Author's Reply to Review 1:

We would like to thank reviewer #1 for his/her 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 reviewers are numbered and typeset in italic font. Line, page and figure numbers in the reviewer's comments refer to the original manuscript. Following every comment we give our reply. We provided a revised version of the manuscript in which all changes are indicated: Newly added sections are typeset in blue, sections that have been removed are struck through. In our reply we give page and line numbers that refer to the revised manuscript.

Major Items

1.) While the manuscript explains the essential details of the retrieval algorithm, I found it difficult to follow the data processing flow from its beginning to end. The algorithm involves several steps of filtering and a variety of external data sources. I believe readers would benefit tremendously by the inclusion of a flowchart showing the different steps of retrieval processing and the various models and data sources that are exploited in operational data processing. This flowchart should appear in Section 2, before the sensitivity studies.

Changed: We agree. A flowchart is added in Figure 2 and a short explanation is given line 98-104 at the beginning of section 2:

"Figure 2 summarizes the SICOR algorithm. The dynamic input includes S5P level 1b data, which comprises solar irradiance and Earth radiance spectra in the spectral range 2315-2338 nm, forecast (FCST) data on atmospheric pressure **p**, temperature **T** and water vapor profiles, the terrain's elevation from a digital elevation model (DEM), and a priori information on the CO and CH_4 vertical distribution of the observed atmosphere coming from a chemical transport model (CTM). These input data and their required accuracy will be discussed in more detail in Section 3. The first processing step screens the data to filter out observations with high and optically thick clouds. Subsequently, we utilize a physics-based retrieval approach to infer CO columns from SWIR measurements together with the atmospheric H_2O abundances, surface albedo and a spectral calibration of the measurement spectrum. The spectral absorption by methane is used to infer information on atmospheric scattering by clouds and aerosols. Finally, the algorithm has as output the retrieved CO total column, the corresponding column averaging kernel and an estimate of the random error component. The theoretical baseline of our algorithm is described in detail in the following."

2.) The presented theory of the retrieval algorithm is also somewhat confusing. Three different equations (Eqs. 6, 8, and 17) all reportedly describe the cost function of the TROPOMI CO retrieval algorithm. Even after reading the text a second time, it is unclear which of these cost functions is most relevant. It appears that Eq. 6 simplifies to become Eq. 8 as the result of forcing the DFS value to 1. If that is true, I would suggest deleting Eq. 6 entirely since it really is not relevant to the operational retrieval algorithm. Eq. 17 seems to only be applied when it is necessary to suppress singularities caused when the measurement is insensitive to specific elements of the state vector. Is that correct, or is Eq. 17 actually the "true" cost function for all retrievals? If Eq. 8 and Eq. 17 are used for different situations, then the authors should describe which cost function applies to which scenarios and how exactly that decision is made. If Eq. 17 is applied for all retrievals, then there is no obvious need to present Eq. 8.

Text added: We disagree with the reviewer's suggestion to omit equation 6 and 8 in Section 2.3. Our intension is not only to present the equation, which is finally implemented in the algorithm, but also to summarize the motivation for the final minimization problem in Eq. 17, based on recent research. For this purpose, we find it essential to discuss also Eq. (6) and (8). Also reviewer #2 acknowledges this approach. However, we realize from the reviewer comment that the text not clearly states, which equation is finally applied in the software. Therefore we have added the text at line 287-292:

"In a nutshell, the implementation of the SICOR inversion algorithm is based on the minimization problem (18). It comprises the CO profile scaling approach as a particular regularization of the CO profile retrieval in Eq. (6). For the software implementation, we make use of the fact that its solution is identical to the least squares problem (9), where the atmospheric abundance of CO abundance is adjusted through scaling of a reference profile. Finally to prevent numerical instabilities in the retrieval, we introduce a second regularization in Eq. 18, which mainly affects the inversion of cloud and surface parameters and its effect on the retrieved CO column can be neglected."

3.) The manuscript reports simulations of the effects of water and ice clouds on retrieval bias, but there seems to be no analysis of the effects of aerosols. Is this because aerosols have been shown to have negligible effect on CO retrievals, or is it simply because this effect has not yet been investigated?

Text added: We agree that the manuscript is not clear on the algorithm performance for aerosol-loaded atmosphere. However to avoid duplication of previous results by Vidot et al., 2012, we added the following text at line 343-347:

"This approach appeared to be appropriate to describe the effect of optically thick clouds and boundary layer aerosols in the retrieval and similar small retrieval biases are achieved with the latest version of SICOR described here. However, in case of an optically thin scattering layer due to an elevated dust layer, optically thin clouds and cirrus above a bright surface, the previous version of SICOR (Vidot et al., 2012) could not account for any path enhancement of the observed light due to light trapping between the scattering layer and the surface."

and at line 370-371:

"Similar small biases are found for an elevated dust layer and optically thin clouds (not shown)."

Finally, aerosol information is considered in the orbit simulation together with cloud and cirrus properties as mentioned at line 540-541.

Minor Revisions and Technical Corrections

1.) p. 2, l. 38 - The MOPITT instrument observes CO using both thermal-infrared and SWIR bands.

Changed: see line 40-41

2.) p. 4, l. 75 - Please include reference and spell out the acronym 'SICOR'.

Changed: Acronym is spelled out at line 76-77, corresponding reference Vidot et al. 2012 was already provided in the original manuscript.

3.) p. 7, l. 75 - For readers who are unfamiliar with the general topic of this paper, it would be worthwhile to explain qualitatively how scattering can lead to either pathlength shortening or path-length enhancement.

Changed: We modified the text at line 117-120 and added an appropriate reference.

"The difference ΔCH_4 between the retrieved CH_4 column and a priori methane information coming from a chemical transport model indicates light path modification, either shortening or enhancement in comparison with the direct lightpath from the sun to the spectrometer via reflection at the Earth surface, due to atmospheric scattering by clouds and aerosols. Here the net effect depends on the scattering properties such as scattering height and optical depth, surface reflection and solar and observation geometry (e.g. Aben et al., 2007)."

4.) p. 7, l. 80 - Suggest deleting 'thoroughly'.
Deleted

5.) p. 7, l. 136 - Suggest rewriting sentence 'The fit window ... ' to emphasize the factors considered in the process of selecting the spectral window for CO retrievals.

Not changed: We do not understand the reviewer's suggestion and think that the sentence at line 148-150

"The fit window compromises about optimal CO sensitivity, little interference with water vapor and methane absorptions, and small forward model errors due to the assumed cloud model."

already emphasize the criteria for the selection of the spectral fit window.

5.) p. *7, l. 161 - Please provide more detail with respect to the assumed cloud geometry; for example, what is meant by a 'triangular height profile'? Are single-scattering albedo and asymmetry parameter fixed or are they assumed to vary with height?*

Changed: text adapted at line 174-176:

"In the forward model, clouds and aerosols are represented by a scattering layer with a triangular height profile in optical depth with a center height z_{scat} and a fixed full width at half maximum of 2.5 km. Within the scattering layer we assume a constant single scattering albedo and scattering phase function."

6.) p. 7, l. 169 - Instead of 'microphysical properties', it appears that the authors mean 'optical properties'.

Changed: see line 182

7.) p. 7, l. 170 - References should be provided indicating that these are reasonable values for single-scattering albedo and asymmetry parameter for the CO spectral band; are these values appropriate for both clouds and aerosols?

Text added:

The scattering parameters assumed in the retrieval are not meant to represent true scattering parameters but to describe clouds, cirrus and aerosols effectively, in the sense that the simulated lightpath is of sufficient accuracy to retrieve CO with in the required uncertainty. Here the retrieved scattering parameters may compensate for the ad hoc assumption on the single-scattering albedo and asymmetry parameter but also for other forward model errors. So, the question if '*these values are appropriate for both clouds and aerosols'* can only be answered by analyzing the introduced bias on the retrieved CO column. This is the subject of Sec. 3 and in particular of Sec. 4. To clarify this, we added at line 189-191 the text

These a priori assumptions on the optical parameters of clouds and aerosols will be justified by the performance analysis in Sec. 4, which is based on measurement simulations of one TROPOMI orbit for a realistic variation of cloud and aerosol properties.

8.) p. 8, l. 207 - What is the 'vertical derivative operator' and what is its significance? Does this constraint essentially keep the shape of the profile fixed as the total column amount varies?

Changed: At line 222, we added the text:

"γ is the regularization parameter and

$$L1 = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix}$$

is the first-difference operator and so the regularization favors constant solutions and penalizes the "roughness" off x."

9.) p. 8., l. 212 - What are the consequences of assuming a linear dependence of surface albedo on wavelength? Is this a reasonable approximation?

Not changed: To our knowledge no information is available on the spectral variation of the 2.3 µm surface albedo between 2324 and 2338 nm averaged over a spatial domain of several square kilometers. We do have experiences with CO_2 and CH_4 retrieval from GOSAT measurements in the 1.6 µm spectral range, where spectral fit residuals show that the assumption of the linear dependence of surface albedo on wavelength is appropriate. The validity of this assumption for TROPOMI measurements of CO can only be demonstrated with real measurements.

10.) p. 9, l. 215 - The meaning of the sentence 'To account for the CO sensitivity ...' is unclear.

Changed: At line 234-238 we changed the text to:

~~To account for the CO sensitivity of the SWIR measurements in the inversion,~~ We regularize the solution in Eq. (6) such that one degree of freedom for signal (DFS) of the retrieved CO profile is inferred from the measurement, bearing in mind that the TROPOMI SWIR measurements are only sensitive to the total amount of CO along the path of the observed light through the atmosphere.

11.) p. 10, l. 263 - In Eq. 17, it appears that the product Wx should be dimensionless. However, some of the elements of the state vector x (such as z_cld) are not dimensionless. Therefore, it appears that at least some of the diagonal elements of W should not be dimensionless either. So, shouldn't the diagonal elements of W depend on the units of the corresponding elements in x? The text states that the diagonal values of W are all either 1 or 0, and no units are given.

Changed: The reviewer is correct and we added the following remark at line 233-234:

"To keep all elements of the state vector dimensionless, also these entries of the state vector are normalized to a reference value."

So **x** is now dimensionless and hence also the weighting matrix **W**.

12.) p. 12, l. 298 - The listed telescope aperture is equivalent to 6 mm^2, which is surprisingly small. Is there a typo?

Changed: The values provided in the original manuscript were not the most recent design values. We updated the text accordingly at line 324-326:

The radiance spectra are perturbed by measurement noise from the TROPOMI noise model by Tol et al. (2011) for an instantaneous view with a footprint of $3.5 \times 7\,\mathrm{km}^2$ and a telescope aperture of 12 mm. (The etendue of the SWIR channel is 4.3×10^{-10} m^2 sr.)

13.) p. 14, l. 341 - For consistency with the order of panels in Fig. 5, consider discussing the results of the cirrus experiment after the middle-cloud experiment.

Changed: See text at line 367-371:

"To investigate the effect of photon trapping between clouds and the surface, Fig. 6 depicts the CO bias for a cloud between 4 and 5 km altitude with a small optical depth $\tau_{\mathrm{scat}} = 2$ as a function of surface albedo and cloud coverage. Here, the CO bias reaches 1.5 % with increasing cloud coverage. For a cirrus layer between 9 and 10 km of varying optical depth as function of the surface albedo, the light trapping effect at high surface albedo results in a CO biases $b_{\mathrm{CO}} \leq 0.5\,\%$."

14.) p. 14, l. 342 (and Fig. 5) - In both the left and right panels in Fig. 5, why are there no bias results reported in the lower right corners of the figures (i.e., the grey-colored area)? Are these cases filtered out somehow? Please explain this in the text and in the figure caption.

Changed: Figure caption adjusted:

"**Fig. 6.** Left panel: Retrieval bias in case of a cloud atmosphere. The CO bias is shown as a function of surface albedo A_s and cloud fraction f for a cloud between 4 and 5 km altitude with optical depth τ_{scat} = 2 and a VZA of 0 degree. Right panel: CO retrieval bias for measurements in presence of optically thin cirrus, which overcasts the entire scene, as a function of surface albedo and cirrus optical depth that defined at 2300 nm. The grey area indicates measurement simulations, which were rejected by the cloud filter."

15.) p. 14, l. 343 (and entire paragraph following this line) - While the rest of the manuscript was generally well written, I was unable to understand the material presented in this paragraph. I would suggest rewriting it. First, clearly describe the source of the retrieval error. Which implicit assumptions or approximations in the retrieval algorithm produce this error?

Changed, paragraph at line 373-391 rewritten:

"Moreover, we investigated the implications of the retrieved cloud parameters being effective cloud parameters. These parameters differ from the truth because of the limited information available from the satellite measurements. Here, the retrieval forward model has to describe clouds in a simplified manner with a few free parameters and all remaining cloud properties have to be fixed a priori (see e.g. Koelemeijer et al., 2002; van Diedenhoven et al., 2007). In our case, the cloud model includes several simplifications, e.g. a horizontally homogenous cloud with the triangular height distribution in optical depth and a two stream radiative transfer model to describe the cloud radiative properties. Considering the measured radiometric signal as a mean of a photon ensemble with different light paths through the atmosphere, the retrieval adjusts the cloud parameters and the simulated light paths such that the methane absorption features can be fitted by the forward model. This may include erroneous light paths, which effects average out in the simulated measurement for the particular height distribution of methane. However for another trace gas with a different vertical profile, such as CO, the relevance of the individual photons for the observed signal may differ and so the simulated light paths introduce spectral errors in the simulated CO absorption features. Subsequently adjusting the trace gas concentrations in the retrieval, CO biases are introduced for cloudy atmospheres.

Obviously, this retrieval error depends on the particular CO profile and the altitude at which the simulated light path deviates from its truth. So to characterize this inherent bias of our retrieval approach, we simulate SWIR measurements for a cloudy atmosphere adding CO abundance in a 1 km thick, vertically homogenous layer with varying layer top height z_{per}."

16.) p. 19, l. 432. For the TROPOMI instrument, what are the design goals for radiometric offset error and multiplicative radiometric error?

Changed: Text added at line 473-475:

"The corresponding TROPOMI observation requirement for a radiometric offset is < 0.1 % and for a multiplicative radiometric error on the Earth radiance measurement it is < 2 % (Buscaglione, 2011)."

17.) p. 27. Conclusion. Please address the impact of aerosols. Are they unimportant, or has their effect on retrieval bias not yet been studied?

Changed: text added at line 617-618,

"Similar good accuracy could be achieved for boundary layer aerosols and elevated dust layers."