# Response to reviewer 1:

We would like to thank the reviewer for their valuable comments and suggestions.

Reviewers comments are quoted in italic, our answers in roman, changes in the manuscript are in courier font. Page and line numbers refer to the original AMTD publication.

## Major comments:

*1. The main question concerns the motivation for and usefulness of the merged six satellite-based dataset. The idea of an unique temperature dataset in the upper troposphere and lower stratosphere (UTLS) region is very interesting, but each dataset has different characteristics which make it difficult to merge the data, even if the statistical model the authors describe in Section 3 is comprehensible and the focus is not the investigation of absolute differences between the iVRT and MSU4/AMSU9 data. I think the authors should give more motivation for the merging of the different datasets which is completely missing in the current version (introduction and also later in the text).*

The research and subsequent analyses presented in this paper was shaped in large part by the requirements of the ESA SPARC Initiative (SPIN) under which the work was funded. One of the main objectives of SPIN was to combine ESA and ESA Third Party Mission (TPM) data sets to improve the characterization of existing upper-air Climate Data Records (CDRs), and to produce new CDRs. As such, the choice of data sets used was dictated primarily by the scope of the SPIN project. While we understand that there are good reasons to only use data sets of a similar type when deciding whether to use only ESA/ESA-TPM data or to combine them with the RO data sets, we felt that combining these data sets would give us a more robust time series with smaller uncertainties as there is sufficient temporal and spatial overlap with the RO data to account for systematic biases between the different instruments.

**P. 238, l. 14, insert:**

The purpose of this study is to work towards the inclusion of ESA and ESA-TPM data in stratospheric climate data records (CDRs). As a number of key US-based CDRs ended in 2005/2006 while ESA and ESA-TPM vertical profile observation records begin a few years before 2005/2006 and continue to the present, ESA and ESA-TPM are potential candidates to extend existing CDRs in time. While similar comparisons of RO data only with (A)MSU have been performed (Ho et al. (2009), Ladstädter et al. (2011),Steiner et al. (2011) the motivation to use temperatures from different data sources, not only RO, was to create a new, multi-instrument temperature CDR with smaller uncertainties. The temporal and spatial overlap with the RO data should be sufficient to account for systematic biases between the different instruments.

*Would it not be more useful to perform the analysis with the MIPAS and combined CHAMP/GRACE/TSX data separately? A similar result as shown in Fig. 7 should be possible for the MIPAS (2002 to 2010) and radio occultation (2001 to 2012) data. This would make the results regarding the detected break-points in the merged MSU4/AMSU9 dataset more significant.*

Validations of (A)MSU data sets with RO data have been carried out (see review by Steiner et al. (2011)). While our break-point analysis could have been carried out with a merged RO record and the MIPAS data set, separately, it is not clear that this would make the results more significant.

*In addition to that one could compare also absolute values.*

This is true but it is not the purpose of this analysis i.e. our goal was not to validate the absolute values of the (A)MSU data but rather to focus purely on the statistical robustness of the transition from MSU4 to AMSU9. We feel that a comparison of absolute values would require a higher degree of diligence in weighting our merged vertically resolved temperature data set rather than using the single weighting profile that we have used.

*2. Another points is the choice of the datasets. Why the GFZ dataset and not the CHAMP, GRACE, TSX, and COSMIC data from UCAR with has a much broader database. The amount of ACE-FTS and SMR data are very small compared to the other data. Why not SABER data (since 2002) above 20 km?*

The SPIN activity specified the GFZ RO data sets over those available from UCAR. ACE-FTS and SMR data were used in this analysis as they are ESA/ESA-TMP instruments, while SABER is not.

*3. The radio occultation temperature data are dry temperatures. This causes a negative temperature bias between the dry and real temperatures in regions where atmospheric water vapor cannot be neglect. The lowest level the authors use is 300 hPa. In the extra-tropics the dry air assumption should be justified at and above 300 hPa, but could be violated in the tropical upper troposphere. Did the authors consider this when merging with the other data (MIPAS)?*

Thank you for bringing this to our attention. As our data set only starts at 300hPa, we believe it is safe to assume that the dry air condition holds in the extra-tropics. In the tropics, this assumption might be violated under some conditions. However, as the regression fit of Eq. (2) is done for each latitude zone and pressure level, separately, it allows for a latitudinal and altitudinal dependence that can produce bigger adjustments in the tropics at low altitudes than in the extra-tropics. As CHAMP is the initial data set used in the merging process, the resulting temperatures in our merged VRT data set should be considered dry temperatures. However, it is possible that the correction in the tropics could be better quantified and this would explain the differences seen in the tropics in comparison to COSMIC (Fig. 3).

**P. 241, l. 6, insert:**
RO instruments measure dry temperatures while the ESA and ESA-TMP instruments record physical temperatures. Our merging algorithm makes no explicit distinction between the two. As our data set only starts at 300hPa, it is safe to assume that the dry air condition holds in the extra-tropics (Danzer et al. (2014), Ladstädter et al. (2015)). In the tropics, this assumption might be violated under some conditions. However, as the regression fit of Eq. (2) is done for each latitude zone and pressure level, separately, it allows for a latitudinal and altitudinal dependence that can produce bigger adjustments in the tropics at low altitudes than in the extra-tropics. As CHAMP is the initial data set used in the merging process, the resulting temperatures in our merged VRT data set should be considered dry temperatures.

**References**

Danzer, J., Foelsche, U., Scherllin-Pirscher, B., and Schwärz, M.: Influence of changes in humidity on dry temperature in GPS RO climatologies, Atmos. Meas. Tech., 7, 2883--2896, doi: 10.5194/amt-7-2883-2014, 2014.

Ho, S.-P., Goldberg, M., Kuo, Y.-H., Zou, C.-Z. and Schreiner, W.: Calibration of Temperature in the Lower Stratosphere from Microwave Measurements Using COSMIC Radio Occultation Data: Preliminary Results, Terr. Atmos. Ocean. Sci., 20(1), 87--100, doi:10.3319/TAO.2007.12.06.01(F3C), 2009.

Ladstädter, F., Steiner, A. K., Foelsche, U., Haimberger, L., Tavolato, C., and Kirchengast, G.: An assessment of differences in lower stratospheric temperature records from (A)MSU, radiosondes, and GPS radio occultation, Atmos. Meas. Tech., 4, 1965-1977, doi:10.5194/amt-4-1965-2011, 2011.

Ladstädter, F., Steiner, A. K., Schwärz, M., and Kirchengast, G.: Climate intercomparison of GPS radio occultation, RS90/92 radiosondes and GRUAN from 2002 to 2013, Atmos. Meas. Tech., 8, 1819--1834, doi:10.5194/amt-8-1819-2015, 2015.

Steiner, A. K., Lackner, B. C., Ladstädter, F., Scherllin-Pirscher, B., Foelsche, U. and Kirchengast, G.: GPS radio occultation for climate monitoring and change detection, Radio Sci., 46(6), doi:10.1029/2010RS004614, 2011.