

Final author comments on the manuscript amt-2016-355, reviewer 1

We would like to thank reviewer 1 for the constructive comments that aided us to improve our manuscript. In this document we provide our replies to the reviewer’s comments. The original comments made by the reviewer are numbered and typeset in italic and bold face font. Following every comment we give our reply. We provide a new version of the manuscript but in our replies to the comments we provide line numbers, page numbers and figure numbers referring to the original version of the manuscript, if not stated differently.

1. ***Extension of the spatial coverage (especially over the ocean) by analyzing cloudy scenes is desirable. However, for thick clouds the sensitivity to CO below the cloud is very limited. As shown in the manuscript this does not hamper validation, because the corresponding null-space error can be taken into account mathematically by applying the averaging kernels. One could argue that this makes good validation results even easier to achieve, because potentially highly variable CO partial columns below clouds are essentially substituted by a priori partial columns for both the satellite data and the validation data set in the comparison. This makes physical interpretation of the retrieval results in cases of high clouds, in particular above CO source regions, difficult, because the retrieved parameter is rather the CO partial column above the cloud (extended by the a priori below the cloud) than the CO total column and the most interesting part may be hidden below the cloud. In this sense, validation and physical significance can potentially be two different stories. Please elaborate a little more on this issue in the manuscript. Therefore, I assume that retrievals for (too) high clouds are filtered out in the post-processing. Please describe in more detail what the corresponding threshold for the presented SCIAMACHY data set is (e.g., in Figure 10). I would propose to omit scenes with cloud heights larger than about 2 or 3 km: This would still allow retrievals above low-eta clouds like stratocumulus (Sc) over the ocean.***

adjusted The reviewer argues that the evaluation of the CO bias is distorted in case of high clouds because both the satellite data and the validation data set are substitute by the same a priori partial columns in the comparison and so this effectively reduces relative errors on the column. To refute this argument, we’d like to outline an important feature of the profile scaling approach used in our study. Here, the retrieval approach scales a reference profile to fit a forward model to the measurement. In case of a cloudy atmosphere, the scaling parameter relies on the CO measurement sensitivity above the cloud but adjusts the CO concentrations at all altitudes due to the scaling. If the scaling parameter is retrieved with a certain error, this effects the CO concentrations at all altitudes in the same manner.

So for a cloudy-sky retrieval, a measurement or forward model error affects the CO sub-column above the cloud in the same manner as it does for the total column. Therefore, the physical interpretation of the relative biases is the same for clear-sky and cloudy retrieval.

The fact that the CO bias for the site Windhoek in Fig. 7 is similar for clear-sky and cloudy retrievals, indicates a similar retrieval accuracy for CO at altitudes the measurement is sensitive to. In this sense, we are convinced that the comparison does not favor a validation of cloudy retrieval over clear-sky retrieval.

To emphasize this aspect, we added the following sentences at p6,l24:

“For the interpretation of errors, it is important to note that for the scaling of a reference profile, the interpretation of relative biases is very similar for clear-sky and cloudy retrieval. In both cases, it indicates the relative error to the CO concentrations at all altitudes the retrieval is sensitive to.”

Moreover, we add at p9,l14:

“ Consequently, the fact that clear-sky and cloudy retrievals result in the same bias indicates that the accuracy for CO is comparable at altitudes the measurements are sensitive to. ”

and at p4, l23:

“ Hence, when considering c_{eff} as an estimate of the true column, data interpretation does not rely on the use of the total column averaging kernel but the null-space error e_{null} becomes part of the error budget. When c_{eff} is seen as an effective column with its vertical sensitivity described by the column averaging kernel, data analysis focuses more on the information that can be inferred from the measurements and reduces the impact of the a priori choice of the reference profile. This requires a proper use of the total column averaging kernel as shown in Eq. (4) and the null-space error does not contribute to the error budget (see also Wassmann et al. (2015)). ”

For validation with TCCON and NDACC in Fig 9 and to illustrate the global CO distribution in Fig. 10, it is important that the CO column have similar vertical sensitivity and can be interpreted as an estimate of the true total column. Therefore, for this figures, we use only data with a retrieved cloud top height ≥ 1.5 km. However, based on the MOZAIC comparison, we are convinced that also CO retrievals with

higher clouds are very useful when exploiting the data set, e.g. due to the assimilation of the CO product in a chemical transport model. Therefore we are reluctant to filter the data a posteriori as suggested by the reviewer.

2. *Page 8, Lines 13-14 and Figures 4-6: Why are cloudy conditions defined as high clouds here (condition (3) in Sect. 2.3)? The comparison with airborne measurements is mainly above CO source regions (Teheran, Beijing). This is an example of the interpretation problems described in the General Comments. Figure 6 confirms that a lot of CO is hidden below the clouds (comparison of yellow and pink bars). I would prefer to use condition (2) here or an alternative condition (2b) with h_{cl} > 2 or 3 km.*

changed De Laat et al. 2012 showed that the validation of SCIAMACHY clear sky measurements at these sites close to CO source regions suffer from representation errors, i.e. that due to the large pixel size of SCIAMACHY the aircraft measurements cannot agree with the satellite observations because of probing different CO air masses. The same holds for low clouds. However, for high clouds, the horizontal distribution of CO above the cloud is more homogenous and so representation errors are less relevant. As explained in our reply to the reviewer’s general comment, the comparison between satellite and aircraft measurements are not distorted because of the profile scaling approach, and so the good agreement for SCIAMACHY observations with high clouds, support the interpretation of de Laat et al., 2012 and in addition gives confidence in the SCIAMACHY cloudy data product. To clarify this, we change the discussion in the manuscript (p8,l26) from

“The time series in Fig. 5 show strong outliers in the MOZAIC/IAGOS CO column densities, which are not found by SCIAMACHY. de Laat et al. (2012) explained this with strong local pollution at the airports, which affects the airborne measurements but are not seen by the satellite due to the coarse spatial resolution. Figure 7 confirms this reasoning. Applying the column averaging kernels of the cloudy-sky retrievals to MOZAIC/IAGOS measurements, the enhanced CO column values are strongly reduced for both Teheran and Beijing and brings the aircraft and satellite observations into much better agreement. Thus confirms that the column averaging kernel of the cloudy-sky retrieval describes well the cloud shielding of the lower atmosphere (see Fig. 2) where the local pollution is primarily located. “

to

“de Laat et al. (2012) explained the larger errors for clear-sky observations with strong local pollution at the airports, which affects the airborne measurements but are not seen by the satellite due to the coarse spatial resolution. This is concert with the extremely large values of the MOZAIC/IAGOS CO column densities time series in Fig. 4. Hence, we expect a better agreement, when we consider SCIAMACHY retrievals for high clouds because the atmospheric shielding of the spatial heterogeneity of CO. Figs. 4 and 2 confirms this. Applying the column averaging kernels of the cloudy-sky retrievals to MOZAIC/IAGOS measurements significantly improve the comparison with SCIAMACHY retrievals and so supports the error interpretation by (de Laat et al., 2012) but also demonstrates the data quality of SCIAMACHY cloudy-sky retrievals. “

3. *Page 9, Lines 23-24: How about applying TCCON/NDACC averaging kernels to profile information from the TM5 CO model?*

not changed The paragraph at p9, l23-24 is about MOZAIC/IAGOS profile measurements and their use to validate our product. These airborne measurements can be used to validate clear-sky and cloudy SCIAMACHY retrievals and so provide interesting information on the consistency of the data product and the relevance of representation errors for the validation. This is different for TCCON/NDACC measurements, which can only be used to validate clear-sky retrievals and cloudy-sky retrievals for low clouds at remote regions. Furthermore, applying averaging kernels on TM5 profiles will not help because those are used as reference profiles in our retrieval. In Borsdorff et al. (2014) we showed that the reference profile cannot be used to fill up the null-space, i.e. $(C - A_{col})x_{TM5} = 0$.

4. *Page 10, Line 16: The given range of h_{cl} corresponds to condition (3) in Sect. 2.3, not (2). Which condition is actually used, (2) or (3)? I would recommend to do the analysis for low clouds only (Condition (2) or (2b)).*

corrected Correct, we use condition (2) here. We changed the sentence at (p10, l16) from

“When analyzing SCIAMACHY cloudy-sky CO retrievals with $1.5 \text{ km} < h_{cl} < 8 \text{ km}$, $\tau_{cl} > 2$ corresponding to condition (2) ...”

to

“When analyzing SCIAMACHY cloudy-sky CO retrievals with $h_{cl} < 1.5 \text{ km}$, $\tau_{cl} > 2$ corresponding to condition (2) ...”

5. *Page 10, last paragraph and Figure 10: 1) Which conditions for the full-mission SCIAMACHY cloudy-sky measurements are used? As described above, I would propose to omit scenes with cloud heights larger than about 2 or 3 km.*

changed We clarified this point in the manuscript as described in our answer to comment 4 of this review (p10, l16). Furthermore, we changed the sentence at p10,26 from: “The full-mission averaged CO product is shown in Fig. 10 and indicates . . .”

to

“The full-mission averaged CO product is shown in Fig.10 and is based on clear-sky and cloudy observation contaminated by optical thick low clouds ($h_{cld} < 1.5$ km). This figure indicates . . .”

6. *2) You describe CO outflow over the ocean. As the retrievals over the ocean are from cloudy-sky measurements only and the retrievals over land are a mixture from clear-sky and cloudy-sky measurements, are there any indications of land-sea-biases due to cloud shielding? This is hard to see in Figure 10, because the color-scale is often saturated at the mentioned source regions.*

changed To consider potential error correlation with land-sea transitions, we investigated biases for the coastal site Wollongong but did not find any indication for a land-sea bias in the data. To emphasize this point, we changed the sentence (p10, l21) from:

“For Wollongong, the mean bias for clear-sky and cloudy-sky retrievals is comparable (about 4.0 ppb), but the correlation improves from about $r = 0.5$ to 0.8 for the cloudy retrievals.”

to

“For Wollongong, the mean bias for clear-sky and cloudy-sky retrievals is comparable (about 4.0 ppb), and the correlation improves from about $r = 0.5$ to 0.8 for the cloudy retrievals. Here the majority of the cloudy measurements are observations over the oceans and so these results indicate that there is no evidence for a land-sea bias of the SCIAMACHY CO retrieval.”

7. *3) It would also be interesting to show a global map illustrating the fraction of cloudy-sky measurements. Are the source regions dominated by clear-sky or cloudy-sky measurements?*

added We included the figure suggested by the referee. The discussion is added to the manuscript at p7,19

“Figure 4 shows that much more data become available when considering cloudy-sky retrievals under optical thick low and high cloud conditions in addition to clear-sky retrievals. Between January 2003 to the end of the SCIAMACHY mission in April 2012, 70 % of all data are inferred from cloudy-sky observations (40% over the oceans and 30% over land). Over the oceans only cloudy-sky retrievals are possible but over land the fraction of clear-sky to cloudy-sky retrievals varies. For example, the data coverage over the Sahara region is dominated by clear-sky retrievals while the Amazon region and Indonesia show cloudy-sky retrievals be in a majority. Averaged globally, 55% of all measurements over land are cloudy-sky retrievals. Hence for SCIAMACHY, our CO retrieval algorithm for cloudy conditions permits the CO data product over the ocean, and in addition has the potential to improve the product over land.”

8. *Figure 8: Why is TCCON data release GGG2012 (instead of GGG2014) used for Ny-Alesund?*

added We added this explanation to the figure description: “. . .for Ny-Ålesund data release GGG2012 is used because the newer release GGG2014 is not yet available for this site.”

9. *Page 8, Line 30: Typo? Thus This*
corrected

10. *Page 9, Line 26: Please add :: ...is most reliable: Ny-Alesund, ...*
corrected

References

- Borsdorff, T., Hasekamp, O. P., Wassmann, A., and Landgraf, J.: Insights into Tikhonov regularization: application to trace gas column retrieval and the efficient calculation of total column averaging kernels, *Atmos. Meas. Tech.*, 7, 523–535, doi:10.5194/amt-7-523-2014, 2014.
- de Laat, A. T. J., Dijkstra, R., Schrijver, H., Nédélec, P., and Aben, I.: Validation of six years of SCIAMACHY carbon monoxide observations using MOZAIC CO profile measurements, *Atmos. Meas. Tech.*, 5, 2133–2142, doi:10.5194/amt-5-2133-2012, 2012.

Wassmann, A., Borsdorff, T., aan de Brugh, J. M. J., Hasekamp, O. P., Aben, I., and Landgraf, J.: The direct fitting approach for total ozone column retrievals: a sensitivity study on GOME-2/MetOp-A measurements, *Atmospheric Measurement Techniques*, 8, 4429–4451, doi:10.5194/amt-8-4429-2015, 2015.

Final author comments on the manuscript amt-2016-355, Ruediger Lang

We would like to thank Ruediger Lang for the constructive comments that aided us to improve our manuscript. In this document we provide our replies to the reviewer's comments. The original comments made by the reviewer are numbered and typeset in italic and bold face font. Following every comment we give our reply. We provide a new version of the manuscript but in our replies to the comments we provide line numbers, page numbers and figure numbers referring to the original version of the manuscript, if not stated differently.

1. *Retrievals involving the knowledge of certain above-ground mean-scattering heights with negligible information retrieved from below the scattering layer require special pre-cautionary measures taken by the producer to avoid mixing of total and partial column information. The authors mention in Section 2.3 on page 6 the potential of providing complementary information due to this mixing, but do not go in any detail how this complementary information (what is meant is probably profile information) can be retrieved or is used.*

changed We modified the paragraph (p6,l32)

“CO retrievals for cloudy conditions represent an interesting addition of our data product because of the different vertical sensitivity, the retrieved CO column probes different altitude ranges and so providing complementary information. . .” to

“CO retrievals under cloudy conditions represent an interesting addition to our clear-sky data product. Here, the retrieved column is mostly sensitive to CO above the cloud and for varying cloud height and due to the shielding effect of clouds, different altitudes are probed. For the interpretation of errors, it is important to note that for the scaling of a reference profile, the interpretation of relative biases is the same for clear-sky and cloudy retrieval. In both cases, it indicates the relative error to the CO concentrations at all altitudes the retrieval is sensitive to. Furthermore, ”

2. *I find the way the retrieved column is interpreted relative to the true profile in equation 4 quite misleading with respect to the formulation in the conclusion saying: Generally for clear-sky observations the null-space error is small, but for cloudy conditions it can easily exceed 30 percent. Here, clouds shield the atmosphere below and the a priori CO profile shape is used to add the lacking information. Here it is formulated as if ghost column information from the scaled a-priori profile is added to complement the lacking information of the retrieved column above the scattering layer, while in Equation 4 it looks like the profile information is over-weighted with an averaging kernel above the cloud, therefore using only (or predominantly) a priori profile information from above the scattering layer in the retrieval. What is true?*

text adjusted The reviewer's point does not represent a contradiction the two mentioned statements are mathematically equivalent as described in Borsdorff et al. (2014). To clarify this, we adjusted the text at p6 line 12, referring to our previous work to keep the discussion short.

“ Column averaging kernel values of one represent the ideal value for a vertical integration (see Eq. (2)) and values < 1 indicate reduced retrieval sensitivity e.g. because of atmospheric shielding by clouds and aerosols. Values > 1 are typical for a profile scaling approach. In case of an optically thick cloud the sensitivity of the measurement for CO below the cloud is lost. The inversion scales the entire vertical profile of CO based on its sensitivity to CO above the cloud. Hence, the sensitivity of the retrieved CO column with respect to the CO concentration above the cloud is enhanced (> 1) and at the same time the sensitivity for CO below the cloud is reduced (< 1). These features are clearly reflected in Fig. 2 and discussed in more detail by Borsdorff et al. (2014). ”

3. *In addition, the potential draw-backs of providing this mix of effective and total column information is not discussed in much detail in the paper. While this type of data-set can potentially be used relatively safely in model assimilation, where the use of the averaging kernels is an integral part of the assimilation process, it is less clear how such a data-set can or should, if at all, be used in the way it is displayed in Figure 10. This does not render the data-set useless of course. However it requires a discussion on the purpose of the data-set presented here.*

changed We agree with the reviewer that in general the use of the column averaging kernel is essential for proper data interpretation, as it is stated in the manuscript (p6 l22-24). For validation, CO profile measurements are required, which are only available at a few sites. A corresponding validation strategy is discussed in Sec. 3.1. We also argue that for low clouds over unpolluted NDACC/TCCON sites, the retrieved columns can be considered as a reasonable estimate of the true column density of CO within

the overall uncertainty of the SCIAMACHY CO measurement. This is demonstrated in Sec. 3.2 showing validations with TCCON and NDACC measurements. To illustrate the new SCIAMACHY data product, we included these low cloud observations in Fig 11 (previous 10) and added the following sentence to the manuscript (p10,l26): “In particular, the validation with NDACC and TCCON measurements revealed that the CO column, retrieved from measurements with low, optically thick clouds, can be used as an estimate of the true column (i.e. ignoring effects of the column averaging kernel) in absence of local sources of CO.”

4. *From my point of view, such a SWIR CO SCIAMACHY data-set is only complete once the results presented have been assimilated in a model, i.e. as a level-3 or 4 products, and model information has been added to the part of the profile not visible to the retrieval. From this perspective the scope of this paper would then be the presentation of a major step towards such a final level-4 data-set. I therefore suggest to add a discussion of this aspect (maybe even consider to change the title to sub-column retrievals) and consider to add an outline on how to complete the task, in a next step. E.g. by providing a final SCIA CO SWIR data-set by adding missing sub-column information for individual cloudy-sky retrievals involving a model or other means (i.e. reducing the null-space error). We should not forget that there are currently many atmospheric composition and other retrieval methods available which potentially can provide an-above-cloud-column (e.g. water vapour retrievals from the same type of instrumentation) but choose instead to provide data-gaps in order to avoid confusion and biases, even though averaging kernels are available as well for these retrievals. While overcoming instrument SN deficiencies are the means to an end it cannot be the end for the means.*

changed We are reluctant on a statement that the data set is only completed after assimilating the retrieval results. Each data level has its value set and should be validated based on the data uncertainty of the subjacent data level. Therefore, reporting on the level-2 CO SCIAMACHY data product and its validation we consider as a valuable contribution. We are also convinced that applications of the level-2 product other than data assimilation are of scientific value and should not be excluded a priori. On the other hand, we also agree that data assimilation as e.g. considered by the European CAMS project is an important application and should be investigated in future studies. Therefore we added to the conclusions (p12,l19):

“... of the SCIAMACHY CO data product. The correct interpretation of these data requires the use of the column averaging kernel. For varying cloud height and due to the shielding of the atmosphere below the cloud, the data may be used in future work to discriminate the vertical distribution of CO in the atmosphere, e.g. by means of data assimilation.”

5. *Abstract: line 10: low bias...: Some SCIAMACHY CO retrievals apply a bias correction to the CO product, e.g. to agree with MOPITT at source free locations (Pacific) or with trusty ground stations. Is any bias correction applied prior to this comparison?*

changed We do not apply any a posteriori bias correction on CO. However, as shown in Borsdorff et al. (2016), Fig 4, we apply a radiometric correction to the spectral measurements of SCIAMACHY. This correction is estimated from high albedo measurements over Sahara assuming that the chemical transport model TM5 provides the true methane profile, e.g. p5,l5.

“To account for radiometric errors of the SCIAMACHY measurements, we apply a temporal dependent radiometric correction derived from SCIAMACHY observations of the Sahara and reliable model prediction of methane over this area. Moreover, the measurement noise is estimated from regular eclipse measurements and both aspects are described by Borsdorff et al. (2016).”

6. *Abstract: line 20: overall the cloudy-sky... Considering the improvements exposed above, would not cloudy-sky measurements be something more than simply a "valuable addition"?*

changed We agree and adjusted the abstract at (p1,l20) from

“Overall the cloudy-sky CO retrievals from SCIAMACHY short wave infrared measurements present a valuable addition to the clear-sky only data set.”

to

“Overall the cloudy-sky CO retrievals from SCIAMACHY short wave infrared measurements present a major extension of the clear-sky only data set, which more than triples the amount of data and adds unique observations over the oceans.”

7. *Introduction: line 7: two stream radiative transfer solver.. Are two streams sufficient to accurately simulate observed radiances at top of the atmosphere? This simplification could lead to observation geometry dependent biases. Are sphericity effects taken into account?*

changed This point is discussed in detail by Landgraf et al. (2016). For individual cases, the induced errors may reach several percent but becomes insignificant when considering regional scales. To address this aspect, we change the sentence at (p3,17) from:

“ These parameters are used in a two stream radiative transfer solver called 2S-LINTRAN (Landgraf et al., 2016) to account for the changes to light path that occur due to scattering by clouds and aerosols. ”

to

“ These parameters are used in a two stream radiative transfer solver 2S-LINTRAN to account for changes to the light path due to light scattering by clouds and aerosols. For cloudy conditions, the CO error induced by this radiative transfer solver reaches occasionally several percent (Landgraf et al., 2016) and becomes marginal for regional measurement ensembles. The SICOR data product”

8. *Section 2.1, line 23: Are the SCIAMACHY operational L1 noise values reliable? Otherwise - although formally correct - the consideration of the measurement noise could worsen the CO estimates. Or is the CO product derived from a different SCIAMACHY L1 product?*

changed We are not using the operational L1 noise as described by Borsdorff et al. (2016), p232: “The channel 8 SCIAMACHY measurement noise is dominated by detector dark noise, which is estimated from SCIAMACHYs daily dark state measurements taken during the orbit eclipse. ”We found that this approach improves significantly our retrieval because the noise of every detector pixel is basically measured as a function of time.

A reference and a corresponding remark is added at p5 l29 as stated in reply 5: “ For SCIAMACHY CO retrievals, we use the recalibrated SCIAMACHY spectra including estimates of the measurement noise as described by Borsdorff et al. (2016). ”

9. *Section 2.2, page 5, 1st paragraph: A short motivation for increasing or choosing this spectral window would be helpful (like for the reduction of noise. information content, type of absorbers, ice layer contamination etc...).*

changed The selection of the retrieval window is described in detail by Borsdorff et al. (2016). To clarify this we add the sentence (p5,15): “. . .This is necessary to account for the instrument degradation e.g. the considerable loss of detector pixel by radiation damage and the light scattering of the ice-layer on the detector array of SCIAMACHY’s channel 8 (Borsdorff et al., 2016). “

10. *Section 2.2, page 5, line18: If the used cloud model has a fixed half-width of 1.5km, the lowest possible cloud top height in the forward model is 3km. Does this not constrain the retrieval in an unnecessary way?*

changed The retrieval is not restricted to a lowest cloud height > 3 km but the vertical profile of the cloud is truncated at the Earth surface. To clarify this we add the following sentence (p5,l19):

“In case of cloud heights < 3 km the cloud profile is cut off at the Earth surface and is renormalized”

11. *General: I find the term ”physics-based” retrieval bit problematic. It suggests that you may do certain things somewhere else, which are not physics based. Do you? I hope not! Full-physics would be also problematic, since many effects have not been considered in the retrieval, e.g. horizontal variability inside the spatial pixels.*

changed We change the term “physics-based retrieval” to “scattering retrieval” through out the manuscript.

12. *Section 2.2, page 5 last paragraph: The cloud optical properties depend on the micro-physical ones. Accordingly, the retrieval results depend on the assumed cloud micro- physics. Could you be more specific on the cloud microphysical model which has been used?*

not changed

For the retrieval of effective cloud properties as done by SICOR, the step to derive optical properties from microphysical properties is not required. For example, it is sufficient to describe the cloud by its optical depth and the asymmetry factor of the scattering phase function without analyzing the underlying droplet size distribution and refractive index. The approach is discussed in detail by Landgraf et al. (2016) and an appropriate reference is already given in the manuscript.

13. *Section 2.3, page 6: After 2005, SCIAMACHY channel 8 detector suffered irreversible damage decreasing the quality of the spectra. Do the FRESCO and SICOR cloud heights also correlate that good for other years after 2005?*

changed The year 2005 is not a specific year concerning the SCIAMACHY channel 8 detector performance, which actually suffered from a continuous degradation. Nevertheless, the SICOR cloud height retrieval is only affected marginally by this degradation. We correlated the FRESCO data with the SICOR cloud height for the year 2010 and found a similar good comparison. Therefore, we added a sentence at (p6,l5) and (p11,l17)

“Even for later years of the mission we find a similar agreement (e.g. $r = 0.8$, slope= 0.9, intercept=431 m for the year 2010). “

14. *Section 2.3, page 6, line 14ff: In case of an optically thick cloud... It is not clear from this formulation whether scaling of the profile in a retrieval under cloudy conditions is performed differently or in the same way as in the clear sky case. When it is stated that "... using only the measurement sensitivity above the cloud", does it mean that "the atmosphere bellow the cloud is neglected and not taken into account during the fitting process" or that "the forward model, which includes a cloud layer, is practically insensitive to CO bellow the cloud and mainly fit the CO information above the cloud"? Please clarify.*

changed The retrieval still fits a scaling of a reference profile but the forward model becomes insensitive due to the presence of a cloud. To clarify this, we changed the manuscript according to comment 2 of this reviewer.

15. *Section 2.3, I am missing a discussion of the fact that you expect/want the cloud parameters to be different from the FRESCO once retrieved in the NIR in first place. Since otherwise you could simply use the latter to at least adjust your h_cld. I think you should mention here the reasons why and how you expect them to be different in the SWIR. For sure one expects them to be higher in the SWIR, which is what one can see.*

changed FRESCO and SICOR use a different way to describe clouds in an effective manner. For example, FRESCO uses a elevated Lambertian reflector of predefined albedo which partially covers the ground scene. SICOR uses a homogenous scattering layer with triangular height profile and variable optical depth. Moreover, the SCIAMACHY NIR and SWIR measurements are not necessarily collocated, which may introduce a pseudo-random error contribution. Therefore, we add the following sentence at (p6,l5):

“The remaining small difference between the retrieved cloud height and the FRESCO cloud top height may origin from the different cloud models used. For example, FRESCO simulates clouds as an elevated Lambertian surface, whereas SICOR fits the center height of the triangular height profile.”

16. *Section 2.3, page 6, line 27: Also for the validation.... Strictly speaking the data-use is limited. Since its limited to the profile part where the null-space is small. While the sentence is correct in principle it paints a too broad picture of data use.*

changed We change the sentence starting from (p6,l25)

“ Generally, the null-space error does not limit the data use if the averaging kernel is applied properly. For example, for data assimilation the averaging kernel provides the altitude sensitivity, which is needed to adequately adjust the atmospheric state. “ to

”The null-space does not provide a problem for application like data assimilation, if the averaging kernel is applied properly. Exploiting the altitude sensitivity given by the averaging kernel, the atmospheric state of the model can be adjusted adequately by the assimilation scheme. Also for the validation of the CO data product, the null-space error does not impose principle limitation, where Eq. (5) accounts for cloud effects in our retrieval when comparing satellite measurements with independent profile soundings.“

17. *Section 2.3, page 6, line 32: CO retrievals for cloudy conditions... Is the complementary use of CO columns with different averaging kernels to retrieve vertical profile information demonstrated somewhere in the paper (see general remark)? In contrast, the reduction of the SN and the consequent reduction on the retrieval noise is clearly demonstrated and should function as the key-motivation to use cloudy scene retrievals.*

adjusted In the manuscript we discuss the use of the retrieved column for cloudy atmospheres with varying cloud height to infer CO height information as an potential future application, which can be achieved by data assimilation. To make this point clear, we adjusted the text accordingly to our reply to reviewer’s comment 1.

18. *Section 2.3, page 7 last paragraph, last sentence: ...and in addition has the potential to improve the CO product over land. ->In cases where the true and the a-priori CO profile are similar, and where no significant deviations in CO concentration (true vs. a-priori) below the cloud layer occur.*

not changed We disagree, the product is improved in more general terms (e.g. data coverage, lower noise error, ...). The data filtering mentioned by the reviewer is only required to validate our data product with TCCON/NDACC measurements. When profile measurements are available, the validation is not hampered by this, as discussed in Sec. 3.1. Here, the total column averaging kernel can be deployed and the validation is possible without any restriction on the profile shape and a priori knowledge.

19. *Section 3: retrieval conditions: $SN > 20$: This condition will change the CO retrieval densities favoring Earth location with high surface albedos. Have you notice any data gaps in dark surface locations?* **statistics not changed** The SNR filter threshold is carefully chosen to thin out data as less as possible. Figure 3 shows that a SNR of about 20 corresponds to clear-sky measurements over the oceans and so by definition the data filter does cause some data gaps over ocean. However, its effect on land observation is minor.

20. *Section 3: retrieval conditions number 3): The quality of the SCIA channel 8 spectra and, consequently, also the CO data record, gets worse after the end of 2005. Accordingly one could think in time variable noise thresholds to select for useful CO retrievals. Are the given noise thresholds constant over the whole mission? $\epsilon_{CO} \sim 1e19$ represents roughly a relative noise error of some hundreds of percent. Please justify this quality criterion.*

changed Yes, the noise threshold is constant over the whole mission. It is a weak filter on the data to remove the largest outliers, which cannot be explained by the noise statistics. Therefore, it is not needed to adjust the threshold as a function of time. To clarify this we add the following sentence to the manuscript (p7,l21):

“ These criteria represent a weak data filtering to remove outliers due to unphysical retrievals. For example, the signal-to-noise threshold corresponds to clear-sky measurements over the oceans (see Fig 3) for which a stable inversion is not possible. The retrieval would result in unphysical outliers and are thus rejected by the filter. ”

21. *Section 3.1, page 8, line 5: Furthermore, towards... Whereas this is an acceptable assumption for unpolluted regions, I would expect that CO concentrations increase towards the surface in polluted areas, since CO emissions happens at surface (vehicle and industrial emissions, biomass burning, etc). Maybe the assumption is backed by transport over the selected circular area of 850km?*

changed We agree with the reviewer that for source regions the CO concentration may increase towards the surface. However, our experience is that a linear schemes as proposed by the reviewer introduces the risk to unstable extrapolation due to the uncertainty of the profile shortly after take-off. This is avoided by our method, which is also in agreement with the work of (de Laat et al., 2012) estimating the error of the approach to be a view percent of the total column.

We add the following sentence to the manuscript (p8,l6):

“ ... which ensures the numerical stability of the extrapolation. When the CO column density is derived from the aircraft measurement in this manner its accuracy is estimated to be a few percent (de Laat et al., 2012). Numerical experiments using the lower data points of the profile to estimate the gradient in the CO concentration for linear extrapolation led to instabilities and thus this approach is not used here.”

22. *Section 3.1, page 8, line 9: ...chosen dynamically...: Can you provide the maximum and minimum temporal difference considered in the collocation?*

changed We changed the sentence (p8,l9) from: “The temporal collocation criterium is chosen dynamically for each individual MOZAIC/IAGOS measurement.”

to

“The temporal collocation criterium is chosen dynamically for each individual MOZAIC/IAGOS measurement and varies typical between 7 and 30 days.”

23. *Section 3.1, page 9, line 3 and 4: Considering a collocation area higher than 2e6km2, the difference in city areas of about $\sim 1000km^2$ do not a play a role explaining the CO results.*

changed we agree and remove the part

“Beijing with 18.6 million urban inhabitants and an urban area of about 1400 km² has about double the size of Teheran with an urban population of about 8.8 million and an urban area of 730 km². Moreover the metropolitan area in Beijing is much larger than that of Teheran, indicating that the CO sources in Beijing may be spatially more extended than those in Teheran.”

from the manuscript (p9,l2-4):

24. *Section 3.1, page 9, 1st Paragraph and Figure 6: How does the here presented reasoning explain the fact that cloudy sky retrievals compare so much better to MOZAIC measurements with averaging kernel applied than the corresponding clear-sky retrievals. Is the above cloud column more representative to the above cloud in-situ measurement? Or what is going on here. Especially given the fact that the cloudy and clear sky columns by Mozaic are not so different in first place.*

changed To clarify this point, we adjusted the discussion in the document (p8,126-33)

“ The time series in Fig. 4 show strong outliers in the MOZAIC/IAGOS CO column densities, which are not found by SCIAMACHY. de Laat et al. (2012) explained this with strong local pollution at the airports, which affects the airborne measurements but are not seen by the satellite due to the coarse spatial resolution. Figure 6 confirms this reasoning. Applying the column averaging kernels of the cloudy-sky retrievals to MOZAIC/IAGOS measurements, the enhanced CO column values are strongly reduced for both Teheran and Beijing and brings the aircraft and satellite observations into much better agreement. This confirms that the column averaging kernel of the cloudy-sky retrieval describes well the cloud shielding of the lower atmosphere (see Fig. 2) where the local pollution is primarily located. ...”

to

“ de Laat et al. (2012) explained the larger errors for clear-sky observations with strong local pollution at the airports, which affects the airborne measurements but are not seen by the satellite due to the coarse spatial resolution. This is concert with the extremely large values of the MOZAIC/IAGOS CO column densities time series in Fig. 4. Hence, we expect a better agreement, when we consider SCIAMACHY retrievals for high clouds because the atmospheric shielding of the spatial heterogeneity of CO. Figs. 4 and 2 confirms this. Applying the column averaging kernels of the cloudy-sky retrievals to MOZAIC/IAGOS measurements significantly improve the comparison with SCIAMACHY retrievals and so supports the error interpretation by (de Laat et al., 2012) but also demonstrates the data quality of SCIAMACHY cloudy-sky retrievals.

Another interesting feature of our comparison is the better agreement for Beijing than for Teheran, where at the same time the SCIAMACHY ...”

25. *Section 3.1, page 9, Last Paragraph: An additional reason for the good agreement is also the better SNR at Windhoek due to higher surface reflectivity*

changed We agree and add the following text at (p9,115):

“We complete our SCIAMACHY CO validation that is based on MOZAIC/IAGOS aircraft measurements by analyzing observations from the airport at Windhoek. Figure 5 shows good agreement between SCIAMACHY and MOZAIC/IAGOS measurements with a clear seasonality. Here, the spatial CO distribution is less affected by local sources and so representation errors are less relevant. Moreover, the high surface albedo ensures clear-sky retrievals with high precision.”

26. *Section 3.2, page 9, line 23 and 24: Because of lack of... Is the collocation area in this case also a circle of 850km radius?* **changed** Yes, we use the same criteria as for MOZAIC as already mentioned in the manuscript (p9,28). To emphasize this, we change this sentence from: “ Using the same criteria as in Sec. 3.1 ...” to “Using the same criteria (i.e. collocation radius and quality filtering) as in Sec. 3.1 ...”

27. *How does the criterion of 30-day window compare with the dynamically chosen time window allowing an error of the mean of $< 1e17$?*

changed For clear-sky retrievals, this relation depends on the retrieval precision and so mainly on the surface albedo. Thus, for some NDACC/TCCON stations the dynamical averaging corresponds to a 7-day to 30-day averaging but for others it would be even averaging over years. We addressed this point in our answer to comment 22 of the reviewer and changed the manuscript accordingly.

28. *Section 3.2, page 10, line 6: >17 ppb: According the the Mauna Loa and Reunion plots, the ϵ_S amounts to values larger than 100 ppb. How is this number here then to be interpreted according to the plots?*

not changed The reviewer refers here to the scatter of individual retrievals (which can be > 100 ppb for low albedo sites) and the standard error of the mean bias averaged over 10 years. These are different diagnostic tools and cannot be compared one-by-one.

29. *Section 3 generally: Wouldnt it be an option the application of the AK to TM5 profiles? Of course, the null space error would remain.*

not changed No. The TM5 profiles is used as reference profile to be scaled by the inversion. Borsdorff et al. (2016) (Eq. (26)) showed that in this case $Ax_{TM5} = Cx_{TM5}$ and so the TM5 profile cannot be used to add null-space contributions.

30. *Conclusion, page 11, line 31: on biases: Are these biases shown in Fig. 6?*

changed We change the the statement at (p11,l29) form:

“Direct comparison of the MOZAIC/IAGOS CO columns estimated at Beijing, Teheran, and Windhoek with collocated SCIAMACHY clear-sky CO retrievals of the CO column showed a small bias of 0.5 – 9.5 ppb, in agreement with previous studies. However for cloudy SCIAMACHY observations, the bias exceeds 120.0 ppb for Teheran and 30.0 ppb for Beijing. ”

to

“Direct comparison of the MOZAIC/IAGOS CO columns estimated at Beijing, Teheran, and Windhoek with collocated SCIAMACHY cloudy-sky retrievals of the CO column show a small bias of 0.5 – 9.5 ppb. However, for clear-sky SCIAMACHY observations, the bias exceeds 120.0 ppb for Teheran and 30.0 ppb for Beijing, which is in agreement with previous studies (Borsdorff et al., 2016, de Laat et al., 2012).”

31. *Figure 4, Caption: If clouds are thick and high enough, the cases compared here represent basically the CO column from ToA down to the upper troposphere. Considering that the MOZAIC/IAGOS CO profiles are extended to ToA using TM5 data, in how far do the comparison shown here for thick high cloud cases - and after applying the SCIA AK really refer to a comparison between SCIA vs. MOZAIC/IAGOS and not SCIA vs TM5?*

changed We agree that the validation for high and optically thick clouds with MOZAIC/IAGOS profiles is subject to an additional uncertainty due to the flight height of the aircraft. As described in the manuscript (p8,l3) we only select airborne profiles, which at least reach an altitude of 9 km but overall with mean maximum altitude of 11 km. On the other hand, we select cloudy retrievals with an cloud altitude lower than 8 km and so there is a clear overlapping altitude range between the aircraft measurements and the sensitivity of the retrieval. To stress this point, we changed the following paragraph (p7,l29):

“In this study, we consider the vertical profile measurements at three cities: Teheran and Beijing, which are known to be affected by strong local pollution events (de Laat et al., 2012), and Windhoek, which is situated in an area with higher surface albedo and thus high radiometric precision of SCIAMACHY clear-sky measurements. Moreover, at Windhoek the atmospheric CO abundance is only slightly influenced by local pollution resulting in spatially homogeneous CO fields. Table 1 provides more details on these sites. From the MOZAIC/IAGOS data set, we select only those profiles that cover at least an altitude range up to 9 km with data gaps \leq 1 km.“ to

“In this study, we consider the vertical profile measurements at three cities: Teheran and Beijing, which are known to be affected by strong local pollution events (de Laat et al., 2012), and Windhoek where the atmospheric CO abundance is only slightly influenced by local emissions resulting in spatially homogeneous CO fields. The Windhoek measurement site is situated in an area with higher surface albedo and thus high radiometric precision of SCIAMACHY clear-sky measurements. Table 1 provides more details on these sites. From the MOZAIC/IAGOS data set, we select only those profiles that reach at least an altitude of 9 km with data gaps \leq 1 km. The mean maximal altitude of the selected profiles is 11 km, which overall indicates a sufficient overlap between the aircraft measurements and the sensitivity of the retrieval with cloud heights $<$ 8 km. ”

32. *Figure 10: To which extend is this plot dominated by the data of the "good SCIA SWIR years" 2003, 2004, 2005? Is the CO product displayed here considering only data from 2006 up to 2012 maintain the smoothness and the coverage properties of the given plot?*

not changed There is not a specific time range of “good SCIAMACHY years” because the detector degrades continuously with time. Figure. 7 and Fig. 8 show the time series of NDACC/TCCON validations for cloudy SCIAMACHY observation covering the full range of the mission with a good precision over the entire time range. Here, early SCIAMACHY years do not dominate the validation. So, we are reluctant to speak here about ‘good’ and ‘bad’ SCIAMACHY years. Particular, the year 2005, mentioned by the reviewer also in previous comments, can not be confirmed as any break point of the data quality.

33. *Abstract: line 9: This improves... -> The situation improves...*

corrected

34. *Section 2.1 line 7: $C = (1, \dots, 1)^T \rightarrow$ I would suggest to describe it in the form of: the transpose of the column vector or similar.*

changed After reflecting this formulation, we agree that it does not make to much sense to define here C explicitly as a row or column vector. To become independent on the specific representation of vectors, we adjusted the text as follows: “ $c = \vec{C}^T \vec{\rho}_{\text{ref}}$. Here, T indicates the transposed of the vector \vec{C} with all

elements $C_i = 1$. So Eq. (2) approximates the vertical integration assuming that the entries of the CO profiles are given in subcolumns.

Equation (1) is inverted with respect to state vector \vec{x} with the solution \vec{x}_{ret} using the least squares fitting approach, where we apply the Gauss-Newton algorithm to account for the non-linearity of the forward model. “

Mathematically the suggested formulation is equivalent to the formulation in the manuscript. We see no benefit in a reformulation.

35. *Section 2.1, line 12: Define gain matrix G for completeness and readability (all other quantities have been defined).*

modified The gain matrix is defined and discussed in detail in Borsdorff et al. (2014). Repeating the definition here would mean to increase the math in the manuscript substantially, without increasing the readability of the text, to our opinion. Therefore, we suggest to add the following sentence to the manuscript (p4, l26):

“Here \mathbf{G} is the gain matrix of the inversion and its definition is given by Borsdorff et al. (2014). ”

36. *Section 2.1 line19: an -> a*
corrected

37. *p.9, line 1: "...a larger representation errors..."*
corrected

38. *Section 2.2, page 5, line 9: clouds -> aerosol and clouds*
corrected

39. *Section 2.2, page 5 line 10: Moreover? I think this is the purpose in first place...*
corrected Moreover is removed.

40. *Section 2.2, page 5, line 16 and 17: To establish This sentence should probably go in the previous paragraph, since it is relevant for the retrieval of h_{cld} and τ_{cld} . But I guess the CH₄ profile is also fixed for the "physics-based" retrieval. Is it?*

not changed The paragraph describes the scattering (so physics-based) retrieval, where cloud information is inferred from CH₄ absorptions. Therefore, CH₄ is fixed to the TM5 profile. So, this sentence is correct and we think well placed.

41. *Section 3.1, page.8, line 30: "Thus confirms that...";*

corrected

42. *Section 3.2, page 9, line 20: air column -> dray-air column?*
corrected

43. *Figure 2, Caption: cloud at... -> cloud center height? cloud top height? Please specify.*
corrected It is the center height, corrected in the manuscript.

References

- Borsdorff, T., Hasekamp, O. P., Wassmann, A., and Landgraf, J.: Insights into Tikhonov regularization: application to trace gas column retrieval and the efficient calculation of total column averaging kernels, *Atmos. Meas. Tech.*, 7, 523–535, doi:10.5194/amt-7-523-2014, 2014.
- Borsdorff, T., Tol, P., Williams, J. E., de Laat, J., aan de Brugh, J., Nédélec, P., Aben, I., and Landgraf, J.: Carbon monoxide total columns from SCIAMACHY 2.3 m atmospheric reflectance measurements: towards full-mission data product (2003-2012), *Atmos. Meas. Tech.*, 9, 227–248, doi:10.5194/amt-9-227-2016, 2016.
- de Laat, A. T. J., Dijkstra, R., Schrijver, H., Nédélec, P., and Aben, I.: Validation of six years of SCIAMACHY carbon monoxide observations using MOZAIC CO profile measurements, *Atmos. Meas. Tech.*, 5, 2133–2142, doi:10.5194/amt-5-2133-2012, 2012.
- Landgraf, J., aan de Brugh, J., Scheepmaker, R., Borsdorff, T., Hu, H., Houweling, S., Butz, A., Aben, I., and Hasekamp, O.: Carbon monoxide total column retrievals from TROPOMI shortwave infrared measurements, *Atmos. Meas. Tech.*, 2016, 1–39, doi:10.5194/amt-2016-114, 2016.

Carbon monoxide column retrieval for clear-sky and cloudy atmospheres: a full-mission data set from SCIAMACHY 2.3 μm reflectance measurements

Tobias Borsdorff¹, Joost aan de Brugh¹, Haili Hu¹, Philippe Nédélec², Ilse Aben¹, and Jochen Landgraf¹

¹SRON Netherlands Institute for Space Research, Utrecht, the Netherlands

²Laboratoire d'aérodologie (LA), CNRS UMR-5560 et Observatoire Midi-Pyrénées, Université Paul-Sabatier, Toulouse, France

Correspondence to: T. Borsdorff (t.borsdorff@sron.nl)

Abstract. We discuss the retrieval of carbon monoxide (CO) vertical column densities from clear-sky and cloud contaminated 2311-2338 nm reflectance spectra measured by the Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (SCIAMACHY) from January 2003 until the end of the mission in April 2012. These data was processed with the Shortwave Infrared CO Retrieval algorithm SICOR that we developed for the operational data processing of the Tropospheric Monitoring Instrument (TROPOMI) that will be launched on ESA's Sentinel-5 Precursor (S5P) mission. This study complements previous work that was limited to clear-sky observations over land. Over the oceans, CO is estimated from cloudy-sky measurements only, which is an important addition to the SCIAMACHY clear-sky CO data set as shown by NDACC and TCCON measurements at coastal sites. For Ny-Ålesund, Lauder, Mauna Loa, and Reunion, a validation of SCIAMACHY clear-sky retrievals is not meaningful because of the high retrieval noise and the few collocations at these sites. ~~This~~ The situation improves significantly when considering cloudy-sky observations, where we find a low mean bias $\bar{b} = \pm 6.0$ ppb and a strong correlation between the validation ~~data set~~ and the SCIAMACHY ~~data sets~~ results with a mean Pearson correlation coefficient $r = 0.7$. Also for land observations, cloudy-sky CO retrievals present an interesting complement to the clear-sky data set, ~~which is less sensitive to the spatial representativeness of the satellite and validation measurement~~. For example, at the cities Teheran and Beijing the agreement of SCIAMACHY clear-sky CO observations with MOZAIC/IAGOS airborne measurements is poor with a mean bias of $\bar{b} = 171.2$ ppb and 57.9 ppb because of local CO pollution, which cannot be captured by SCIAMACHY. ~~The validation improves significantly for~~ For cloudy-sky retrievals ~~with $\bar{b} = 52.3$ ppb and 5.0 ppb, respectively. This is due to a reduced retrieval sensitivity to CO below,~~ the validation improves significantly. Here the retrieved column is mainly sensitive to CO above the cloud and so ~~to the altitude range, which is mostly affected by~~ not affected by the strong local surface emissions. Adjusting the MOZAIC/IAGOS measurements to the vertical sensitivity of the retrieval, the mean bias add up to $\bar{b} = 52.3$ ppb and 5.0 ppb, respectively. At the less urbanised region around the airport Windhoek, local CO pollution is less prominent and so MOZAIC/IAGOS measurements agree well with SCIAMACHY clear-sky retrievals with a mean bias of $\bar{b} = 15.5$ ppb, but can be even further improved considering cloudy SCIAMACHY observations with a mean ~~CO~~ bias of $\bar{b} = 0.2$ ppb. Overall the cloudy-sky CO retrievals from SCIAMACHY short wave infrared measurements present a ~~valuable addition to~~ major extension of the clear-sky only data set, which more than triples the amount of data and adds unique observations over the oceans. More-

over, the study represents the first application of the S5P algorithm for operational CO data processing on cloudy observations prior to the launch of the S5P mission.

1 Introduction

The Tropospheric Monitoring Instrument (TROPOMI) will be launched on board of the Copernicus Sentinel-5 Precursor (S5P) satellite. Besides the ultraviolet, visible and near infrared spectral range, it will measure Earth radiance and solar irradiances ~~irradiance~~ in the 2.3 μm short wave infrared (SWIR) spectral range during its expected lifetime of seven years (Veeffkind et al., 2012). The TROPOMI SWIR spectrometer builds upon the Scanning Imaging Absorption Spectrometer for Atmospheric Char-
tography (SCIAMACHY) that was operational from January 2003 to April 2012 on ESA's ENVISAT satellite (Bovensmann et al., 1999). The SWIR measurements of TROPOMI and SCIAMACHY will be similar with respect to spectral coverage and resolution, but TROPOMI will have better spatial resolution ($7\times 7\text{ km}^2$), a larger signal-to-noise ratio, and daily global coverage. Therefore TROPOMI will extend the record of SWIR measurements from space, and because of its similarity to SCIAMACHY, it is possible to use the same CO retrieval algorithm for both instruments and thereby provide a consistent data product for long-term study of CO (Borsdorff et al., 2016; Landgraf et al., 2016b).

The SWIR measurements around 2.3 μm of SCIAMACHY have been used to retrieve the vertical column densities of various atmospheric trace gases: HDO (Frankenberg et al., 2009; Scheepmaker et al., 2015), water vapor (Schrijver et al., 2009), and CO (Frankenberg et al., 2005; Buchwitz et al., 2004; Gloudemans et al., 2008; Gimeno García et al., 2011; Borsdorff et al., 2016). In particular, the SCIAMACHY CO product has been used to analyse biomass burning events (Buchwitz et al., 2004), monitor the transport of atmospheric pollutants (Gloudemans et al., 2006), and detect pollution from mega cities (Buchwitz et al., 2007).

~~The~~ A major limitation of the SCIAMACHY CO data set is its high retrieval noise, which can exceed 100 percent of the retrieved value for individual columns. Hence, in practice the data needs to be averaged spatially and temporally to reduce the noise contribution (de Laat et al., 2007; Gloudemans et al., 2006). The SCIAMACHY instrument has had issues with degradation and calibration in the SWIR spectral range, and efforts to retrieve CO are hampered by the failure of a significant number of detector pixels and the growth of a layer of ice on the detector array (Gloudemans et al., 2005, 2008). Borsdorff et al. (2016) discussed a re-calibration of the measurements and presented a full mission CO data set, which is restricted to clear-sky scenes over land. This SCIAMACHY CO data product for clear skies was validated with ground-based Fourier Transform Spectrometer (FTS) measurements ~~that were~~ provided by the Network for the Detection of Atmospheric Composition Change (NDACC) and the Total Carbon Column Observing Network (TCCON) (de Laat et al., 2010; Borsdorff et al., 2016), and it was also validated with airborne measurements provided by the MOZAIC/IAGOS project (de Laat et al., 2012; Borsdorff et al., 2016). The successful validation of this recalibrated data set showed that SCIAMACHY observations reveal meaningful information about atmospheric CO over the full mission period.

SCIAMACHY's SWIR measurements of clear skies over land have a good sensitivity for the vertical column density of CO. However, the clear-sky CO data product requires a strict cloud filter over land and the ~~wholesale~~ rejection of all observations

over oceans – due to the low reflectivity of the ocean surface in the SWIR. Hence, a retrieval that works for cloudy scenes is ~~necessary~~ very desirable in order to extend spatial coverage beyond the small fraction of SCIAMACHY observations ~~which are made of,~~ which are clear skies over land.

Gloude-mans et al. (2009) and Buchwitz et al. (2006) demonstrated that cloud contaminated measurements can be useful for ~~CO retrieval~~ the CO retrieval, because the high reflectivity of clouds decreases retrieval noise. By using SCIAMACHY observations under cloudy conditions, Gloude-mans et al. (2009) investigated the outflow of CO from Asia and Indonesia over the oceans. However, to properly interpret these measurements, one has to account for the shielding and scattering effect of clouds. For this purpose, Vidot et al. (2012) and Landgraf et al. (2016b) proposed the Shortwave Infrared CO retrieval (SICOR) algorithm for the operational processing of ~~CO data for~~ the S5P mission. In addition to trace gas absorption, SICOR estimates two cloud parameters: cloud optical thickness and cloud height. These parameters are used in a two stream radiative transfer solver ~~called~~ 2S-LINTRAN (~~Landgraf et al., 2016b~~) to account for ~~the changes to~~ changes to the light path due to light path ~~that occur due to~~ scattering by clouds and aerosols. ~~SICOR produces a data product that~~ For cloudy conditions, the CO error induced by this radiative transfer solver reaches occasionally several percent (Landgraf et al., 2016b) and becomes marginal for regional measurement ensembles. The SICOR data product includes the retrieved CO column, its noise estimate and the column averaging kernel. The averaging kernel reflects the vertical sensitivity of the retrieval to CO, and it depends on the cloud parameters that were estimated for the observed scene.

In this study, we apply the SICOR algorithm to ~~measurements from~~ SCIAMACHY's 2.3 μm SWIR ~~channel~~ measurements. Doing this allows us to evaluate the SICOR algorithm's performance on real cloud-contaminated measurements and to expand upon the data set produced by Borsdorff et al. (2016). Whereas that CO data set was limited to clear-sky retrievals over land, here we provide SCIAMACHY CO data for cloudy scenes over the oceans and both clear and cloudy scenes over land. The new data product includes ~~the~~ column averaging kernels for individual measurements, ~~and we use these to help~~ which helps to validate the data product with vertical concentration profiles of CO ~~that were measured~~ provided by the MOZAIC/IAGOS project. We also validate the data product with TCCON and NDACC ground-based FTS measurements. The paper is structured as follows. Section 2.1 introduces the inversion of vertical column densities by profile scaling, Sect. 2.2 explains the retrieval setup for the SCIAMACHY measurements, and Sect. 2.3 analyses the quality of the retrieved cloud parameters. In Sect. 3.1 we present the data validation using MOZAIC/IAGOS aircraft measurements while Sect. 3.2 focuses on the validation with NDACC and TCCON measurements. A summary and conclusions are given in Section 4.

2 Inversion method

The inversion of CO vertical column densities from SCIAMACHY's 2.3 μm reflectance spectra utilises the profile scaling approach. This approach has been applied to SCIAMACHY data before by Gloude-mans et al. (2008) and Borsdorff et al. (2016), and its regularisation type is discussed in detail by Borsdorff et al. (2014). Here, we summarise the aspects of ~~inversion theory~~ the retrieval that are needed for the later discussion.

2.1 Profile scaling approach

The inversion is based on the assumption that the ~~m -dimensional~~ measurement \mathbf{y}_{meas} can be described by the forward model \mathbf{F} in the bounds of the measurement error e_y , namely

$$\mathbf{y}_{\text{meas}} = \mathbf{F}(\mathbf{x}, \mathbf{b}) + e_y. \quad (1)$$

- 5 Here, the vector \mathbf{b} comprises forward model parameters, which are known a priori. The state vector \mathbf{x} contains all parameters to be retrieved including the column density of CO and other trace gases.

The forward model \mathbf{F} requires as input the vertical concentration profile of the atmospheric trace gases, which we obtain by scaling a reference profile ρ_{ref} with the corresponding column density of the trace gas in \mathbf{x} . The reference profile does not change during the retrieval and is normalised by its vertical column density c ,

$$10 \quad c = \mathbf{C}^T \rho_{\text{ref}}. \quad (2)$$

~~where the n -dimensional vector $\mathbf{C} = (1, \dots, 1)^T$~~ Here, T indicates the transposed of the vector \mathbf{C} with all elements $C_i = 1$. So Eq. (2) approximates the vertical integration assuming that the entries of the CO profiles are given in subcolumns.

Equation (1) is inverted with respect to state vector \mathbf{x} with the solution \mathbf{x}_{ret} using the least squares fitting approach, where we apply the Gauss-Newton algorithm to account for the non-linearity of the forward model. ~~In the linear approximation the~~

- 15 ~~solution \mathbf{x}_{ret} can be expressed by the gain matrix \mathbf{G} of the inversion,~~

$$\underline{\mathbf{x}_{\text{ret}}} = \mathbf{G} \mathbf{y}_{\text{meas}}.$$

The retrieved vertical column density c_{ret} is an element of the solution vector \mathbf{x}_{ret} and describes an effective column density due to the regularisation inherent to the profile scaling approach. In the linear approximation, the effective column and the true atmospheric abundance are related by the equation,

$$20 \quad c_{\text{ret}} = \mathbf{a}_{\text{col}} \rho_{\text{true}} + e_c, \quad (3)$$

where \mathbf{a}_{col} is the column averaging kernel, e_c is the column retrieval error due to the measurement error e_y and ρ_{true} is the true trace gas profile. Borsdorff et al. (2014) presented a numerically efficient algorithm for calculating \mathbf{a}_{col} . The column averaging kernel represents ~~an a~~ weighted integration of the true vertical profile taking into account the particular retrieval sensitivity. The differences between the true column, $c_{\text{true}} = \mathbf{C} \rho_{\text{true}}$, and ~~the effective column,~~ $c_{\text{eff}} = \mathbf{a}_{\text{col}} \rho_{\text{true}}$, cannot be inferred from the measurement and is also known as the null-space or smoothing error of the retrieval (Borsdorff et al., 2014; Rodgers, 2000),

$$25 \quad e_{\text{null}} = (\mathbf{C} - \mathbf{a}_{\text{col}}) \rho_{\text{true}}. \quad (4)$$

Hence, when considering c_{eff} as an estimate of the true column, data interpretation does not rely on the use of the total column averaging kernel but the null-space error e_{null} becomes part of the error budget. When c_{eff} is seen as an effective column with its vertical sensitivity described by the column averaging kernel, data analysis focuses more on the information that can be

inferred from the measurements and reduces the impact of the a priori choice of the reference profile. This requires a proper use of the total column averaging kernel as shown in Eq. 3 and the null-space error does not contribute to the error budget (see also Wassmann et al. (2015)).

Finally, the measurement noise described by the measurement covariance matrix \mathbf{S}_y introduces noise on our retrieval product, which is characterised by the retrieval noise covariance matrix,

$$\mathbf{S}_x = \mathbf{G}\mathbf{S}_y\mathbf{G}^T. \quad (5)$$

~~In this manner~~Here \mathbf{G} is the gain matrix of the inversion and its definition is given by Borsdorff et al. (2014). Therewith, we have defined all diagnostic tools for our retrieval. A detailed overview of the profile-scaling approach is given in Borsdorff et al. (2014).

10 2.2 Retrieval Settings

For SCIAMACHY CO retrievals, ~~the~~we use the recalibrated SCIAMACHY spectra including estimates of the measurement noise as described by Borsdorff et al. (2016). The SICOR algorithm settings are very similar to those of the TROPOMI CO data processor (Landgraf et al., 2016b, a) but with adaptations to account for the instrument degradation of the SCIAMACHY instrument (Borsdorff et al., 2016). It comprises two main processing steps: first a non-scattering retrieval for cloud detection and subsequently a ~~physics-based~~physics-based scattering retrieval to infer CO column abundances. ~~To account for the limited radiometric accuracy and precision of~~For the SCIAMACHY measurements, we choose the spectral window 2311-2338 nm ~~(Borsdorff et al., 2016)~~ for both retrieval steps. This represents a wider retrieval window compared to the TROPOMI settings, which are 2315-2324 nm for the non-scattering retrieval and 2324-2338 nm for the ~~physics-based~~physics-based scattering retrieval (Landgraf et al., 2016b). This is necessary to account for the instrument degradation e.g. the considerable loss of detector pixel by radiation damage and the light scattering of the ice-layer on the detector array of SCIAMACHY's channel 8 (Borsdorff et al., 2016). To account for radiometric errors of the SCIAMACHY measurements, we apply a temporal dependent radiometric correction derived from SCIAMACHY observations of the Sahara and reliable model prediction of methane over this area. Moreover, the measurement noise is estimated from regular eclipse measurements and both aspects are described by Borsdorff et al. (2016).

The settings for the non-scattering retrieval are identical with the one described by Borsdorff et al. (2016) and in first instance we retrieve the column density of CO, CH₄, H₂O, and HDO without accounting for atmospheric scattering and clouds in particular. ~~The~~In that case, the difference between the retrieved CH₄ column and the a priori CH₄ column from the chemical transport model TM5 (Williams et al., 2013, 2014) reflects the light-path effect of aerosols and clouds on the observations due to shielding and photon path enhancement by multi-scattering ~~(Vidot et al., 2012)~~. ~~Moreover for~~Vidot et al. (2012). For the subsequent retrieval of CO, we use the retrieved non-scattering CH₄ column to provide a first estimate for the cloud height. For this purpose we integrate the TM5 vertical CH₄ profile from the cloud center to the top of the model atmosphere and adjust the cloud height to match the retrieved non-scattering methane column.

The subsequent ~~physics-based~~physics-based scattering retrieval is described in detail by Landgraf et al. (2016b). It estimates the trace gas columns of CO, H₂O, and its isotopologue HDO simultaneously with the cloud height h_{cld} and cloud optical thickness τ_{cld}

assuming a fully overcast scene. To estimate the cloud parameters from the SWIR measurements, the atmospheric CH₄ vertical profile concentration is fixed to accurate model estimates provided by the TM5 model. The ~~physics-based-scattering~~ retrieval simulates a cloud as a horizontally homogeneous scattering layer with a triangular height profile in optical depth with a fixed half width of 1.5 km and infers the optical thickness and the cloud height from the methane absorption signal. In case of cloud heights < 3km the cloud profile is cut off at the Earth surface and is renormalized. The remaining a priori information and the molecular absorption spectroscopy is described by Borsdorff et al. (2016).

2.3 Evaluation of the retrieved cloud parameters

Compared to the previous work by Borsdorff et al. (2016), the new aspect of this study is the retrieval of CO column densities from cloudy SCIAMACHY measurements, where we infer cloud parameters from the 2.3 μm methane absorption band. The effect of clouds on the retrieved CO column density is described with the column averaging kernels for each individual measurement. Obviously the retrieved cloud properties are effective parameters and so do not necessarily represent the real atmospheric situation but describe clouds, cirrus and aerosols such that the simulated lightpath is sufficiently accurate to retrieve CO within the required accuracy. In that sense the retrieved cloud parameters depend on the forward model of the retrieval algorithm and the spectral fit window, which makes their verification difficult.

To gain ~~first~~-experience with the SICOR cloud parameters, we compared the retrieved cloud height h_{cld} with the SCIAMACHY cloud top height product, processed with the Fast Retrieval Scheme for Clouds from the Oxygen A Band (FRESCO) algorithm. In contrast to SICOR, this algorithm infers cloud top height and cloud fraction from the ~~SCIAMACHY the~~ Oxygen A band around 760 nm assuming an elevated Lambertian reflector as a cloud model (Wang et al., 2008). This comparison is particular worthwhile since SCIAMACHY's NIR measurements do not suffer from the formation of an ice layer on the detector module like the SWIR measurements (Gloude-mans et al., 2008). Figure 1 shows a scatter plot of the SICOR and FRESCO cloud heights for one year of SCIAMACHY observations in year 2003 for an ocean region with a latitude range from 30° S to 0° and a longitude region from 30° W to 0°. We only consider measurements with a significant cloudiness indicated by SICOR cloud optical depth of $\tau_{cld} > 1$. The cloud parameters are highly correlated with a Pearson correlation coefficient $r = 0.9$. A linear regression provides a slope of 1.01 and an intercept of 300 m, indicating a valuable cloud product from the SWIR measurements of SCIAMACHY. However, this Even for later years of the mission we find a similar agreement (e.g. $r = 0.8$, slope= 0.9, intercept=431 m for the year 2010). The remaining small difference between the retrieved cloud height and the FRESCO cloud top height may origin from the different cloud models used. For example, FRESCO simulates clouds as an elevated Lambertian surface, whereas SICOR fits the center height of the triangular height profile. This comparison should not hide the fact that the validity of the SICOR effective cloud parameters can only be demonstrated by comparing the SCIAMACHY CO data product for cloudy observations with independent validation measurements, which is the subject of the next section.

In the following, we use the SICOR cloud parameters to classify the retrievals with respect to cloudiness. We distinguish between three CO retrieval conditions: (1) clear-sky observations over land with $\tau_{cld} < 0.3$ and $h_{cld} < 1$ km, (2) cloudy observations contaminated by optically thick low clouds with $\tau_{cld} > 2$ and $h_{cld} < 1.5$ km, and (3) cloudy observations contaminated by

optically thick high clouds with $\tau_{cloud} > 2$ and $1.5 \text{ km} < h_{cloud} < 8 \text{ km}$. For these categories, Fig. 2 shows typical column averaging kernels selected from the SCIAMACHY CO retrievals over Australia for the year 2003.

Column averaging kernel values of one represent the ideal value for a vertical integration (see Eq. 2 (2)) and values < 1 indicate reduced retrieval sensitivity e.g. because of atmospheric shielding by clouds and aerosols. Values > 1 are typical for a profile scaling approach. In case of an optically thick cloud the ~~profile is adjusted at all altitudes including sensitivity of the measurement for~~ CO below the cloud ~~using only the measurement sensitivity is lost. The inversion scales the entire vertical profile of CO based on its sensitivity to CO~~ above the cloud. ~~This means that Hence,~~ the sensitivity of the retrieved CO column with respect to the CO concentration above the cloud is enhanced (> 1) and at the same time the sensitivity for CO below the cloud is reduced (< 1). These features are clearly reflected in Fig. 2 ~~and discussed in more detail by Borsdorff et al. (2014)~~.

So under clear-sky conditions we can retrieve CO vertical column densities with a sensitivity close to 1 at all altitudes (see yellow line in Fig. 2). Hence, the corresponding null-space error is small (Buchwitz et al., 2004; Gloudemans et al., 2008) and a direct comparison of the retrieved column with CO column reference measurements is possible while neglecting the effect of column averaging kernels (Borsdorff et al., 2016). For cloudy retrievals the comparison must account for averaging kernel effects. Clouds can easily lead to null-space errors $> 30 \%$ depending on the discrepancy between ~~the~~ true vertical trace gas profile and the reference profile to be scaled by the retrieval (Borsdorff et al., 2014).

~~Generally, the~~ The null-space ~~error does not limit the data use does not provide a problem for application like data assimilation,~~ if the averaging kernel is applied properly. ~~For example, for data assimilation the averaging kernel provides~~ Exploiting the altitude sensitivity ~~, which is needed to adequately adjust given by the averaging kernel,~~ the atmospheric state ~~of the model can be adjusted adequately by the assimilation scheme.~~ Also for the validation of the CO data product, the null-space error does not impose principle limitation. ~~Using, where~~ Eq. 3, ~~we can account~~ (3) ~~accounts~~ for cloud effects in our retrieval when comparing satellite measurements with independent profile soundings. Wassmann et al. (2015) demonstrated this strategy for ozone column retrievals from GOME-2 measurements using ozonesonde measurements for validation. In practice, ~~however,~~ this approach is often hampered by the small number of available validation measurements.

CO retrievals ~~for under~~ cloudy conditions represent an interesting addition ~~of our data product because of the different vertical sensitivity to our clear-sky data product. Here,~~ the retrieved ~~CO column probes different altitude ranges and so providing complementary information. Additionally~~ column is mostly sensitive to CO above the cloud and for varying cloud height and due to the shielding effect of clouds, different altitudes are probed. For the interpretation of errors, it is important to note that for the scaling of a reference profile, the interpretation of relative biases is the same for clear-sky and cloudy retrieval. In both cases, it indicates the relative error to the CO concentrations at all altitudes the retrieval is sensitive to. Furthermore, the radiometric precision of SWIR measurements for cloudy scenes is generally much better than for clear-sky observations because of the brightness of clouds. In the $2.3 \mu\text{m}$ spectral range this represents an important asset for data exploitation, particularly for SCIAMACHY, which suffers from a poor radiometric precision. Also for TROPOMI, with a much better radiometric performance, CO retrievals over cloudy scenes represent an important addition for ocean scenes where due to the low reflectivity of ocean water no information about atmospheric CO can be inferred from clear-sky observations.

To illustrate this, Fig. 3 shows the signal-to-noise ratio (SNR) and CO retrieval noise as a function of the SICOR cloud optical thickness τ_{cld} for an ocean region with a latitude range from 30° S to 0° and a longitude range from 30° W to 0° . For a cloudless scene " $\tau_{cld} \rightarrow 0$ ", the signal-to-noise is less than 25 and the CO retrieval error is very high. Already for SICOR cloud optical thickness of 1-2 the SNR rises above 100 and the CO retrieval noise reduces to $5 \cdot 10^{17}$ molec cm^{-2} . For ~~higher-larger~~ cloud optical thickness, the SNR and CO precision saturate around $\tau_{cld} = 8$ where the SNR reaches 250 and the corresponding CO retrieval noise is about $2 \cdot 10^{17}$ molec cm^{-2} .

Figure 4 shows that much more data becomes available when considering cloudy-sky retrievals under optical thick low and high cloud conditions in addition to clear-sky retrievals. Between January 2003 to the end of the SCIAMACHY mission in April 2012, 70 % of all data are inferred from cloudy-sky observations (40% over the oceans and 30% over land). Over the oceans only cloudy-sky retrievals are possible but over land the fraction of clear-sky to cloudy-sky retrievals varies. For example, the data coverage over the Sahara region is dominated by clear-sky retrievals while the Amazon region and Indonesia show cloudy-sky retrievals be in a majority. Averaged globally, 55% of all measurements over land are cloudy-sky retrievals. Hence for SCIAMACHY, our CO retrieval algorithm for cloudy conditions permits the ~~derivation of CO~~ CO data product over the ocean, and in addition has the potential to improve the ~~CO~~ product over land.

15 3 Validation

The quality of our CO retrieval for cloudy atmospheres needs to be demonstrated through validation. ~~Here we begin by validating~~ First, we validate the SCIAMACHY CO retrievals with airborne profile measurements of the MOZAIC/IAGOS project, based on in-situ measurements on commercial airliners (Nédélec et al., 2003). Following this, we complete our validation using CO ground-based measurements of the TCCON and NDACC networks. To ensure the quality of the SCIAMACHY CO column retrievals, we perform an a posteriori quality filter on the individual CO retrievals:

1. The χ^2 values of the spectral fit must be < 40 .
2. The mean signal-to-noise ratio of the measurements in the fit window must be > 20 .
3. The noise ϵ of the retrieved CO, HDO and H₂O column must be below an upper threshold, namely $\epsilon_{\text{CO}} < 1 \times 10^{19}$ molec cm^{-2} , $\epsilon_{\text{HDO}} < 1 \times 10^{20}$ molec cm^{-2} , and $\epsilon_{\text{H}_2\text{O}} < 4 \times 10^{22}$ molec cm^{-2} .

These criteria represent a weak data filtering to remove outliers due to unphysical retrievals. For example, the signal-to-noise threshold corresponds to clear-sky measurements over the oceans (see Fig 3) for which a stable inversion is not possible. The retrieval would result in unphysical outliers and are thus rejected by the filter.

3.1 Airborne CO profile measurements (MOZAIC/IAGOS)

The MOZAIC/IAGOS project provides profile measurements of reactive gases performed on board of long-distance passenger airliners. Since 1994, in-situ profile measurements are ~~performed~~ recorded during the ascent and descent phases of more than

40,000 flights. These observations are used to derive vertical column densities of CO with a precision of about 5 % (Nédélec et al., 2003). The recorded profiles do not represent a strict vertical intersection of the atmosphere but at lower altitudes atmospheric constituents are measured close to the airport where at higher altitudes the distance to the airport may reach 200-400 km. ~~The study by~~ de Laet et al. (2014) showed that in most cases the corresponding collocation errors with SCIAMACHY observations imposed by the aircraft ascent and descent phases average out to a large extent when ~~analysing~~ analyzing temporal averages.

In this study, we consider the vertical profile measurements at three cities: Teheran and Beijing, which are known to be affected by strong local pollution events (de Laet et al., 2012), and Windhoek, ~~which where the atmospheric CO abundance is only slightly influenced by local emissions resulting in spatially homogeneous CO fields.~~ The Windhoek measurement site is situated in an area with higher surface albedo and thus high radiometric precision of SCIAMACHY clear-sky measurements. ~~Moreover, at Windhoek the atmospheric CO abundance is only slightly influenced by local pollution resulting in spatially homogeneous CO fields.~~ Table 1 provides more details on these sites. From the MOZAIC/IAGOS data set, we select only those profiles that ~~cover reach~~ at least an altitude ~~range up to of~~ 9 km with data gaps < 1 km. ~~This follows the procedure of de Laet et al. (2014) and ensures the availability of sufficient CO profile information to validate our data product.~~ ~~The mean maximal altitude of the selected profiles is 11 km, which overall indicates a sufficient overlap between the aircraft measurements and the sensitivity of the retrieval with cloud heights < 8 km.~~ Above the maximum flight altitude, the CO profiles are extended by simulated CO profiles of the TM5 chemical transport model. Furthermore, towards the ground level we extend the profiles assuming a constant CO mixing ratio, ~~which ensures the numerical stability of the extrapolation.~~ When the CO column density is derived from the aircraft measurement in this manner its accuracy is estimated to be a few percent (de Laet et al., 2012). ~~Numerical experiments using the lower data points of the profile to estimate the gradient in the CO concentration for linear extrapolation led to instabilities and thus this approach is not used here.~~

To collocate SCIAMACHY retrievals with MOZAIC/IAGOS measurements, we select all SCIAMACHY CO retrievals with a ground pixel in a radius of 850 km around the airport site. The temporal collocation criterium is chosen dynamically for each individual MOZAIC/IAGOS measurement ~~and varies typically between 7 and 30 days.~~ Centered around the recording time of the aircraft measurements, the temporal collocation window is chosen such that the average of all spatially collocated SCIAMACHY measurements have a precision < 10^{17} molec cm⁻². Therewith, we ensure that the precision of the SCIAMACHY validation measurements is sufficient for our purpose. ~~We~~ ~~For the validation with MOZAIC/IAGOS observations, we~~ consider SCIAMACHY measurements for clear-sky and cloudy conditions with high, optically thick clouds corresponding to condition (1) and (3) defined in Sec.2.3 and calculate the collocated MOZAIC/IAGOS CO column density from the ~~corresponding profile profiles~~ in two ways: first we perform the vertical integration of the MOZAIC/IAGOS profile after accounting for elevation differences between the SCIAMACHY ground scene and the airport site. Second, we apply the column averaging kernels of all collocated SCIAMACHY observations to the MOZAIC/IAGOS profile and calculate their average. Finally, we average the collocated SCIAMACHY CO column densities accordingly. ~~The two approaches corresponds to the different interpretation of the SCIAMACHY CO product as discussed in Sec. 2.~~ Figures 5 and 6 show ~~these corresponding~~ time series for the three airport sites. Per site, we calculate the mean and standard deviation of the CO time series as a diagnostic tool, which is ~~shown~~

on-depicted in Fig. 7 for the clear-sky and cloudy SCIAMACHY data sets. The figure clearly shows that applying the SCIAMACHY column averaging kernels to the MOZAIC/IAGOS profiles improves the validation for all sites for both clear-sky and cloudy-sky conditions. For clear-sky measurements, this bias improvement is small and is about 9.5 ppb for Teheran, 4.0 ppb for Beijing, and only 0.5 ppb for Windhoek, which confirms the findings by Borsdorff et al. (2014); Gloudemans et al. (2005); Buchwitz et al. (2004). For cloudy-sky measurements, we notice a significant improvement of the bias, 122.2 ppb for Teheran, 22.6 ppb for Beijing, and 7.3 ppb for Windhoek.

The time series in Fig. 5 show strong outliers in the MOZAIC/IAGOS CO column densities, which are not found by SCIAMACHY. de Laat et al. (2012) explained this. de Laat et al. (2012) explained the larger errors for clear-sky observations with strong local pollution at the airports, which affects the airborne measurements but are not seen by the satellite due to the coarse spatial resolution. Figure 5 shows this. This is consistent with the extremely large values of the MOZAIC/IAGOS CO column densities time series in Fig. 5. Hence, we expect a better agreement, when we consider SCIAMACHY retrievals for high clouds because the atmospheric shielding of the spatial heterogeneity of CO. Figs. 5 and 7 confirm this reasoning. Applying the column averaging kernels of the cloudy-sky retrievals to MOZAIC/IAGOS measurements, the enhanced CO column values are strongly reduced for both Teheran and Beijing and brings the aircraft and satellite observations into much better agreement. Thus confirms that the column averaging kernel of the significantly improve the comparison with SCIAMACHY retrievals and so supports the error interpretation by (de Laat et al., 2012) but also demonstrates the data quality of SCIAMACHY cloudy-sky retrieval describes well the cloud shielding of the lower atmosphere (see Fig. 2) where the local pollution is primarily located.

Another interesting feature of the SCIAMACHY cloudy-sky retrievals is that the agreement between SCIAMACHY and MOZAIC/IAGOS measurements integrated by the column averaging kernel is better. Our comparison is the better agreement for Beijing than for Teheran. At the same time the SCIAMACHY mean CO column mixing ratio at Teheran is about 43.0 ppb smaller than for Beijing with about 16.0 ppb less scatter. This may hint at a larger representation errors for Teheran than for Beijing, meaning that the aircraft and satellite observations represent different air masses due to their different spatial and temporal sampling and averaging. Beijing with 18.6 million urban inhabitants and an urban area of about has about double the size of Teheran with an urban population of about 8.8 million and an urban area of . Moreover the metropolitan area in Beijing is much larger than that of Teheran, indicating that the CO sources in Beijing may be spatially more extended than those in Teheran. This would mean that SCIAMACHY with its coarse spatial sampling captures the enhanced CO concentrations over Beijing better than those over Teheran and would explain the differences we see in Fig. 7 for both sites.

We complete our SCIAMACHY CO validation that is based on MOZAIC/IAGOS aircraft measurements by analysing observations from the airport at Windhoek. Figure 6 shows good agreement between SCIAMACHY and MOZAIC/IAGOS measurements with a clear seasonality. Here, the spatial CO distribution is less affected by local sources and so representation errors are less relevant. Moreover, the high surface albedo ensures clear-sky retrievals with high precision. We notice a small improvement when applying the SCIAMACHY cloudy-sky averaging kernels to the aircraft measurements, which removed

~~some high values in the time series.~~ The mean and standard deviation in Fig. 7 show that SCIAMACHY cloudy-sky and clear-sky retrievals agree well with the MOZAIC/IAGOS data for Windhoek.

At this site the reference profile ρ_{ref} from the TM5 model is more accurate and hence applying the averaging kernels to the MOZAIC/IAGOS validation measurements is less important. ~~Furthermore, the spatial CO distribution is less affected by strong~~

5 ~~local sources, which reduces representation errors of both data sets~~ Consequently, the fact that clear-sky and cloudy retrievals result in the same bias indicates that the accuracy for CO is comparable at altitudes the measurements are sensitive to.

3.2 Ground-based Fourier Transform Spectrometers (NDACC/TCCON)

The TCCON and NDACC networks perform direct sunlight measurements with ground based Fourier transform spectrometers under clear-sky conditions. The Infrared Working Group (IRWG) performs measurements in the mid-infrared spectral range at 4.8 μm and is part of NDACC (<http://www.ndsc.ncep.noaa.gov/>). Thereby, it supplies CO vertical column densities that we transformed to column averaged mixing ratios by calculating the ~~air-dry-air~~ column from the surface pressure at the station sites. The TCCON instruments measure in the same spectral range that SCIAMACHY does ~~with the aim of deriving and~~ provide the vertical column densities of trace gases including CO ~~to~~ with high precision (Wunch et al., 2010; Wunch et al., 2011).

15 Because of the lack of profile information, it is not possible to apply averaging kernels, and NDACC/TCCON data must be compared directly with the SCIAMACHY cloudy-sky retrievals. The comparison depends critically on the CO a priori profile information, and so we select near-coast measurement sites in remote areas, where the TM5 model prediction of the relative CO profile is most reliable: Ny-Ålesund, Wollongong (Griffith et al., 2014), Lauder (Sherlock et al., 2014), Mauna Loa, and Reunion . More details about the stations are given in Table 2. Figures 8 and 9 show time series of 30-day median values
20 derived from the ground-based FTS measurements and the collocated SCIAMACHY retrievals. Using the same criteria (i.e. collocation radius and quality filtering) as in Sec. 3.1, missing values of the ground-based FTS time series were interpolated (open circles) using the TM5 model as described by Borsdorff et al. (2016).

The upper panel of Fig. 8 illustrates that data validation ~~with using~~ ground-based measurements ~~sites~~ on small islands like Mauna Loa and Reunion are challenging for the clear-sky SCIAMACHY CO retrievals over land. Only a few satellite land
25 observations can be collocated with the ground-based measurements. Moreover, these measurements have low surface albedo, which results in a large retrieval noise > 100.0 ppb for single soundings. So for these locations, only a few 30-day median values can be determined, which are still dominated by the retrieval noise as shown in the left panel of Fig. 10. The scatter of individual SCIAMACHY retrievals is described by the half difference of the 15.9th and the 84.1th percentile, ~~which is written~~ as ϵ_S and is taken to be an analogue for the standard deviation of a normal distribution. Further, ϵ_N is the mean retrieval noise
30 of the single measurements of the monthly medians.

The correlations between the FTS and clear sky SCIAMACHY measurements over land are poor (about $r = 0.3$) and the high standard error (> 17.0 ppb) indicates that under these conditions even the overall mean value is difficult to estimate. The situation improves for measurement sites at larger islands like Lauder and Ny-Ålesund. We found more collocations between SCIAMACHY land observations and FTS measurements (Fig. 9) which led to smaller standard error of the mean (< 8.0 ppb),

and so a more reliable estimation of the overall mean. However, the 30-day median values are still dominated by the retrieval noise due to the low surface albedo of the observed scenes, and so the correlations between the FTS and SCIAMACHY measurements are still low (about $r = 0.3$, see Fig. 10). For SCIAMACHY clear-sky measurements over land, we achieved the best results for FTS measurements at Wollongong with high surface albedo. The upper panel of Fig. 9 shows that in this case the 30-day CO medians agree well with the FTS measurements. Moreover, the global mean can be estimated with a small standard error and the Pearson correlation coefficient for FTS and SCIAMACHY measurements is $r = 0.5$ indicating that the CO seasonality of the FTS measurements can be captured reasonably well by the SCIAMACHY observations.

When ~~analysing~~ analyzing SCIAMACHY cloudy-sky CO retrievals with ~~$1.5 \text{ km} < h_{\text{cld}} < 8 \text{ km}$~~ , 1.5 km and $\tau_{\text{cld}} > 2$ corresponding to condition (2) defined in Sec. 2.3, the comparison improves for all stations (see lower panel of Figs. 8 and 9). Due to the high reflectivity of clouds, the single measurement retrieval noise is reduced to < 30.0 ppb and much more ground-pixels can be collocated with SCIAMACHY. Figure 10 shows that the standard error of the mean bias falls below 2.0 ppb. For Lauder, the mean bias improves from $\bar{b} = -12.0$ to $\bar{b} = -2.0$ ppb but stays nearly the same for Ny-Ålesund. In both cases, the correlation increases to $r = 0.6$ and 0.7 respectively. For Wollongong, the mean bias for clear-sky and cloudy-sky retrievals is comparable (about 4.0 ppb), ~~but~~ and the correlation improves from about $r = 0.5$ to 0.8 for the cloudy retrievals. ~~Considering~~ Here the majority of the cloudy measurements are observations over the oceans and so these results indicate that there is no evidence for a land-sea bias of the SCIAMACHY CO retrieval because considering all reported cloudy-sky retrievals, we estimate a mean bias $\bar{b} = -6.0$ ppb. This may be due to remaining effects of the retrieval noise or due to the loose collocation criteria. Also issues due to erroneous radiometric calibration of SCIAMACHY SWIR measurements cannot be excluded.

The presented validation analysis provided sufficient confidence to re-process the global full-mission SCIAMACHY CO data set including clear-sky and cloudy-sky observations. In particular, the validation with NDACC and TCCON measurements revealed that the CO column, retrieved from measurements with low, optically thick clouds, can be used as an estimate of the true column (i.e. ignoring effects of the column averaging kernel) in absent of local sources of CO. The full-mission averaged CO product is shown in Fig. 11 and is based on clear-sky and cloudy observation contaminated by optical thick low clouds ($h_{\text{cld}} < 1.5 \text{ km}$). This figure indicates that the cloudy-sky ocean retrieval adds valuable information to the clear-sky data product. For example, over the Atlantic Ocean the CO outflow due to biomass burning in central Africa becomes clearly visible as well as the CO transport over the Pacific Ocean due to pollution from China and the wildfires in Indonesia. All this reveals information about the global transport of pollution in the atmosphere. Future research must demonstrate the merit of this observations for studying global transport of pollution in the Earth atmosphere.

4 Conclusions

In this study, we derived a full mission SCIAMACHY CO column data set that comprises retrievals from clear-sky and cloudy-sky reflectance measurements in the spectral range 2311-2338 nm over land and ocean scenes. The inversion uses the SICOR CO retrieval code that is developed for the operational data processing of the Sentinel 5 precursor mission. It allows us to retrieve effective cloud parameters simultaneously with trace gas columns. The data product includes the CO column and its

column averaging kernel for each individual sounding, and so it provides information on the vertical CO retrieval sensitivity, which changes with the cloudiness of the observed scene. This study focused on the validation of SCIAMACHY CO retrievals for cloudy-sky observations with MOZAIC/IAGOS aircraft measurements and TCCON and NDACC ground based observations. It represents an ~~extension to cloudy skies of~~ significant extension with respect to the previous work by Borsdorff et al.

5 (2016), presenting a similar data set for clear-sky scenes over land.

The effective cloud parameters (cloud height h_{cld} , cloud optical thickness τ_{cld}) are needed to properly simulate the atmospheric lightpath as part of the CO retrieval. Generally, the validation of effective cloud parameters is difficult, and in this study the cloud parameters are validated implicitly by the validation of the CO data product from cloudy-sky SCIAMACHY observations with independent atmospheric measurements. To gain confidence in the cloud product, we compared the cloud
10 height of the SICOR data product from one year of SCIAMACHY observations over the ocean with the corresponding cloud height of the FRESCO data product, which is inferred from SCIAMACHY measurements of the O₂ A band. We found a strong correlation ($r = 0.9$) between both products and an overall offset of the SICOR cloud height versus FRESCO of 300 m. Even for later years of the mission e.g. 2010 we find a similar agreement ($r = 0.8$, slope = 0.9, intercept = 431 m). This agreement shows that the cloud height retrievals from the different spectral ranges of SCIAMACHY are consistent even though the SWIR
15 measurements are affected by radiometric degradation of the instrument.

SICOR uses the profile scaling approach to infer CO vertical column densities from the measurement, which involves a ~~regularisation~~ regularization of the inversion problem. When interpreting the retrieved column as an estimate of the truth, a null-space error is introduced. This null-space error depends on the accuracy of the profile shape to be scaled by the inversion. Generally for clear-sky observations the null-space error is small, but for cloudy conditions it can easily exceed 30 percent.
20 Here, clouds shield the atmosphere below and the a priori CO profile shape is used to add the lacking information. The sensitivity of the retrieved CO column with respect to the true CO density is provided by the column averaging kernel for each individual retrieval. For the validation, we can thus interpret the retrieved column as a vertically integrated CO column density weighted by the column averaging kernel and so the null-space error becomes less relevant for the comparison of the retrieval product with independent atmospheric measurements of the CO profile.

25 Validating SCIAMACHY retrievals with MOZAIC/IAGOS airborne measurements confirmed the approach. Direct comparison of the MOZAIC/IAGOS CO columns estimated at Beijing, Teheran, and Windhoek with collocated SCIAMACHY ~~clear-sky CO cloudy-sky~~ retrievals of the CO column ~~showed show~~ a small bias of 0.5 – 9.5 ppb, ~~in agreement with previous studies. However for cloudy-~~ However, for clear-sky SCIAMACHY observations, the bias exceeds 120.0 ppb for Teheran and 30.0 ppb for Beijing which is in agreement with previous studies (Borsdorff et al., 2016; de Laat et al., 2012). This indicates
30 the large relevance of the null-space error for the comparison of cloudy observations with other data. The difference became much smaller when we accounted for the retrieval sensitivity in the comparison by applying the column averaging kernels to the airborne measurements. Here the bias reduced to 50.0 ppb for Teheran and 8.0 ppb for Beijing. The remaining error for Teheran we attribute to the different sampling of different air masses by the satellite and aircraft measurements. For Beijing with spatially much more extended CO surface emission, these errors are less relevant. For Windhoek, the comparison of air-

craft and satellite observations shows differences of 0.2 ppb when the averaging kernel is applied and 8.0 ppb otherwise. For this site, the vertical shape of the a priori CO profile is well known and so explains the smaller bias for this site.

We completed our validation study using ground-based FTS measurements from the NDACC and TCCON networks for the coastal sites Ny-Ålesund, Lauder, Mauna Loa, Reunion and Wollongong. At these validation sites, independent measurement of the CO profile is not available, and so we assumed a certain shape of the CO profile. Hence, we compare directly the column densities from FTS with those of SCIAMACHY measurements for both clear-sky conditions and cloudy conditions with low clouds. Considering only clear-sky SCIAMACHY observations over land at Lauder, Mauna Loa, and Reunion, the comparison was dominated by the SCIAMACHY retrieval noise because of the low surface albedo and the insufficient number of collocations. Filtering the SCIAMACHY data set to select for optically thick low clouds over land and oceans, we found good agreement with the FTS ground based measurements with a correlation coefficient of $r = 0.6 - 0.7$.

~~For Wollongong~~ For Wollongong, the high surface albedo led to CO retrievals from SCIAMACHY clear-sky measurements that compare well with FTS ground based measurements with a mean bias of $\bar{b} = -4.3$ ppb and a correlation of $r = 0.5$. Even for this site, the cloudy retrievals improved the comparison with a high correlation of $r = 0.8$. The mean bias stayed the same within its uncertainty, which demonstrated the overall consistency of the clear-sky and cloudy-sky data product.

Finally we processed the full-mission data record of SCIAMACHY SWIR measurements. This data-set demonstrates clearly the asset of cloudy-sky CO retrievals over oceans providing a global coverage of the SCIAMACHY CO data product. The correct interpretation of these data requires the use of the column averaging kernel. For varying cloud height and due to the shielding of the atmosphere below the cloud, the data may be used in future work to discriminate the vertical distribution of CO in the atmosphere, e.g. by means of data assimilation.

It is the first time that the operational TROPOMI CO algorithm is tested successfully on real data for clear-sky and cloudy atmospheres, which is an important milestone for the preparation of the S5P mission. Although TROPOMI SWIR measurements will have the same spectral coverage and resolution as SCIAMACHY, we expect a much better data product due to its better radiometric performance and its better spatial resolution and sampling. After launch of SP5 this techniques used in this study will allow us to use IAGOS/MOZAIC and FTS measurements by the TCCON and NDACC networks for the validation of TROPOMI CO. Additionally, using the same retrieval approach for SCIAMACHY and TROPOMI CO retrievals will help to provide a consistent long-term CO data set for both missions.

5 Data availability

The full-mission SCIAMACHY CO data set of this study including clear-sky and cloudy-sky observations is available for download at ftp://ftp.sron.nl/pub/pub/DataProducts/SCIAMACHY_CO/. The underlying data of the figures presented in this publication can be found at <ftp://ftp.sron.nl/open-access-data/>. TCCON data is accessible from <http://tcon.ornl.gov/> and NDACC measurements from <http://www.ndsc.ncep.noaa.gov/>. MOZAIC/IAGOS profiles can be download from <http://www.iagos.fr>.

Acknowledgements. SCIAMACHY is a joint project of the German Space Agency DLR and the Dutch Space Agency NSO with contribution of the Belgian Space Agency. The work performed is (partly) financed by NSO through the SCIAvisie project and in collaboration with the SCIAMACHY Quality Working Group (SQWG) by ESA. We acknowledge the European Commission for the support to the MOZAIC project (1994–2003) and the preparatory phase of IAGOS (2005–2012), the partner institutions of the IAGOS Research Infrastructure (FZJ, DLR, 5 MPI, KIT in Germany, CNRS, CNES, Météo-France in France and the University of Manchester in UK), ETHER (CNES-CNRS/INSU) for hosting the database, the participating airlines (Lufthansa, Air France, Austrian, China Airlines, Iberia, Cathay Pacific) for the transport free of charge of the MOZAIC/IAGOS instrumentation. MACC data were obtained from <http://www.copernicus-atmosphere.eu>. We acknowledge the NDACC-IRWG and TCCON ground-based FTS networks for providing data. This work was carried out on the Dutch national e-infrastructure with the support of SURF Cooperative.

10

References

- Borsdorff, T., Hasekamp, O. P., Wassmann, A., and Landgraf, J.: Insights into Tikhonov regularization: application to trace gas column retrieval and the efficient calculation of total column averaging kernels, *Atmos. Meas. Tech.*, 7, 523–535, doi:10.5194/amt-7-523-2014, 2014.
- 5 Borsdorff, T., Tol, P., Williams, J. E., de Laat, J., aan de Brugh, J., Nédélec, P., Aben, I., and Landgraf, J.: Carbon monoxide total columns from SCIAMACHY 2.3 μm atmospheric reflectance measurements: towards: full-mission data product (2003-2012), *Atmos. Meas. Tech.*, 9, 227–248, doi:10.5194/amt-9-227-2016, 2016.
- Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V. V., Chance, K. V., and Goede, A. P. H.: SCIAMACHY: Mission Objectives and Measurement Modes, *J. Atmos. Sci.*, 56, 127–150, doi:10.1175/1520-0469(1999)056<0127:SMOAMM>2.0.CO;2, 10 1999.
- Buchwitz, M., de Beek, R., Bramstedt, K., Noël, S., Bovensmann, H., and Burrows, J. P.: Global carbon monoxide as retrieved from SCIAMACHY by WFM-DOAS, *Atmos. Chem. Phys.*, 4, 1945–1960, doi:10.5194/acp-4-1945-2004, 2004.
- Buchwitz, M., de Beek, R., Noël, S., Burrows, J. P., Bovensmann, H., Schneising, O., Khlystova, I., Bruns, M., Bremer, H., Bergamaschi, P., Körner, S., and Heimann, M.: Atmospheric carbon gases retrieved from SCIAMACHY by WFM-DOAS: version 0.5 CO and CH₄ and impact of calibration improvements on CO₂ retrieval, *Atmospheric Chemistry & Physics*, pp. 2727–2751, 2006.
- 15 Buchwitz, M., Khlystova, I., Bovensmann, H., and Burrows, J. P.: Three years of global carbon monoxide from SCIAMACHY: comparison with MOPITT and first results related to the detection of enhanced CO over cities, *Atmos. Chem. Phys.*, 7, 2399–2411, doi:10.5194/acp-7-2399-2007, 2007.
- de Laat, A. T. J., Gloudemans, A. M. S., Aben, I., Krol, M., Meirink, J. F., van der Werf, G. R., and Schrijver, H.: Scanning Imaging Absorption Spectrometer for Atmospheric Chartography carbon monoxide total columns: Statistical evaluation and comparison with chemistry transport model results, *J. Geophys. Res. Atmo.*, 112, D12 310, doi:10.1029/2006JD008256, 2007.
- 20 de Laat, A. T. J., Gloudemans, A. M. S., Schrijver, H., Aben, I., Nagahama, Y., Suzuki, K., Mahieu, E., Jones, N. B., Paton-Walsh, C., Deutscher, N. M., Griffith, D. W. T., De Mazière, M., Mittermeier, R. L., Fast, H., Notholt, J., Palm, M., Hawat, T., Blumenstock, T., Hase, F., Schneider, M., Rinsland, C., Dzhola, A. V., Grechko, E. I., Poberovskii, A. M., Makarova, M. V., Mellqvist, J., Strandberg, A., Sussmann, R., Borsdorff, T., and Rettinger, M.: Validation of five years (2003-2007) of SCIAMACHY CO total column measurements using ground-based spectrometer observations, *Atmos. Meas. Tech.*, 3, 1457–1471, doi:10.5194/amt-3-1457-2010, 2010.
- 25 de Laat, A. T. J., Dijkstra, R., Schrijver, H., Nédélec, P., and Aben, I.: Validation of six years of SCIAMACHY carbon monoxide observations using MOZAIC CO profile measurements, *Atmos. Meas. Tech.*, 5, 2133–2142, doi:10.5194/amt-5-2133-2012, 2012.
- de Laat, A. T. J., Aben, I., Deeter, M., Nédélec, P., Eskes, H., Attié, J.-L., Ricaud, P., Abida, R., El Amraoui, L., and Landgraf, J.: Validation of nine-years of MOPITT V5 NIR using MOZAIC/IAGOS measurements: biases and long term stability, *Atmos. Meas. Tech. Discuss.*, 7, 5251–5291, doi:10.5194/amtd-7-5251-2014, 2014.
- 30 Frankenberg, C., Platt, U., and Wagner, T.: Retrieval of CO from SCIAMACHY onboard ENVISAT: detection of strongly polluted areas and seasonal patterns in global CO abundances, *Atmos. Chem. Phys.*, 5, 1639–1644, doi:10.5194/acp-5-1639-2005, 2005.
- Frankenberg, C., Yoshimura, K., Warneke, T., Aben, I., Butz, A., Deutscher, N., Griffith, D., Hase, F., Notholt, J., Schneider, M., Schrijver, H., and Röckmann, T.: Dynamic Processes Governing Lower-Tropospheric HDO/H₂O Ratios as Observed from Space and Ground, *Science*, 325, 1374–1377, doi:10.1126/science.1173791, 2009.
- 35

- Gimeno García, S., Schreier, F., Lichtenberg, G., and Slijkhuis, S.: Near infrared nadir retrieval of vertical column densities: methodology and application to SCIAMACHY, *Atmospheric Measurement Techniques*, 4, 2633–2657, doi:10.5194/amt-4-2633-2011, 2011.
- Gloude-mans, A. M. S., Schrijver, H., Kleipool, Q., van den Broek, M. M. P., Straume, A. G., Lichtenberg, G., van Hees, R. M., Aben, I., and Meirink, J. F.: The impact of SCIAMACHY near-infrared instrument calibration on CH₄ and CO total columns, *Atmos. Chem. Phys.*, 5, 2369–2383, doi:10.5194/acp-5-2369-2005, 2005.
- Gloude-mans, A. M. S., Krol, M. C., Meirink, J. F., de Laat, A. T. J., van der Werf, G. R., Schrijver, H., van den Broek, M. M. P., and Aben, I.: Evidence for long-range transport of carbon monoxide in the Southern Hemisphere from SCIAMACHY observations, *Geophys. Res. Lett.*, 33, L16 807, doi:10.1029/2006GL026804, 2006.
- Gloude-mans, A. M. S., Schrijver, H., Hasekamp, O. P., and Aben, I.: Error analysis for CO and CH₄ total column retrievals from SCIAMACHY 2.3 μm spectra, *Atmos. Chem. Phys.*, 8, 3999–4017, doi:10.5194/acp-8-3999-2008, 2008.
- Gloude-mans, A. M. S., de Laat, A. T. J., Schrijver, H., Aben, I., Meirink, J. F., and van der Werf, G. R.: SCIAMACHY CO over land and oceans: 2003–2007 interannual variability, *Atmos. Chem. Phys.*, 9, 3799–3813, doi:10.5194/acp-9-3799-2009, 2009.
- Griffith, D. W. T., Velazco, V. A., Deutscher, N., Murphy, C., Jones, N., Wilson, S., Macatangay, R., Kettlewell, G., Buchholz, R. R., and Riggenbach, M.: TCCON data from Wollongong, Australia, Release GGG2014R0., TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A., doi:http://dx.doi.org/10.14291/tcon.ggg2014.wollongong01.R0/1149291, 2014.
- Landgraf, J., aan de Brugh, J., Borsdorff, T., Houweling, S., and O., H.: Algorithm Theoretical Baseline Document for Sentinel-5 Precursor: Carbon Monoxide Total Column Retrieval, Atbd, SRON, Sorbonnelaan 2, 3584 CA Utrecht, The Netherlands, 2016a.
- Landgraf, J., aan de Brugh, J., Scheepmaker, R., Borsdorff, T., Hu, H., Houweling, S., Butz, A., Aben, I., and Hasekamp, O.: Carbon monoxide total column retrievals from TROPOMI shortwave infrared measurements, *Atmos. Meas. Tech.*, 2016, 1–39, doi:10.5194/amt-2016-114, 2016b.
- Nédélec, P., Cammas, J.-P., Thouret, V., Athier, G., Cousin, J.-M., Legrand, C., Abonnel, C., Lecoœur, F., Cayez, G., and Marizy, C.: An improved infrared carbon monoxide analyser for routine measurements aboard commercial Airbus aircraft: technical validation and first scientific results of the MOZAIC III programme, *Atmos. Chem. Phys.*, 3, 1551–1564, doi:10.5194/acp-3-1551-2003, 2003.
- Rodgers, C. D.: *Inverse Methods for Atmospheric Sounding : Theory and Practice*, vol. 2 of *Series on Atmospheric, Oceanic and Planetary Physics*, World Scientific, Singapore., River Edge, N.J., réimpression : 2004, 2008, 2000.
- Scheepmaker, R. A., Frankenberg, C., Deutscher, N. M., Schneider, M., Barthlott, S., Blumenstock, T., Garcia, O. E., Hase, F., Jones, N., Mahieu, E., Notholt, J., Velazco, V., Landgraf, J., and Aben, I.: Validation of SCIAMACHY HDO/H₂O measurements using the TCCON and NDACC-MUSICA networks, *Atmospheric Measurement Techniques*, 8, 1799–1818, doi:10.5194/amt-8-1799-2015, 2015.
- Schrijver, H., Gloude-mans, A. M. S., Frankenberg, C., and Aben, I.: Water vapour total columns from SCIAMACHY spectra in the 2.36 μm window, *Atmospheric Measurement Techniques*, 2, 561–571, doi:10.5194/amt-2-561-2009, 2009.
- Sherlock, V., Connor, B., Robinson, J., Shiona, H., Smale, D., and Pollard, D.: TCCON data from Lauder, New Zealand, 120HR, Release GGG2014R0., TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A., doi:http://dx.doi.org/10.14291/tcon.ggg2014.lauder01.R0/1149293, 2014.
- Veefkind, J., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H., de Haan, J., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P.: {TROPOMI} on the {ESA} Sentinel-5 Precursor: A {GMES} mission for global observations of the atmospheric composition for climate,

- air quality and ozone layer applications, *Remote Sensing of Environment*, 120, 70 – 83, doi:<http://dx.doi.org/10.1016/j.rse.2011.09.027>, the Sentinel Missions - New Opportunities for Science, 2012.
- Vidot, J., Landgraf, J., Hasekamp, O., Butz, A., Galli, A., Tol, P., and Aben, I.: Carbon monoxide from shortwave infrared reflectance measurements: A new retrieval approach for clear sky and partially cloudy atmospheres, *Remote Sens. Environ.*, 120, 255 – 266, the Sentinel Missions - New Opportunities for Science, 2012.
- 5 Wang, P., Stammes, P., van der A, R., Pinardi, G., and van Roozendael, M.: FRESCO+: an improved O₂ A-band cloud retrieval algorithm for tropospheric trace gas retrievals, *Atmospheric Chemistry and Physics*, 8, 6565–6576, doi:10.5194/acp-8-6565-2008, 2008.
- Wassmann, A., Borsdorff, T., van de Brugh, J. M. J., Hasekamp, O. P., Aben, I., and Landgraf, J.: The direct fitting approach for total ozone column retrievals: a sensitivity study on GOME-2/MetOp-A measurements, *Atmospheric Measurement Techniques*, 8, 4429–4451, doi:10.5194/amt-8-4429-2015, 2015.
- 10 Williams, J. E., van Velthoven, P. F. J., and Brenninkmeijer, C. A. M.: Quantifying the uncertainty in simulating global tropospheric composition due to the variability in global emission estimates of Biogenic Volatile Organic Compounds, *Atmos. Chem. Phys.*, 13, 2857–2891, doi:10.5194/acp-13-2857-2013, 2013.
- Williams, J. E., Le Bras, G., Kukui, A., Ziereis, H., and Brenninkmeijer, C. A. M.: The impact of the chemical production of methyl nitrate from the NO + CH₃O₂ reaction on the global distributions of alkyl nitrates, nitrogen oxides and tropospheric ozone: a global modelling study, *Atmos. Chem. Phys.*, 14, 2363–2382, doi:10.5194/acp-14-2363-2014, 2014.
- 15 Wunch, D., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Stephens, B. B., Fischer, M. L., Uchino, O., Abshire, J. B., Bernath, P., Biraud, S. C., Blavier, J.-F. L., Boone, C., Bowman, K. P., Browell, E. V., Campos, T., Connor, B. J., Daube, B. C., Deutscher, N. M., Diao, M., Elkins, J. W., Gerbig, C., Gottlieb, E., Griffith, D. W. T., Hurst, D. F., Jiménez, R., Keppel-Aleks, G., Kort, E. A., Macatangay, R., Machida, T., Matsueda, H., Moore, F., Morino, I., Park, S., Robinson, J., Roehl, C. M., Sawa, Y., Sherlock, V., Sweeney, C., Tanaka, T., and Zondlo, M. A.: Calibration of the Total Carbon Column Observing Network using aircraft profile data, *Atmos. Meas. Tech.*, 3, 1351–1362, doi:10.5194/amt-3-1351-2010, 2010.
- 20 Wunch, D., Wennberg, P. O., Toon, G. C., Connor, B. J., Fisher, B., Osterman, G. B., Frankenberg, C., Mandrake, L., O'Dell, C., Ahonen, P., Biraud, S. C., Castano, R., Cressie, N., Crisp, D., Deutscher, N. M., Eldering, A., Fisher, M. L., Griffith, D. W. T., Gunson, M., Heikkinen, P., Keppel-Aleks, G., Kyrö, E., Lindenmaier, R., Macatangay, R., Mendonca, J., Messerschmidt, J., Miller, C. E., Morino, I., Notholt, J., Oyafuso, F. A., Rettinger, M., Robinson, J., Roehl, C. M., Salawitch, R. J., Sherlock, V., Strong, K., Sussmann, R., Tanaka, T., Thompson, D. R., Uchino, O., Warneke, T., and Wofsy, S. C.: A method for evaluating bias in global measurements of CO₂ total columns from space, *Atmos. Chem. Phys.*, 11, 12317–12337, doi:10.5194/acp-11-12317-2011, 2011.
- 25

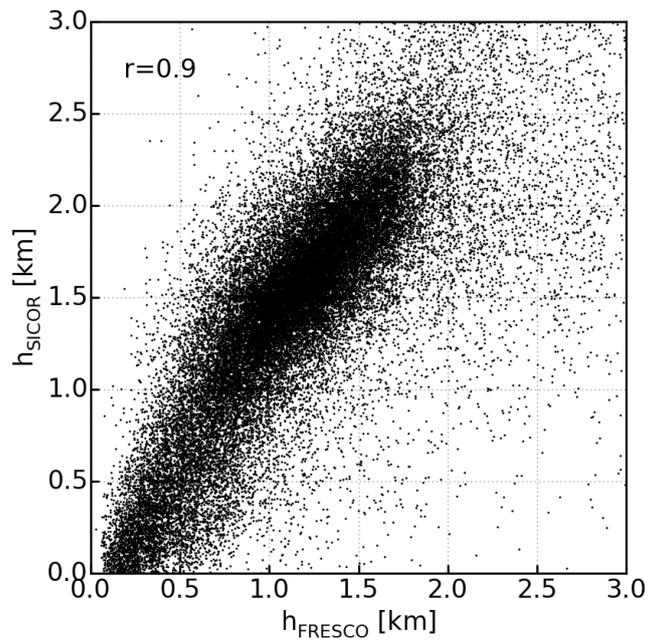


Figure 1. Scatter plot of the retrieved cloud height h_{SICOR} in km with the one provided by Fresco h_{FRESKO} . For the year 2003 over the ocean in the latitude/longitude box $[(30^\circ\text{S}, 30^\circ\text{W}), (0^\circ, 0^\circ)]$ with cloud optical thickness > 1 are shown.

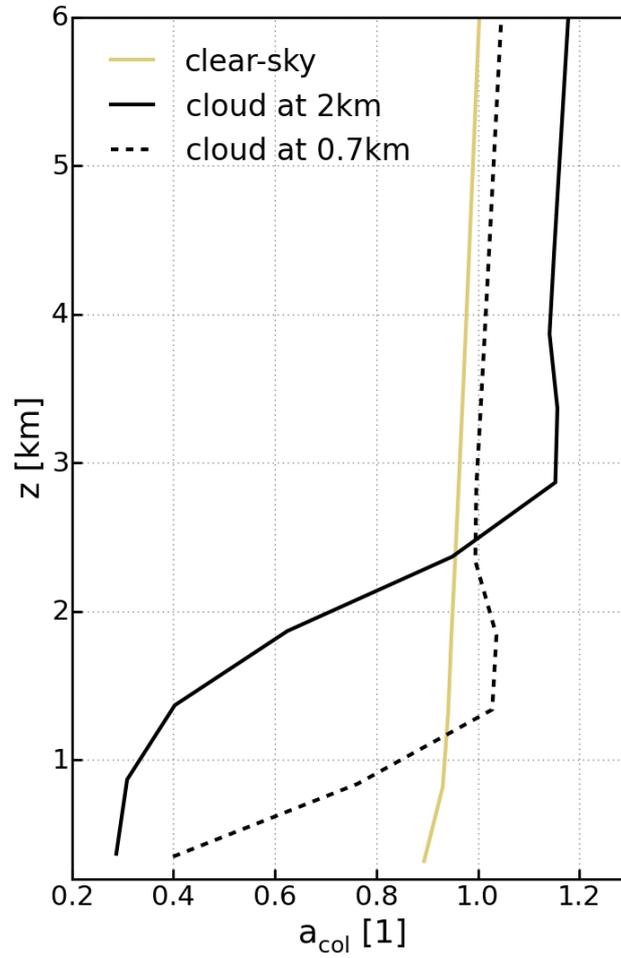


Figure 2. Example SCIAMACHY CO column averaging kernels over Australia in 2003 as a function of altitude for the following cases: clear-sky (yellow line), a cloud at-with 0.7 km center height (dashed line), and a cloud at-with 2 km center height (solid line). Cloud optical thickness of 2.5 for the relevant cases.

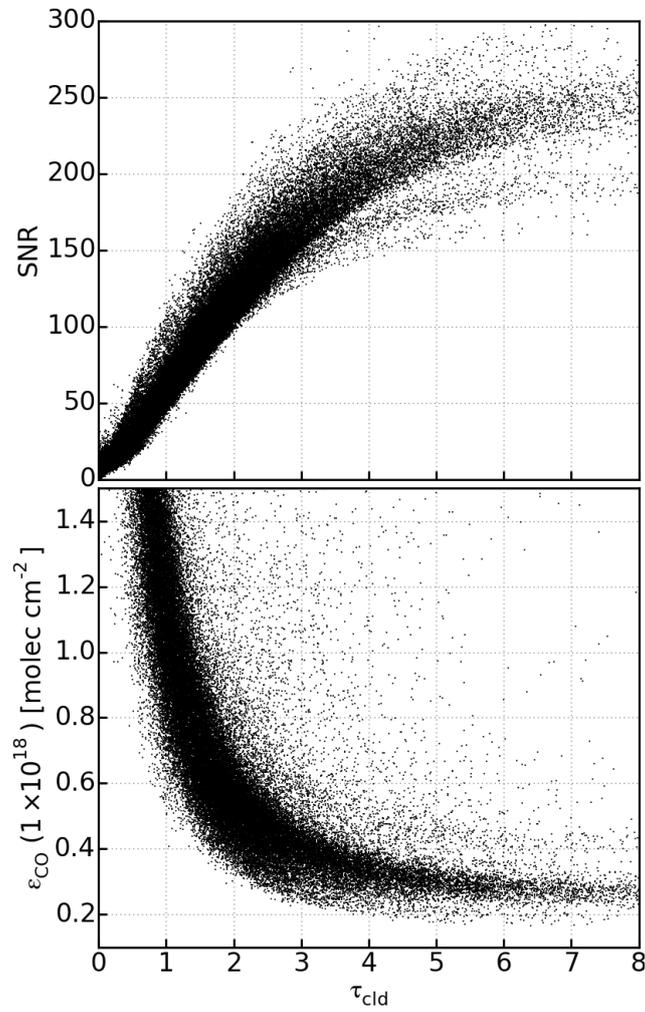


Figure 3. Top panel: SNR in the SCIAMACHY CO retrieval window. Lower panel: CO retrieval error ϵ_{CO} (as a function of the retrieved cloud optical thickness τ_{cld}). Data is shown for the year 2003 for a latitude/longitude box $[(30^{\circ}\text{S},30^{\circ}\text{W}), (0^{\circ},0^{\circ})]$ over the ocean

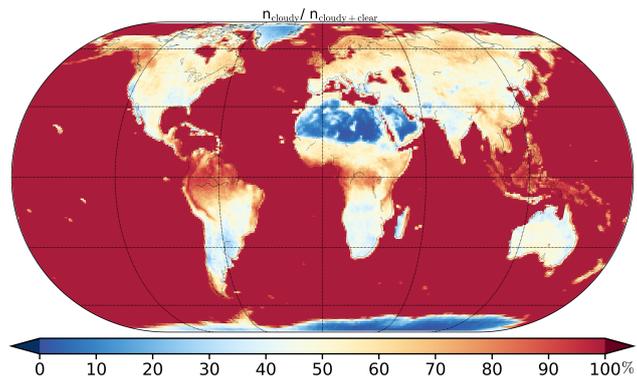


Figure 4. Percentage of cloudy-sky retrievals of all retrievals from January 2003 to the end of the SCIAMACHY mission in April 2012. Here, cloudy-sky retrievals comprise all data products of category 2 and 3, defined in Sec. 2.3.

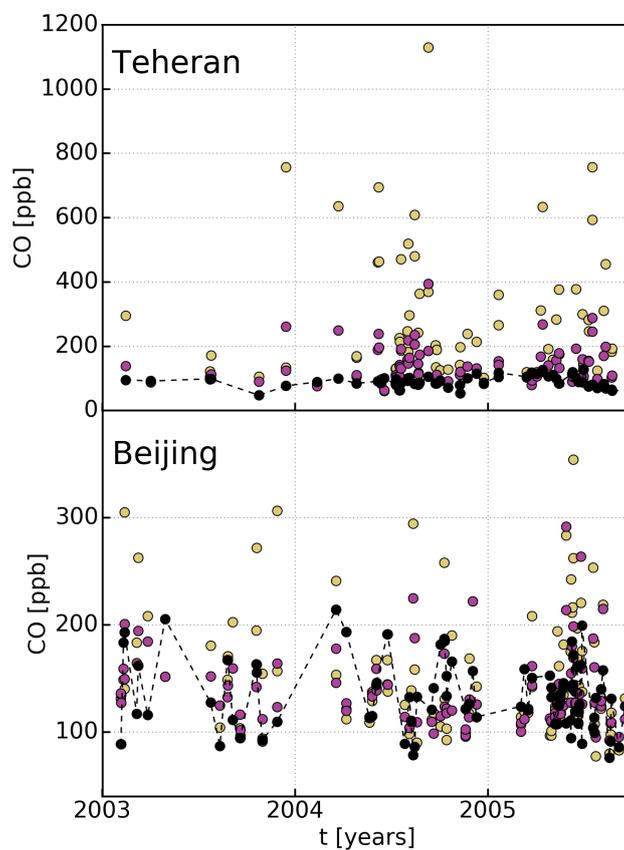


Figure 5. CO column mixing ratios of Teheran and Beijing measured by MOZAIC/IAGOS (yellow), MOZAIC/IAGOS with SCIAMACHY column averaging kernels applied (pink), and SCIAMACHY retrievals for optically thick high cloud conditions (black).

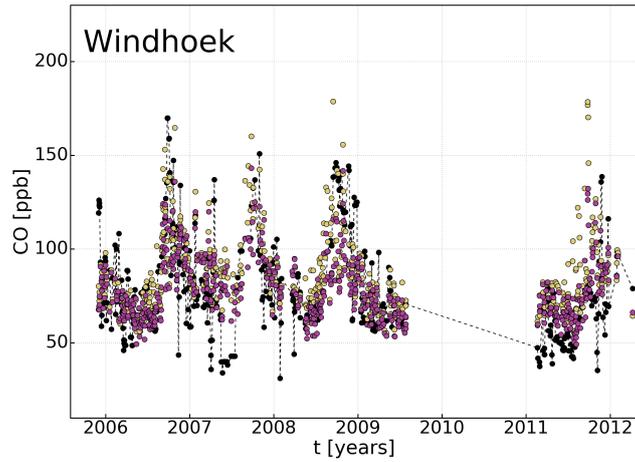


Figure 6. As Fig. 5, but for Windhoek.

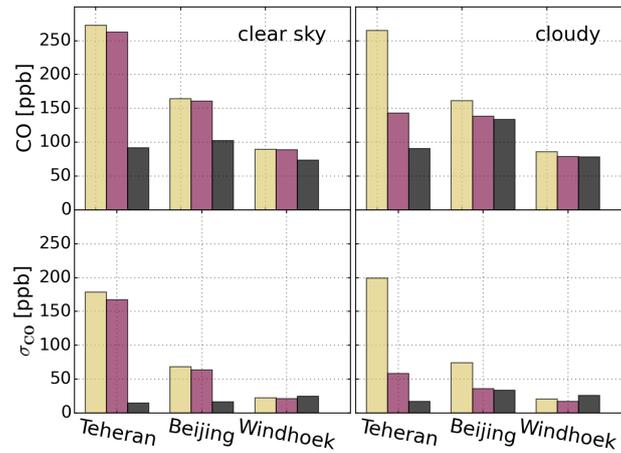


Figure 7. Mean CO column mixing ratio (top panel) and standard deviation (lower panel) of the time series shown in Fig. 5 and Fig. 6, SCIAMACHY (black), MOZAIC/IAGOS (yellow) and MOZAIC/IAGOS with SCIAMACHY column averaging kernels applied (pink). Left panel: clear sky SCIAMACHY retrievals and right panel: SCIAMACHY retrieval for optically thick high cloud conditions.

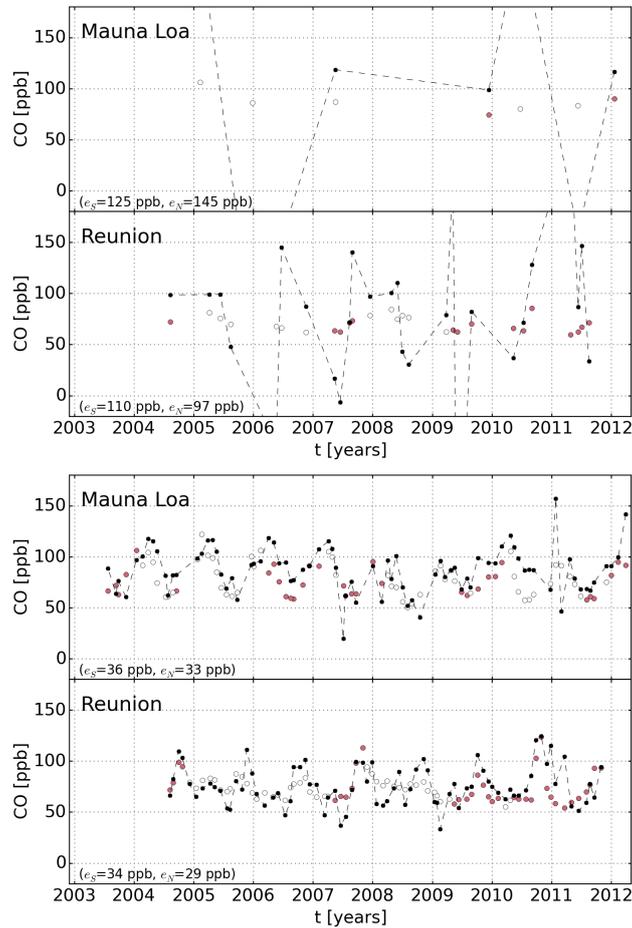


Figure 8. 30-day median of CO column averaged mixing ratios measured by SCIAMACHY (black) and at two NDACC-IRWG stations (pink). Upper two panels: SCIAMACHY clear-sky retrievals, lower two panels: SCIAMACHY cloudy-sky retrievals with optical thick low clouds. Open circles denote interpolated values for periods where no NDACC-IRWG measurements are available. The scatter ϵ_S and the mean retrieval noise ϵ_N for individual retrievals in the monthly bins are shown.

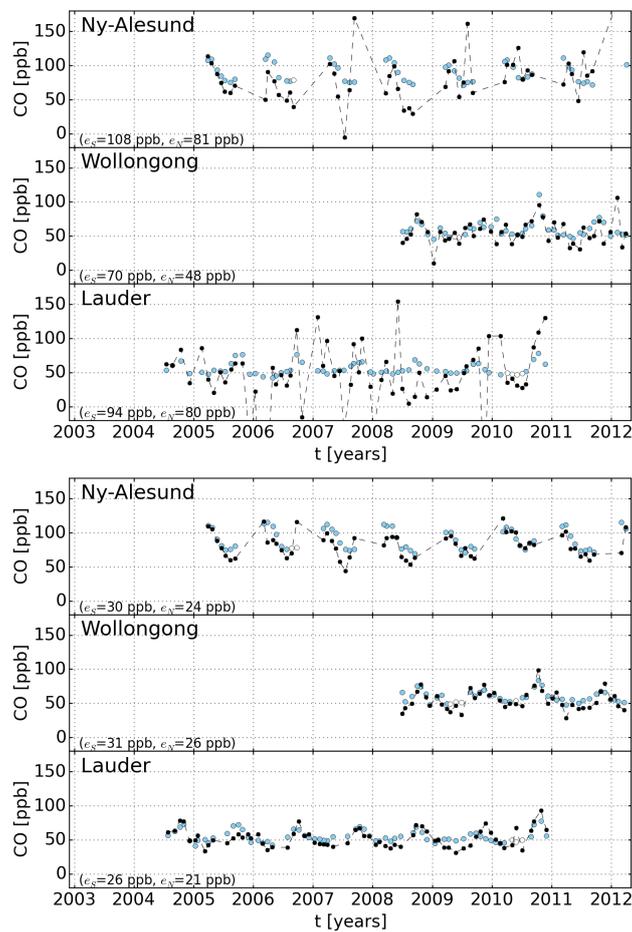


Figure 9. Same as Fig. 8, for TCCON measurements indicated in blue. For Wollongong and Lauder the GGG2014 release of TCCON was used, except for Ny-Ålesund data release GGG2012 is used [because the newer release GGG2014 is not yet available for this site.](#)

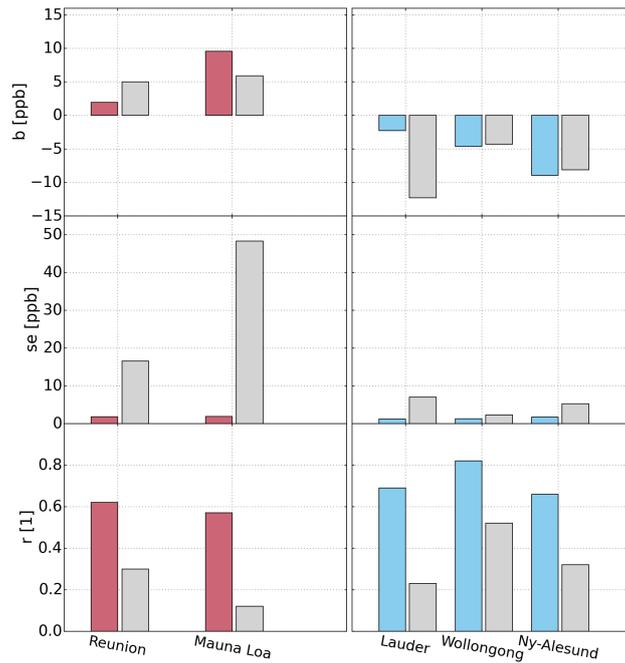


Figure 10. Mean bias SCIAMACHY - FTS (upper panel), standard error of the mean bias (middle panel), and Pearson correlation coefficient of SCIAMACHY with FTS (lower panel) derived from monthly medians as shown in Figs. 8 and 9. Measurements of NDACC (left) and TCCON stations (right) are compared with clear-sky (grey) and cloudy (coloured) SCIAMACHY retrievals.

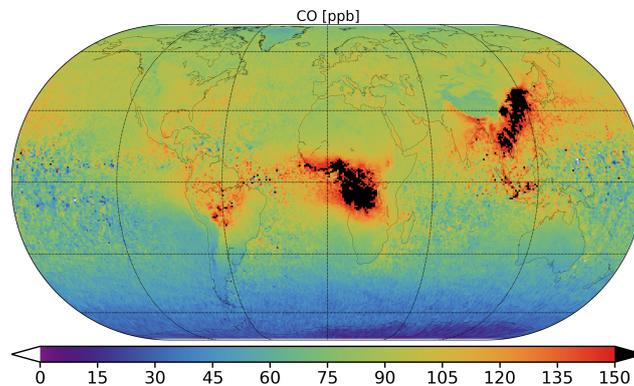


Figure 11. CO column averaged mixing ratios in ppb over land and ocean from clear-sky and cloudy-sky measurements for optically thick low cloud conditions. The values are averaged from January 2003 to the end of the SCIAMACHY mission in April 2012.

Table 1. MOZAIC/IAGOS airports used for validation. The temporal coverage with the SCIAMACHY mission is given in years.

number	name	latitude	longitude	alt	MOZAIC/IAGOS
1	Beijing	40.38°	115.22°	0.04 km	2003–2005
2	Tehran	35.98°	50.24°	1.01 km	2003–2005
3	Windhoek	−21.43°	17.34°	1.72 km	2005–2012

Table 2. Same as Table 1, but for NDACC and TCCON sites.

number	name	latitude	longitude	alt	NDACC	TCCON
1	Ny-Ålesund	78.92°	11.92°	0.02 km	–	2005–2011
2	Mauna Loa	19.54°	−155.57°	3.40 km	2003–2012	–
3	Reunion	−20.90°	55.49°	0.09 km	2004–2011	–
4	Wollongong	−34.41°	150.88°	0.03 km	–	2008–2012
5	Lauder	−45.05°	169.68°	0.37 km	–	2004–2010