Interactive comment on “Evaluation of balloon and satellite water vapour measurements in the Southern tropical UTLS during the HIBISCUS campaign” by N. Montoux et al.

Anonymous Referee #1

Received and published: 15 May 2007

General comments

As the authors point out in their introduction, “water vapor plays a key role in UT/LS climate and chemistry”. It flows that high accuracy water vapor measurements in the UT/LS are critically important. Nevertheless, as shown in the SPARC Assessment of Upper Tropospheric and Stratospheric Water Vapor [Kley et al., 2000], disagreements between various water vapor measurements, especially when water vapor is less than 5 ppmv, are often 30%. These uncertainties in stratospheric water vapor preclude the resolution of many issues surrounding the control of stratospheric water vapor and formation and evolution of cirrus clouds. [e.g. Peter et al., Science, 314, 1399-1402,
Accordingly, the evaluation instruments that measure water vapor in the UT/LS as part of an intercomparison campaign can provide important information. Nevertheless, for the evaluation to be useful, it must provide new and useful information. In this intercomparison, satellite and balloon-borne instruments are intercompared, and the authors provide a brief description of each instrument along with quoted precision and accuracy estimates. So what are the conclusions? Does the paper tell us anything new. The first paragraph of the summary and conclusion section is: “Altogether, individual and statistical comparisons of balloon and satellites water vapour measurements available during HIBISCUS provide a clear picture of the performances of each in the UTLS. The biases compared to AIRS used as reference and the variability of retrieved concentrations indicative of precision, are summarized in Table 7.” But what is that clear picture? Does the information as presented advance our understanding of the accuracy of water vapor measurements in the UT/LS as the data are currently presented? Does it provide any guidance as to which instruments provide the best water vapor data, or are performing within their stated accuracies? Does it even suggest that some instruments are clearly not performing within their stated accuracies? The answer to all these questions is no. Accordingly, this paper needs major revisions before it is published.

Specific issues that need to be addressed.

When clear differences exist between in situ and remote instruments, the potential explanation of air mass inhomogeneities near cloudy regions is offered as a potential explanation. “The two in-situ instruments, SDLA and SAW, flown on the same balloon agree each other, displaying water vapour mixing ratios 100-200% larger than that of HALOE and MIPAS, which could be explained by the difference in space and time between the measurements and by the presence of clouds as shown by the supersaturation up to the tropopause, hardly detectable from the orbit”. This statement, if true, would render the intercomparison with the in situ profiles virtually meaningless, which is unfortunate because the SDLA instrument has the potential for highly accurate
measurements. In fact, looking at Figure 2, the differences exhibited between SDLA and the other instruments extend to 20 km. With the tropopause at about 16 km (for the SF-2 flight anyway, as shown in Marecal et al., [2007]), it is unlikely that the entire profile above the tropopause is significantly moistened by convection. Accordingly, I would submit that the differences exhibited have a significant instrumental component. On the other hand, do the SLDA profiles suffer from balloon or gondola contamination around 19 km? Both the SF-2 and SF-4 profiles show anomalously high water in this region. The authors should address both of these issues as they are critical to questions of instrument accuracy.

There are a number of issues regarding the satellite instrument data.

For SAGE II, while the paper, “A revised water vapor product for the Stratospheric Aerosol and Gas Experiment (SAGE) II version 6.2 data set”, is referenced, below are three quotes from the reference:

“Herein, we describe the process by which we identified an apparent change in the spectral response in the water vapor channel and estimated the new channel spectral response.”

“The spectral response was corrected based on intercomparisons with HALOE”.

“It is important to note that the channel locations parameters were derived to match a mean comparison between SAGE II version 6.1 and the HALOE climatology at a single latitude, averaged over 4 years that has no altitude, seasonal, or temporal component.”

Accordingly, SAGE II and HALOE cannot be viewed as totally independent measurements. This should be stated clearly in the paper.

The HALOE data should be corrected according to HALOE principal investigator Ellis Remsberg [private communication] with increases ranging from a maximum of 1.2 near the tropopause to 1.05 at pressures below 56 mbar. This correction decreases the HALOE “low bias”.

S1789
The Hagan et al. AIRS validation paper covers data to altitudes approaching the tropical tropopause. I can find no validation of AIRS data above 100 hPa. The AIRS data used in the referenced Gettelman validation paper clearly show that for water vapor below 10 ppmv, AIRS is unreliable. In fact, the data below 10 ppmv in this paper are most probably weighted heavily by a priori values of 5 ppmv.

The reference for MIPAS validation does not include any reasonable validation data for UT/LS water vapor. The only intercomparison in that region is with radiosonde data. Above 25 km there is MIPAS intercomparison data with a ground-based microwave radiometer (MIAWARA) that shows good agreement. MIAWARA is validated by intercomparison with HALOE and NOAA frostpoint hygrometers. HALOE’s primary validation is with NOAA frostpoint hygrometers as well. So the validation of these satellite instruments are all tied to a major degree to frostpoint hygrometers. Is this the reason that, as stated in the summary, “HALOE, SAGE II and MIPAS are showing very consistent results”.

There is no reference for GOMOS validation. Should these data be included in the paper if they are not validated?

The zonal mean profile intercomparisons

These profiles illustrate the marked minimum between 17-19 km. As stated by the authors, MIPAS in this region is not reliable based on its variability. AIRS is dominated by it’s a priori. Accordingly, while I am not saying those profiles are correct in this region, and do not have a low bias, it is difficult to base that statement on the data in Figure 4. The difference between HALOE and SDLA in Figure 1 shows the same HALOE bias, requiring the authors to rely on the accuracy of SDLA to make that statement.

Summary

This paper needs major revisions with clearer treatment of instrument accuracies and conclusions regarding instrument performance. For an intercomparison paper to have
merit, these accuracy issues must be appropriately addressed.