AMT-2017-136

Principal criteria	Excellent (1)	Good (2)	Fair (3)	Poor (4)
Scientific significance: Does the manuscript represent a substantial contribution to scientific progress within the scope of Atmospheric Measurement Techniques (substantial new concepts, ideas, methods, or data)?		x		
Scientific quality: Are the scientific approach and applied methods valid? Are the results discussed in an appropriate and balanced way (consideration of related work, including appropriate references)?		x		
Presentation quality: Are the scientific results and conclusions presented in a clear, concise, and well-structured way (number and quality of figures/tables, appropriate use of English language)?		x		

I. General impression

In comparison with the first submitted version of this manuscript, this second version has much improved in many ways. The title, abstract, research intentions and presentation are clear and the work is well motivated, the scientific language and English formulations have strongly improved, and the unclarities in the original version have mostly been addressed and removed or elucidated. Yet two aspects of the paper might still require minor revision (see specific comments below).

II. Specific comments:

1. This work addresses many aspects of SCIAMACHY L1 data and their evaluation in several ways. This certainly has scientific value, but too much information reduces the overall clarity. I would therefore suggest to leave out section 3 and the right column of Figure 3, and integrate the information on the left column of Figure 3 into section 4.1 (and combine Tables 3 and 4 correspondingly). A separate focus on the 10°N to 10°S L1 version effects on L2 seems somewhat redundant when all versions are compared with ozonesonde measurements on a global scale in a later section.
2. It is unexplained and very confusing that the prior comparison results in terms of median bias are typically much better in the UTLS and below for the prior profiles than for the retrieved profiles! This might be due to the fact that the McPeters-Labow climatology is generated from ozonesonde (and MLS) profiles. In terms of random uncertainty however (i.e. comparison spread) the satellite retrievals might do better than the prior profiles due to their better caption of specific ozone dynamics. It is therefore suggested to either leave out the prior comparison (easy option) or add the comparison spread to the plots and explain the differences between satellite retrieval and prior comparisons bias and spread (scientifically preferred option).

III. Technical corrections:

In the abstract it is not fully clear whether the improvement suggestions made in the last paragraph are there for information only, or have really been tested. Please specify that the latter is true.

Page 1 line 61: When referring to Staehelin, please mention 70 “operational” stations worldwide. There are more stations, but not all provide data all the time.

Page 2 line 100: Remove double brackets around reference.

Section 3 and Figure 4: DFS profiles have little quantitative information if the corresponding layer thicknesses are unknown. Please explicitly refer to level-DFS or layer-DFS in the text and use markers instead of lines in the plot, or convert to DFS/km values for clarity.

Page 12 last paragraph of section 4.2 and Figure 8: “The validation results are clearly less noisy and smoother for the case where the AK was applied to the ozone sondes.” is still rather vague. Please mention explicitly that AK-smoothing of the reference data is performed with the intention to remove the vertical smoothing difference error from the comparison error budget between satellite and ground-based observations. With this explicit formulation, Figure 8 even can be considered superfluous and can be left out of the manuscript, as it contains little information that contributes to the remainder of the discussion.