Interactive comment on “Early results and Validation of SAGE III-ISS Ozone Profile Measurements from Onboard the International Space Station” by Michael P. McCormick et al.

Anonymous Referee #1

Received and published: 7 November 2019

General comments:

This work provides the first validation results of the SAGE III-ISS ozone profile observations with respect to ozonesonde, lidar, and ACE-FTS measurements. Although results are usually presented in a clear way, the validation methodology and its description (mainly Section 2.4, see below) should be improved before publication in AMT.

Specific comments:

- Abstract, line 20 (and throughout the text): The meaning of the term ‘accuracy’ is unclear as it is sometimes differently used in different contexts. In fact, ‘accuracy’ should not be used as a synonym of total uncertainty; see VIM, GUM, or Loew et al. (2017) for their application to satellite validation practices. It is recommended to just use ‘total uncertainty’ instead.

- Lines 66-70: The ‘CompositeO3’ product is not explained in the last paragraph of Section 2.1.

- Section 2.2: Information on the ozonesonde launch frequency and lidar observation frequency is missing, while this is of importance in considering the reference data coverage (temporal representativeness). Typical uncertainties of these measurements, with references to the appropriate literature, should be mentioned as well for proper interpretation of the comparison statistics later on.

- Section 2.4 is rather poorly written in terms of English language quality (e.g. lines 91-92), contains duplicate information, and the comparison statistics can be significantly improved. Regarding the latter, three important issues have been detected: 1. The coincidence criteria are not motivated, except for the mentioning of the need for larger values to compensate for the fewer comparisons in the southern hemisphere. The coincidence criteria should be motivated, ideally in terms of the (estimated) spatial extent of the measurements, including references. 2. Using a linear interpolation (line 105) for vertical sampling to a common grid is presently considered to leave important vertical profile information out of the resampling. Improved methods given in Calisesi et al. (2005) or Langerock et al. (2015) (or see Keppens et al. (2019) for an overview) could be implemented. 3. Calculating relative differences by using the averaged observation as a reference (lines 115-117) is not appropriate for satellite validation exercises (it is rather used for model comparisons as the average model is typically closer to the truth than each individual model): The reference averaging process corrupts the independency of the reference and smooths the effects that you want to detect (the unknown satellite errors), making their quantification flawed. If one wants to normalize the satellite deviation, the normalization should be done independently from the satellite data, i.e. using the reference data only.
Section 3: In order to properly address the satellite uncertainty budget, the spatiotemporal sampling difference errors and vertical smoothing difference errors that result from the inexact coincidences should be discussed as part of the uncertainties. E.g. one expects an effect of the larger coincidence criteria in the southern hemisphere.

Line 237: Please provide a reference for the statement that the tropical tropopause area “is most likely impacted by cirrus clouds”.

The conclusions should contain quantitative results (as the abstract).

Technical corrections:

Abstract, line 14: “the average difference of ozone concentration measured by SAGE III-ISS” is a confusing phrasing (as if average differences are measured). Please rephrase.

Line 43: Explain the acronym ‘CFC’.

Line 103: Provide links or references to the product user guides.

Line 190: Explain three-monthly acronyms.

Conclusions, line 258: “This paper represents an early effort to provide validation of ozone to the broad science-user community.” is a too broad statement. Which ozone, and measured by what?