base at 12 km and a cloud top at 12.5 km, a scattering coefficient of 1/km and a g = 0.7. For all cloud types a single scattering albedo of $\omega_0 = 0.999$ is taken (for further details see Knecht (2015), section 4) and the clouds are assumed to be internally homogenous. This is justified to due the radiative smoothing within the cloud and our interest in modelling the disturbance of the radiative field and its effects on the ratio of the measured absorption for the scaling gas and target gas, in particular for the cloud free part of the atmosphere (Marshak and Davis, 2005). Further for the simulations the ground albedo of Pinker (1982) is taken and profiles of O_4 and CH_2O as shown in the insert of Figure 4 are assumed. Finally the clouds are arranged according to Figure 3. Then the positions and orientations of the limb telescope are randomly located in the surrogate cloud field.

For both gases the expected α-factors are then simulated at 343 nm using the RT model McArtim (Deutschmann et al., 2011) for solar zenith angles of 14°(given in blue in Figure 4) and 50°(given in red in Figure 4). The frequency distributions show calculations based on the cloud scenario described above, and additionally the same calculations are carried out assuming clear skies (blue and red lines in Figure 4).

The calculations were performed in support for the interpretation of the measurement taken within the framework of the ACRIDICON campaign (Wendisch et al., 2016) which took place in the Amazon basin in fall 2014. There strong convection lead to cumulus and cumulus nimbus clouds which are largely varied in vertical and horizontal extent, providing an extreme test case to determine the sensitivity of the method as a function of cloud cover. Similar simulations were carried out for a different cloud scenario at 477 nm, yielding similar results (see Knecht, 2015).

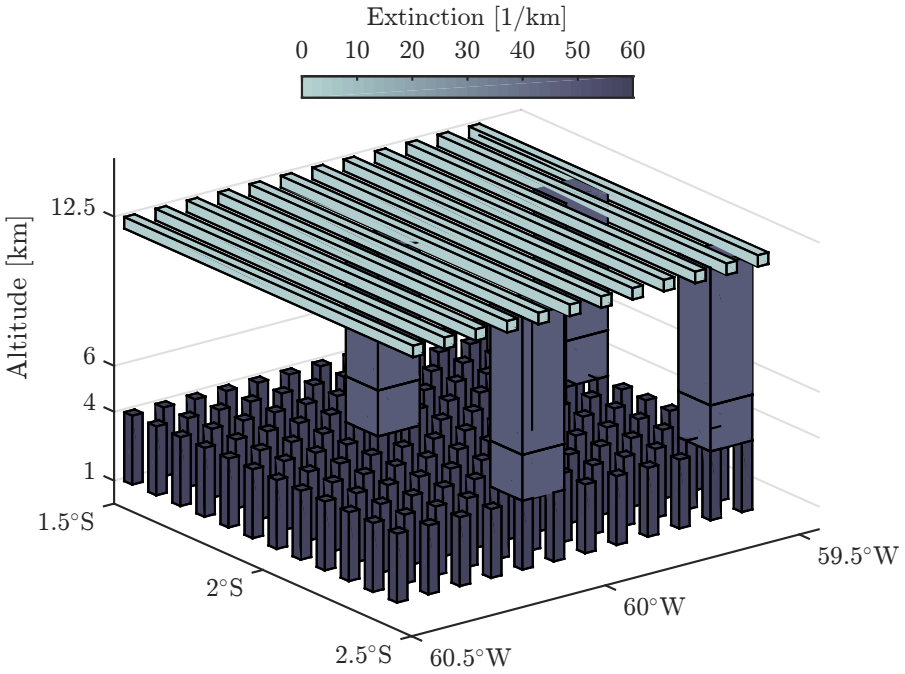

Figure 3: Surrogate cloud field assumed in the α-ratio sensitivity study (for details see text) (Knecht, 2015).

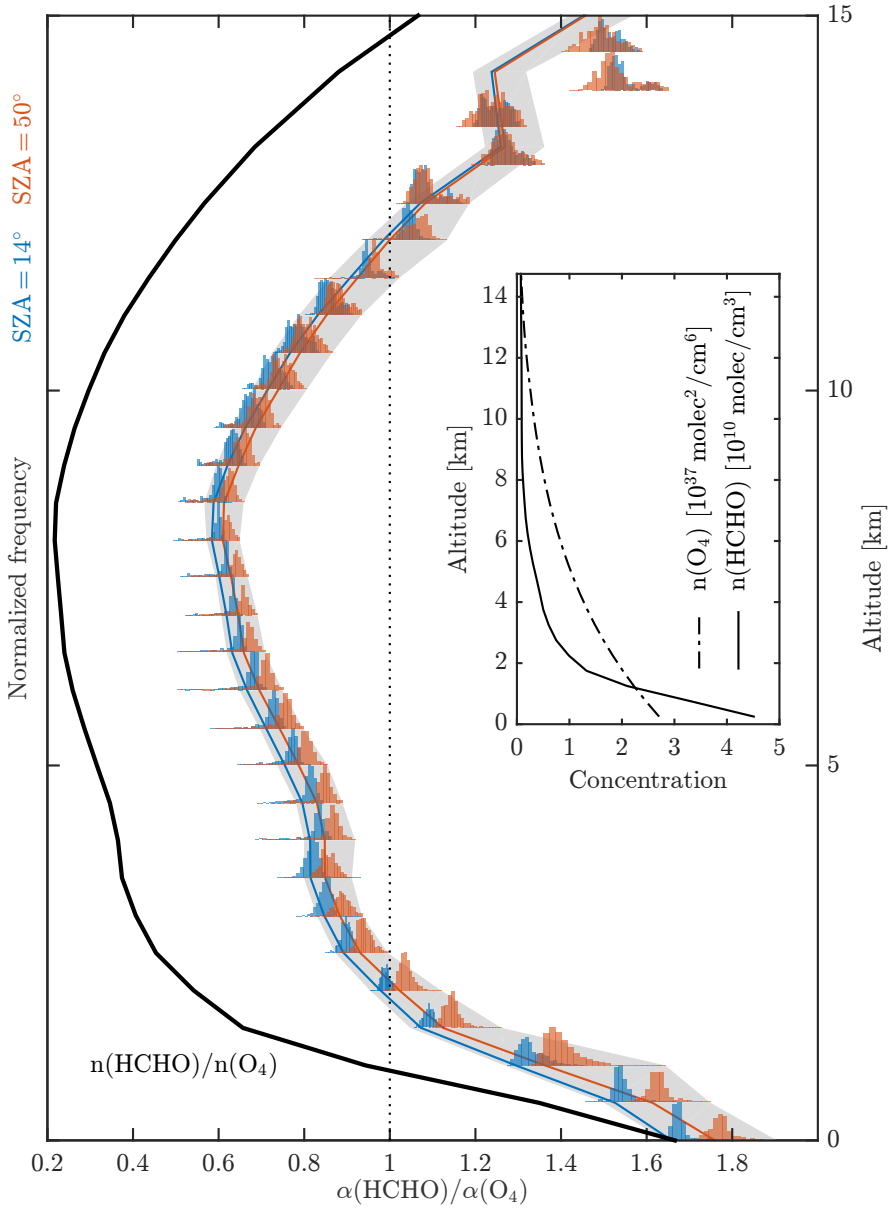

Figure 4: Simulated ratios of $\alpha_{HCHO}/\alpha_{O_4}$ factors for the assumed profiles of HCHO and O_4 as shown in the insert. The concentration ratio n_{HCHO}/n_{O_4} is shown as black profile. In the simulations the telescope is randomly placed and oriented in the azimuth direction but the elevation angle is kept constant at EA = 0°to probe the atmosphere. Two sets of simulations are shown: A purely Rayleigh scattering atmosphere (blue/red lines and grey uncertainty range) and Rayleigh and Mie scattering in the surrogate cloud field shown in Figure 3 (blue/red frequency distributions of simulated Alpha factor ratios at the indicated altitude). Simulations are carried out for solar zenith angles of 15°(blue) and 50°(red). Adopted from (Knecht, 2015).

2.2 Small scale variability

Figure 5 shows the comparison of CLaMS predictions and in situ measurements of O_3, which indicates (a) a constant offset for parts of the data and (b) a scatter of in situ measurements around CLaMS predictions, which is interpreted as small scale variabilities which are not covered by the model. The in-situ measured O_3 mixing ratios are smoothed from 100 seconds before start of each spectrum integration to 100 seconds after end of each spectrum integration in order to account for the horizontal sensitivity of each measurement due to multiple scattering.

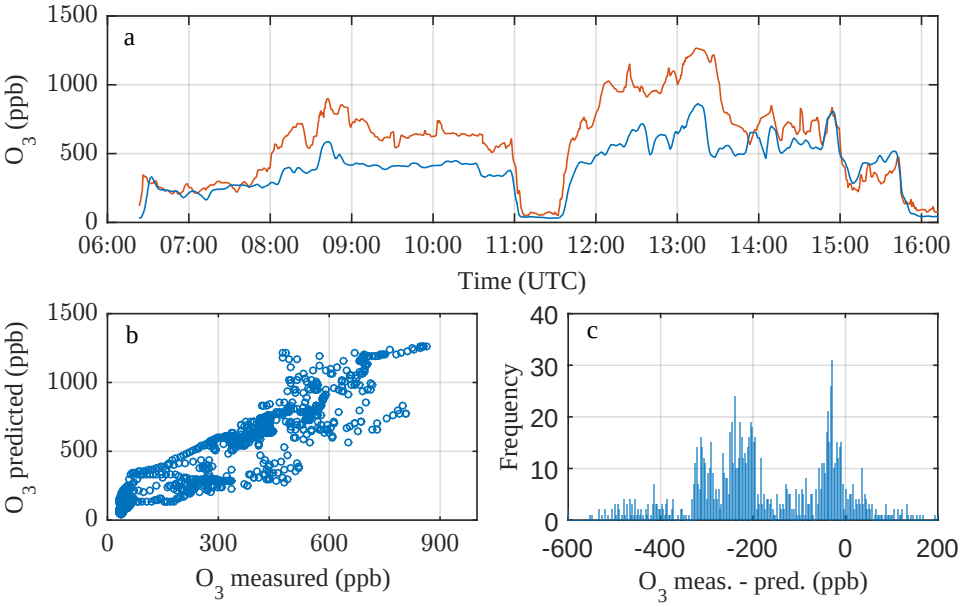

Figure 5: Comparison of in-situ measured (panel a, blue) and CLaMS predicted (panel a, red) O_3 mixing ratio for the HALO flight from Cape Town to Antarctica on 13 September 2012. Panel b: Correlation of in-situ measured (Fairo) and CLaMS predicted O_3 mixing ratios. Panel c: Frequency distribution of the difference in in-situ measured (Fairo) and CLaMS predicted O_3.

2.3 Sum of random error contributions

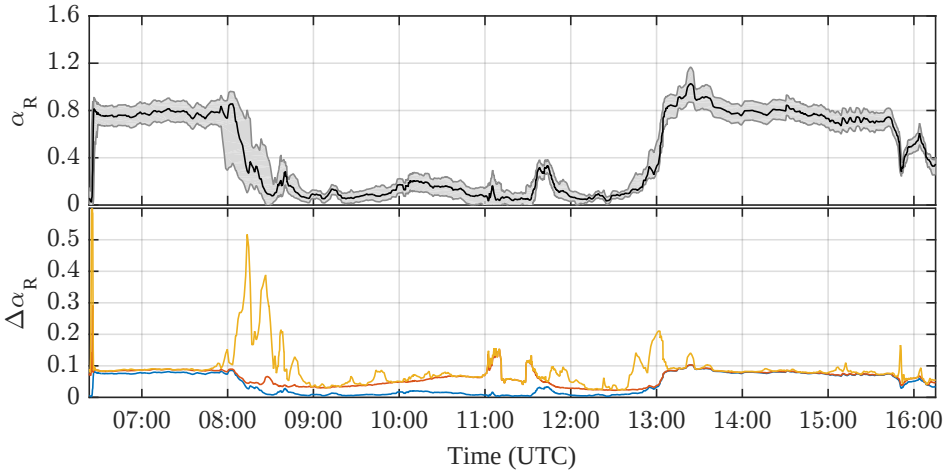

Figure 6: Summary of major random error contributions to $\Delta\alpha_R$ for the HALO flight on 13 September 2012. The α_R ratio (black line) and it's uncertainty $\Delta\alpha_R$ (grey shaded area) including all statistical components (top panel). Contribution of individual systematic errors to $\Delta\alpha_R$ subsequently added up (bottom panel): (a) The blue line show a 10% uncertainty in the Mie extinction α_R, (b) the red line shows the uncertainty due to small scale variability added to contribution (a) and the yellow line shows the vertical sampling uncertainty added to (b).

3 α_R systematic errors

This section gives further details on the calculations to estimate potential systematic errors of the calculated α_R due to erronous predictions of the trace gas profile shapes by the employed CTMs, as described in section 3.7.2 of the paper. As mentioned there, the trace gas curtains are altered in such a way that (a) the concentration of O_3 at flight altitude changed to match the in situ measured concentration, and (b) an altitude offset is calculated by comparing in situ measured and predicted N_2O, and the curtains of O_3 and NO_2 are both shifted accordingly.

Figure 7 shows how α_{R} and the inferred $[NO_2]$ are affected in both approaches. In panel (a) the predicted and measured N_2O are compared, and panels (b) and (c) demonstrates how the N_2O profiles need to be vertically shifted in order to obtain an agreement between the model predicted and in situ measured N_2O. Note that the subsidence of higher stratospheric air in the polar vortex (at 09:00 – 13:00 UTC, see Rolf et al., 2015) is not reproduced by the EMAC model. The resulting differences in measured and predicted N_2O (panel a) indicate a vertical misalignment of up to 4 km (panel b). The trace gas profiles of O_3 and NO_2 are shifted together with N_2O. Panel (d) compares α_{R} based on the different assumptions regarding different profile shapes of the involved gases, and panel (e) the compares the inferred $[NO_2]$. When forcing the predicted concentrations to the in-situ measured O_3 at flight altitude (while keeping the predicted O_3 profile fixed elsewhere) the change in inferred $[NO_2]$ (green dots in panels d and e) mostly falls into the uncertainty range given by the random error (the gray shaded area in panel e). Evidently the largest discrepancy as compared to the standard run in the simulated α_{R} and inferred $[NO_2]$ occurs when the trace gas profiles are altitude shifted (the purple dots in panel d and e). This uncertainty becomes in particular large for flight sections where O_3 has strong vertical gradients (e.g. 8:30 – 9:30 UTC and 12:00 – 13:00 UTC). Hence, vertical transport processes which are not covered by EMAC may change the profile shapes in such a way that systematic errors in the scaling method retrieval arise. However, since trace gases such as NO_2 and O_3 are chemically reactive, the altitude shift carried out here may not reflect the change in trace gas profile shape experienced by NO_2 and O_3. This is indicated by the fact that the modifications mentioned previously (green dots) do not give rise to large uncertainties in these flight sections, i.e. measured $[O_3]$ and $[NO_2]$ do not differ strongly from predictions at flight altitude.

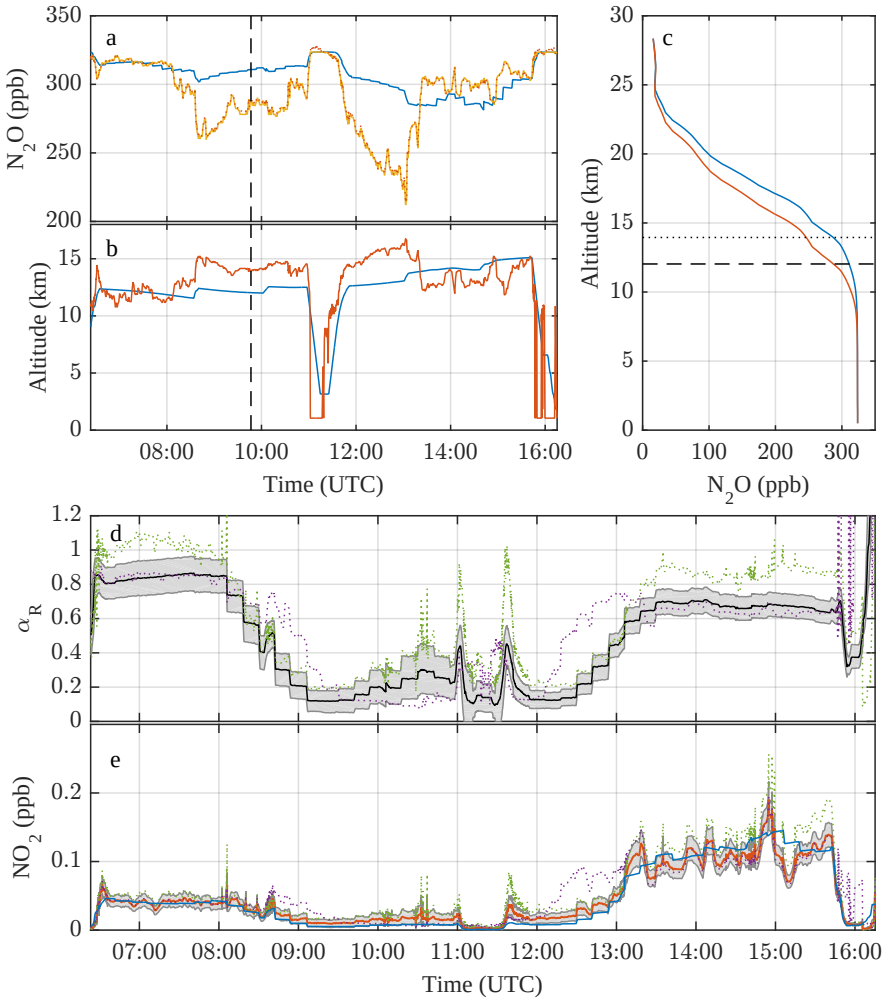

Figure 7: Comparison of sources of potential systematic errors investigated for the HALO flight on 13 September 2012. Panel a: EMAC predicted N_2O mixing ratio (blue), N_2O when shifting the predicted EMAC profiles as indicated in panel c (yellow) until predicted N_2O agrees with in-situ measured N_2O at flight level (red dots). Panel b: Comparison of aircraft GPS altitude (blue) and apparent altitude (red) where in-situ measured and EMAC predicted N_2O would agree. Panel c: EMAC predicted N_2O and altitude shifted N_2O for the time indicated by the black dashed line in the panels a and b. The broken line marks the aircrafts altitude, the dotted line marks the altitude of the respective O_3 mixing ratio in the unchanged EMAC profile. Panel d: α_R (black line) and $\Delta\alpha_R$ due to random errors (gray shaded area). Additionally, α_R resulting from either forcing the simulated to the in-situ measured O_3 concentration at flight altitude (green dots) and those resulting from the altitude shifting the O_3 profiles (purple dots) are shown. Panel e: Inferred NO_2 mixing ratios for (a) the unchanged O_3 profile (red line ~~the~~ and uncertainty range displayed as gray shaded area), (b) when forcing the EMAC predicted to the measured O_3 at flight altitude (green dots), (c) when altitude shifting the profile until the simulated and measured N_2O agree at flight altitude (purple dots), and (d) for comparison the predicted EMAC [NO_2] (blue line).