Manuscript prepared for Atmos. Meas. Tech. Discuss.
with version 2014/09/16 7.15 Copernicus papers of the LaTeX class copernicus.cls.
Date: 1 May 2015

Validation of merged MSU4 and AMSU9 temperature climate records with a new 2002–2012 vertically resolved temperature record

A. A. Penckwitt[1], G. E. Bodeker[1], P. Stoll[1], J. Lewis[1], T. von Clarmann[2], A. Jones[3], and A. I. Jonsson[3]

[1]Bodeker Scientific, Alexandra, New Zealand
[2]Karlsruhe Institute of Technology, Institute for Meteorology and Climate Research, Karlsruhe, Germany
[3]Department of Physics, University of Toronto, Toronto, Ontario, Canada

Correspondence to: A. A. Penckwitt (andreas@bodekerscientific.com)

Abstract

A new database of monthly mean zonal mean (5° zones) temperature time series spanning 17 pressure levels from 300 to 7 hPa and extending from 2002 to 2012 was created by merging monthly mean time series from two satellite-based mid-infrared spectrometers (ACE-FTS and MIPAS), a microwave sounder (SMR), and from three satellite-based radio occultation experiments (GRACE, CHAMP, and TSX). The primary intended use of this new temperature data set is to validate the merging of the Microwave Sounding Unit channel 4 (MSU4), and Advanced Microwave Sounding Unit channel 9 (AMSU9) temperature time series conducted in previous studies. The six source data sets were merged by removing offsets and trends between the different measurement series. Weighted means were calculated of the six source data sets where the weights were a function of the uncertainty on the original monthly mean data. This new temperature data set of the upper troposphere and stratosphere has been validated by comparing it to RATPAC-A, COSMIC radio occultation data as well as the NCEPCFSR reanalyses. Differences in all three cases were typically < 2 K in the upper troposphere and lower stratosphere, but could reach up to 5 K in the mid-stratosphere. The data across the 17 pressure levels have then been vertically integrated, using the MSU4/AMSU9 weighting function, to provide a deep vertical layer temperature proxy of the merged MSU4+AMSU9 series. Differences between this vertically integrated data set and two different versions of the MSU4+AMSU9 data set – one from Remote Sensing Systems and one from the University of Alabama at Huntsville – were examined for discontinuities. No statistically significant discontinuities were found in either of those two data sets suggesting that the transition from the MSU4+AMSU9 data to AMSU9 data only does not introduce any discontinuities in the MSU4+AMSU9 climate data records that might compromise their use in temperature trend analyses.

1 Introduction

Reliable long-term data records are necessary to detect and understand climate change. Vertically resolved trends in temperature provide a sensitive test of the mechanisms driving climate change (Thompson and Solomon, 2005; Karl et al., 2006; Randel et al., 2009) which in turn are vital to improving climate model projections of future climate. The National Oceanic and Atmospheric Administration (NOAA) operated Microwave Sounding Units (MSUs) on polar-orbiting satellite platforms from 1978 to 2005 and Advanced Microwave Sounding Units (AMSUs) from 1998 onwards. These nadir-looking instruments became the primary source of long-term temperature climate records of the troposphere and stratosphere even though their measurements were originally intended for short-term weather forecasts, not climate applications. To create long-term homogeneous climate data records, measurements need to be adjusted to account for inter-satellite biases, orbital changes and calibration deficiencies, as well as small differences in radiometer frequency and bandwidth between the MSU and AMSU instruments. Three such long-term records are currently maintained by independent groups – Remote Sensing Systems (RSS) (Mears and Wentz, 2009a, b), the University of Alabama at Hunstville (UAH) (Christy et al., 2003), and the Centre for Satellite Applications and Research (STAR) (Zou and Wang, 2011). While substantial differences in temperature trends were found between the RSS and UAH data sets in the stratosphere (Randel et al., 2009), trends in the lower stratosphere between (A)MSU data sets are generally comparable (Thompson et al., 2012).

Satellite-based infrared sounders, as well as Global Positioning System (GPS) radio occultation (RO) experiments, offer high vertical resolution making them prime candidates to extend the MSU temperature record and improve the monitoring of upper-air temperature changes. Temperature measurements from infrared sounders typically have a measurement uncertainty of $< 2\,\mathrm{K}$ with good horizontal resolution. RO measurements have moderate horizontal resolution ($\sim 300\,\mathrm{km}$), but good vertical resolution ($\sim 1\,\mathrm{km}$) and are well suited for high quality climatologies because of their low systematic errors ($< 0.5\,\mathrm{K}$) and high stability (Borsche et al., 2007). Ladstädter et al. (2011) found good agreement be-

tween RO and radiosonde temperature anomalies, but statistically significant differences in trend estimates from (A)MSU data compared to RO measurements which they believed resulted mainly from the (A)MSU data.

In this study, monthly mean zonal mean (5° zones) temperature time series from six satellite-based instruments, ~~including three~~ three European Space Agency (ESA) and Third Party Missions (TPMs) and three RO instruments, are merged to create a new time series of monthly mean temperatures on 17 pressure levels from 300 hPa to 7 hPa. This data set is then vertically integrated using the (A)MSU vertical weighting functions to create a (A)MSU proxy record suitable for assessing the continuity of the merged (A)MSU temperature data sets by the RSS and UAH groups, focusing on the Temperature Lower Stratosphere (TLS) product. The TLS series spans altitudes from 13 to 22 km and results from merging the MSU channel 4 (MSU4) and AMSU channel 9 (AMSU9) temperature time series.

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.

Section 2 describes the data sets used to compile our new vertically resolved temperature data set. In Sect. 3, the method used to merge the six source data sets is described. The merged data set is then validated against RATPAC (Radiosonde Atmospheric Temperature Products for Assessing Climate) radiosonde data, COSMIC (Constellation Observing System for Meteorology, Ionosphere, and Climate) RO data as well as the NCEPCFSR (National Centers for Environmental Prediction Climate Forecast System Reanalysis) reanalyses in Sect. 4. In Sect. 5, the new data set is vertically integrated using an appropriate

weighting function to create an (A)MSU proxy. The resulting temperature time series is then used to verify the merging of MSU4 and AMSU9 temperature series.

2 Data

Monthly mean zonal mean temperature records from satellite missions by ~~the European Space Agency (ESA) and Third Party Missions (TPMs)~~ESA and TPMs, as well as RO data, were created in collaboration with SPARC (Stratosphere–troposphere Processes And their Role in Climate) as a project of SPIN (ESA SPARC Initiative). In the following sections, we briefly describe the characteristics of the different instruments, and how the monthly mean zonal mean data sets used in this study were produced from the individual temperature profiles.

2.1 ESA and ESA-TPM temperature profiles

Temperature measurements from three instruments of ESA/ESA-TPM were used – the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS), the Atmospheric Chemistry Experiment – Fourier Transform Spectrometer (ACE-FTS) and the Sub-Millimeter Radiometer (SMR).

MIPAS is a mid-infrared limb emission Fourier transform spectrometer on board the Environmental Satellite (Envisat). This study is based on temperature data processed with the MIPAS research processor at the Institute for Meteorology and Climate Research (IMK) at the Karlsruhe Institute of Technology (KIT) in cooperation with the Institute de Astrofìsica de Andalucìa (IAA-CSIC). In particular, the high spectral resolution data V3O_T_8 were used from June 2002 to March 2004 with an uncertainty of 0.4–0.8 K at $\sim$ 3 km vertical resolution for stratospheric altitudes (von Clarmann et al., 2003). For the time period from April 2004 to December 2004 no MIPAS data were available. From January 2005 to April 2010 the reduced resolution data V5R_T_220 were used with an uncertainty of 0.5–1.4 K and a vertical resolution of 2–3.5 km (von Clarmann et al., 2009b; Stiller et al., 2012). The horizontal resolution of MIPAS temperatures is about 300 km for the high resolution data pre 2004

and about 120 to 400 km for the reduced resolution data post 2004 (von Clarmann et al., 2009a).

ACE-FTS is a Fourier transform spectrometer on the Canadian SCISAT satellite using solar occultation mid-infrared spectra to retrieve vertical profiles of temperature. The uncertainty on the ACE-FTS measurements has not been published, but the systematic error of the Version 2.2 data is estimated to be 2 K at 10–50 km (Sica et al., 2008). The vertical resolution of ACE-FTS measurements is $\sim$ 3–4 km (Bernath et al., 2005). As ACE-FTS makes no more than 30 measurements a day, its spatial and temporal resolutions are lower than those of the other instruments used in this study. Furthermore, ACE-FTS only measures certain latitudes for a given time of the year.

Our temperature data set was created from Version 3.0 data covering February 2004 to March 2011. The ACE-FTS product uses a priori information for pressure and temperature at low altitudes from the Canadian Meteorological Center (CMC). After our analysis was well underway, it was discovered that Version 3.0 data should only be used until the end of September 2010 because of problems with how the pressure and temperature information was extracted from the CMC models. These issues give inaccurate results for all of the retrievals past this date (Boone et al., 2013). In addition, quality flags were recently introduced for all level 2 version 2.5 and 3.5 data (Sheese et al., 2014), but no updated version of the monthly mean zonal mean temperature data set used in this study has been made available yet.

SMR is a passive microwave sounder on the Odin satellite with four tunable receivers in the $\sim$ 486–581 GHz spectral range as well as one mm-wave receiver. Temperature measurements in the stratosphere are made in the 544.6 GHz band. Version 2.0 data from August 2001 to April 2012 were used. The uncertainty of the temperature data is estimated to be 1–3 K while the vertical resolution for the 544.6 GHz band is 4–6 km. No information on any systematic errors is available.

2.2 RO temperature profiles

The GPS RO limb sounding technique has many advantages compared to infrared sounding measurements from space. The measurements are self-calibrating and independent of the mission. Thus, there is expected to be no satellite-to-satellite bias or instrumental drift. RO observations achieve global coverage with high vertical resolution of about 1 km in the lower stratosphere. Measurements are minimally affected by aerosols, clouds or precipitation, leading to very small systematic biases (equivalent to < 1 K; average systematic bias is < 0.1 K) and small measurement uncertainties (0.02–0.05 K) (Anthes, 2011).

The CHAMP (CHAllenging Minisatellite Payload) satellite was launched in 2000 (Wickert et al., 2005). It provided the first long-term continuous GPS RO data set from May 2001 to October 2008. The temperature profiles have a vertical resolution of 1.5 km and have a systematic bias of ~ 0.5 K between 5–20 km (Borsche et al., 2007).

The GRACE (Gravity Recovery and Climate Experiment) is a twin satellite configuration based on CHAMP heritage (Beyerle et al., 2005) with similar resolution and uncertainty to CHAMP (Wickert et al., 2005). The data set extends from January 2006 to December 2011.

The most recent RO data come from the Tracking, Occultation and Ranging instrument package on board of the TerraSAR-X (hereafter referred to as TSX) that was launched in 2007 (Beyerle et al., 2011). The vertical resolution in the upper troposphere/lower stratosphere is about 500 m. The data set spans July 2008 to March 2012.

CHAMP, GRACE and TSX data were obtained directly from the GeoForschungsZentrum (GFZ) Helmholtz Centre using Potsdam Occultation Software version 6.0.

2.3 Monthly mean zonal mean temperature records

The monthly mean zonal mean temperature records were derived in a format following the specifications of the SPARC Data Initiative (SPARC-DI)(Hegglin and Tegtmeier, 2015). The individual temperature profiles used to calculate the monthly mean zonal mean data sets were first carefully screened according to recommendations given in relevant quality control documents, in published literature, and/or according to best knowledge of the instrument

scientists. Linear interpolation in log pressure coordinates was used to interpolate individual temperature profiles onto the ~~pressure grid~~ SPARC-DI pressure grid, consisting of the following 28 pressure levels: 300, 250, 200, 170, 150, 130, 115, 100, 90, 80, 70, 50, 30, 20, 15, 10, 7, 5, 3, 2, 1.5, 1, 0.7, 0.5, 0.3, 0.2, 0.15, and 0.1 hPa. For instruments providing data on an altitude grid, conversion from altitude to pressure levels was done using retrieved temperature/pressure profiles or meteorological analyses. Zonal means were calculated as the average of all of the measurements on a given pressure level within ~~each~~ 5° latitude ~~zone~~zones (with midpoints at 87.5° S, 82.5° S, 77.5° S, …, 87.5 °N). The standard error of the mean of the measurements was used as uncertainty on the mean values within each latitude zone at each pressure level.

Additionally, the RO data were screened such that temperatures below 150 K were omitted as were temperatures above 330 K. ~~The interpolated values were corrected for their zonal mean and monthly mean representativeness using NCEPCFSR 6hourly temperature fields on pressure surfaces as~~

~~$$T_{\text{corr}} = T_{\text{RO}}(\theta, \phi, P, t) \times \frac{\overline{T}_{\text{CFSR}}(5°, P, \text{ month})}{T_{\text{CFSR}}(\theta, \phi, P, t)},$$~~

~~where T_{corr} is the bias corrected temperature value, T_{RO} is the RO temperature measurement interpolated onto pressure P at latitude θ, longitude ϕ at time t, $\overline{T}_{\text{CFSR}}(5°, P, \text{ month})$ is the NCEPCFSR 5monthly mean zonal mean temperature at pressure P, and T_{CFSR} is the NCEPCFSR temperature at the same time and location as T_{RO}. Applying Eq. (1) corrects the RO measurements for their sampling bias both in terms of geographical coverage and temporal coverage within the month of interest.~~

~~It may be noted that this mean representative correction was only applied to the RO measurements though it might also have been a useful correction for the other three instruments, especially in the case of ACE-FTS with its low sampling rate. For this reason, monthly mean temperatures that were obtained from only a small number of measurements are excluded from the merging process as described in the next section.~~

3 Constructing a single merged temperature climate data record

The monthly mean zonal mean temperature data sets from the three ESA/ESA-TPM instruments and the three RO experiments described in Sect. 2 were merged to create new zonal mean monthly mean temperature time series in 5° latitude zones, henceforth referred to as the vertically resolved temperature (VRT) climate record or data set. As not all of the six source data sets had measurements at all ~~27~~ 28 pressure levels mentioned in Sect. 2.3, only the first 17 pressure levels were used. RO measurements have been shown to be an effective climate benchmark below ~ 25 km altitude (Kuo et al., 2004; Alexander et al., 2014). Comparisons of the annual mean zonal mean temperatures calculated from the three RO data sets show agreement within 0.5 K in the lower stratosphere (~ 20 km) but there is a robust bias that increases in magnitude with height, reaching 3–5 K in the upper stratosphere, near 40 km. GRACE is consistently warmer than both CHAMP and TSX in the Southern Hemisphere and consistently colder in the Northern Hemisphere. Because of the systematic bias of GRACE against both CHAMP and TSX, we exclude GRACE as a suitable benchmark data set for the construction of the merged product. Choosing between CHAMP and TSX as an initial data set, CHAMP provides a longer data record.

Figure 1 shows graphically the process of how, starting with CHAMP as the benchmark data set, the remaining data sets are successively merged. At each latitude zone and pressure level, the data set with the biggest temporal overlap with CHAMP is chosen as the first candidate to be merged with CHAMP. In the example of Fig. 1, MIPAS is found to have the biggest overlap with CHAMP. Thus, differences are calculated between the CHAMP and MIPAS monthly mean temperatures and these are then modelled statistically using a regression model that includes an offset and drift where the offset fit coefficient is expanded

in Fourier pairs to also account for seasonality in the differences, i. e.:

$$f(t) = a_1 + a_2 t + a_3 \sin(2\pi t) + a_4 \cos(2\pi t) + a_5 \sin(4\pi t) + a_6 \cos(4\pi t). \quad (1)$$

The uncertainty of the difference function $f(t)$ takes into account possible correlations between the coefficients

$$\sigma_f^2(t) = \sum_{i=1}^{6} \left(\frac{\partial f(t)}{\partial a_i} \right)^2 \cdot \sigma_{a_i}^2 + 2 \sum_{i=1}^{5} \sum_{k=i+1}^{6} \frac{\partial f(t)}{\partial a_i} \frac{\partial f(t)}{\partial a_k} \cdot \mathrm{cov}(a_i, a_k) \quad (2)$$

The statistical model fit from Eq. (1) is then used to correct the MIPAS data

$$T_{\mathrm{MIPAS,corr}}(t) = T_{\mathrm{MIPAS}}(t) + f(t), \quad (3)$$

and the model's uncertainty is used to recalculate the uncertainty estimate of the corrected MIPAS data

$$\sigma_{\mathrm{MIPAS,corr}}(t) = \sqrt{\sigma_{\mathrm{MIPAS}}^2(t) + \sigma_f^2(t)}, \quad (4)$$

where σ_{MIPAS} is the uncertainty on the MIPAS mean temperatures. A new benchmark time series is created by merging the original CHAMP time series with the corrected MIPAS time series by calculating a weighted average

$$T_{\mathrm{merged}} = \frac{w_{\mathrm{CHAMP}} \cdot T_{\mathrm{CHAMP}} + w_{\mathrm{MIPAS,corr}} \cdot T_{\mathrm{MIPAS,corr}}}{w_{\mathrm{CHAMP}} + w_{\mathrm{MIPAS,corr}}}, \quad (5)$$

where the weights are given by $w_x = 1/\sigma_x^2$. The uncertainty in the correction applied to the MIPAS monthly means is incorporated into a new estimate of the uncertainty on the monthly mean

$$\sigma_{\mathrm{merged}} = \frac{1}{\sqrt{w_{\mathrm{CHAMP}} + w_{\mathrm{MIPAS,corr}}}} \quad (6)$$

The process is then repeated using each remaining data set in the order of decreasing temporal overlap with the iteratively revised benchmark data set. For the example in Fig. 1,

after merging MIPAS with CHAMP, this is found to be GRACE. Differences between the GRACE monthly mean temperatures and those of the new benchmark data set of CHAMP merged with corrected MIPAS are then used as input to a new statistical model to derive corrections to be applied to the GRACE monthly means. This process is repeated, consecutively folding the MIPAS, GRACE, TSX, ACE-FTS and SMR into the original CHAMP data to successively extend the CHAMP data set. The order in which the data sets are merged can be different for each latitude zone and pressure level depending on the availability of data.

The inclusion of each additional data set reduces the uncertainty on the final product (as long as they are appropriately corrected). The monthly mean data from an instrument are excluded from this merging process if there are fewer than 4 measurements in a particular month or the number of measurements in the month is less than 5 % of ~~measurements of the month with the highest number of measurements~~ the maximum within that year. The new VRT climate record consists of such a temperature time series for every 5° latitude zone and on the following 17 pressure levels: 300, 250, 200, 170, 150, 130, 115, 100, 90, 80, 70, 50, 30, 20, 15, 10, and 7 hPa.

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. (1) 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.

4 Validation of the VRT climate record

The merged temperature database described above was validated against the following three independent data sets:

1. The radiosonde-based RATPAC-A (Free et al., 2005) database.
2. A COSMIC (Anthes et al., 2008) RO zonal mean monthly mean temperature data set.
3. A NCEPCFSR (Saha et al., 2010) monthly mean zonal mean temperature data set.

The comparison of the vertically integrated temperature data set against the merged MSU4+AMSU9 data set (presented in Sect. 5) also provides a partial validation of the VRT data set.

4.1 Validation against RATPAC-A

The RATPAC-A data set provides annual anomalies on pressure levels and is aggregated over seven large geographical regions, namely a global data set, the Northern and Southern Hemispheres, the tropics (30° S to 30° N) and extra-tropics, as well as a narrower tropical zone from 20° S to 20° N. Area-weighted annual means were calculated for those seven geographical regions from our VRT database for comparison. Because RATPAC-A data are only provided as anomalies, and because the period over which the anomalies are calculated could not be ascertained from the available RATPAC-A documentation (nor from e.g. D. Seidel, personal communication, 2013), the RATPAC-A anomalies were first subtracted from the VRT annual time series at each pressure level and then the mean of the resultant time series was subtracted. If the inter-annual variability in our merged data set and RATPAC-A were identical, the resultant time series would be uniformly zero, i. e. we compare the ability of the two databases to track inter-annual variability in the temperature signals.

Figure 2 shows the comparison of the inter-annual variability for the Southern (a) and Northern (b) Hemispheres as a function of pressure and time. At ~~250~~50 hPa and at

~~300~~100 hPa differences in inter-annual variability can exceed 1 K while at ~~lower~~ higher pressures, differences are generally less than ± 0.5 K.

Differences in the inter-annual variability are smaller on all pressure levels in the tropics (Fig. 2c and d). In extra-tropical regions (Fig. 2e and f), the anomalies are in general ~~also small~~ less than 1 K at all pressure levels~~, but can be as large as ± 1.5~~. However, there are notable inter-annual differences. In 2004, the anomalies are up to 2 K ~~at low pressure levelsof 30 to 5~~higher in the Southern Hemisphere while up to 1 ~~hPa in 2004/05 as well as at 300~~K lower in the Northern Hemisphere across all pressure levels. In 2005 and 2006, positive differences of up to 1 K can be seen in both Hemispheres, but only at low pressure levels (< 100 hPa~~in 2011.~~). Similar comparisons are also available for the whole globe (not shown here). The ability of the VRT data set to track year-to-year variability over most of the upper troposphere and stratosphere at the ± 0.5 K level indicates that this database includes a valid representation of inter-annual variability in temperatures over this region (with the exception of the extra-tropics in the time period from 2004 to 2006).

4.2 Validation against a COSMIC radio occultation data set

The COSMIC RO data set used for this validation was created in the same way as the other RO data sets used in the construction of the VRT climate record described in Sect. 2.3. The COSMIC RO data set was retained as a means of validating the VRT data set, rather than being used in the construction of the data set. Comparisons between the VRT and COSMIC RO data sets were made at all 17 pressure levels for which the databases are available. Examples of comparisons at four different pressure levels are shown in Fig. 3. In general, temperature differences are more pronounced from 2009 onwards after the end of the CHAMP record that was used as an initial benchmark. At 10 hPa the VRT values tend to be up to ~~3~~5 K lower than the COSMIC temperatures south of $\sim 20^\circ$ S and up to ~~3~~5 K higher than the COSMIC temperatures north of $\sim 20^\circ$ S. At 100 hPa the differences are smaller, typically within ~~1~~2 K, with merged temperatures being lower than COSMIC temperatures at higher latitudes and higher than COSMIC temperatures at lower latitudes.

The latitude zone over the tropics where merged temperatures are higher than those of COSMIC diminishes in width with increasing pressure and at around 200 hPa the pattern reverses, i. e. merged temperatures in the extra-tropics are higher than COSMIC and lower in the tropics (see Fig. 3c and d).

South of $\sim 60^\circ$ S, there is a distinct temporal dependence with VRT temperatures considerably cooler in the months of October and November.

The temperature differences relative to the COSMIC data set are typically of the order of ~~1~~1.5 K at lower altitudes up to a pressure level of 50 hPa, but can become relatively large, reaching about 4 K at lower pressure levels of 7 hPa. There is also a distinct difference apparent between the tropics and extra-tropics. At low altitudes (up to 200 hPa) merged temperatures are typically lower in the tropics and higher in the extra-tropics, but the opposite in sign at higher altitudes (< 200 hPa). As this happens mainly after the end of the CHAMP record from 2009 onwards, this bias of VRT in the tropics might indicate problems with adjustments of ESA and ESA-TPM physical temperatures to dry temperatures. Where CHAMP data was available, the differences are with COSMIC are close to zero in the tropics, but increase after 2009 at the 200 hPa level (Fig. 3 d)

4.3 Validation against NCEPCFSR reanalyses

NCEPCFSR reanalyses were linearly interpolated onto the log pressure levels of the VRT data set and then area weighted monthly mean zonal means were calculated for all pressure levels.

Figure 4 shows differences in temperature between the VRT and the NCEPCFSR data sets. The VRT ~~data set can be up to 5~~temperatures are typically within ± 1 K ~~warmer than~~ of NCEPCFSR at 10 hPa~~with a larger bias in the Northern Hemisphere than in the Southern Hemisphere~~. At 90 hPa, ~~differences over the tropics are close to zero but with~~ VRT temperatures are typically lower than NCEPCFSR temperatures with the biggest differences in the tropics (up to 2 K) that also show a clear seasonal cycle. ~~These differences grow to 1.5–2K over higher latitudes.~~ At 250 hPa the differences are generally also between zero and 2 K, but smallest in the tropics. The left-hand column of Fig. 4 shows the mean differ-

ence for each latitude zone averaged over all months. A clear latitudinal bias can be seen with ~~temperature differences increasing from~~ the most negative values at about 20° ~~out to the poles , and also larger temperature differences over the equatorwith~~ and differences becoming smaller towards the poles and the equator. There are some seasonal variations, most striking over the tropics. ~~There is no discernible trend in the difference fields.~~

~~As opposed to other months in the year, VRT temperatures in December seem to be consistently lower than NCEPCFSR temperatures as indicated by the vertical stripes in the right hand column of Fig. 4b and c.~~

The results from these three validations suggest that the use of the ESA/ESA-TPM data together with the RO temperature data in the construction of the VRT data set results in a temperature time series that, with sufficient extension, is likely to be suitable for analyses of temperature trends in the upper troposphere and stratosphere.

5 Validation of merged MSU4 and AMSU9 data sets

One of the purposes of constructing the VRT data record was to assess the quality of the merging of the MSU4 and AMSU9 temperature time series in the lower stratosphere (the TLS product) by both the RSS and UAH group. Of particular interest is whether the transition from the period of overlap between MSU4 and AMSU9 (from 1998 to 2006) to coverage by AMSU9 only after 2006 introduces any discontinuity in the climate data record. A statistical model is used that accounts for any systematic biases between the VRT time series and the MSU4+AMSU9 data sets. The residuals are examined for any anomalies, in particular a step function around the transition to the AMSU9 only measurements around 2006.

5.1 Merged MSU4 and AMSU9 temperature time series

Microwave sounders retrieve vertical temperature profiles by measuring the thermal emission from oxygen molecules at different frequencies. The MSUs operating on a number of polar-orbiting NOAA satellites from 1978 to 2006 had four channels ranging from 50.3 to 57.95 GHz, measuring the atmospheric temperature in four thick layers from the surface

through to the lower stratosphere. Reliable operation of the MSU ceased in 2005, but there was already a significant number of missing data beginning in 2004 (Mears and Wentz, 2009a).

A series of follow-on instruments, the AMSUs, began operation in 1998 using a larger number of channels that not only improved the vertical resolution, but also extended measurements into the upper stratosphere not covered by MSU channels. By using the AMSU channels most closely matching the channels of the MSU instruments, MSU-based temperature series were extended to the present.

The temperature in the lower stratosphere (TLS) is measured by MSU channel 4 and is closely matched by AMSU channel 9. The amalgamated TLS product of MSU4+AMSU9 by RSS version 3.2 is described in Mears and Wentz (2009a). While the version 2.3 data set uses data only from one AMSU instrument, NOAA-15, in our study the updated version 3.3 data set is used that includes data from AMSU instruments on Aqua, NOAA-18, and METOP. Data from NOAA-16 are not used because of an unexplained drift during its lifetime.

Version 5.0 of the UAH data set contains AMSU data from both NOAA-15 and NOAA-16 (Christy et al., 2003). Subsequent versions improve the correction due to diurnal drift between the satellites (Mears and Wentz, 2005; Christy and Spencer, 2005), convert AMSU data to mimic MSU, and add additional AMSU measurements from AQUA, NOAA-18, NOAA-19, and METOP. In our study, we use the UAH version 5.6.

5.2 Vertically weighted layer mean temperature series for comparison with MSU4+AMSU9

As the nadir-looking microwave sounders measure the temperature in thick layers, our VRT time series needs to be integrated across the 17 pressure levels to obtain a single zonal mean monthly mean temperature series (iVRT) comparable to the MSU4+AMSU9 data sets. Usually the averaging kernel of the temperature retrieval is used for this purpose (Rodgers and Connor, 2003). In the case of the MSU and AMSU data products, however, no profile retrieval involving any constrained generalized inversion is performed, but the temperatures obtained from the microwave channels are directly assigned to the altitude

range to which the particular channel is sensitive. Thus, the weighting function of the related radiative transfer problem of this channel, which describes which fraction of the measured signal originates from which altitude, can be used directly in place of the averaging kernel.

RSS provides text files of vertical weighting functions for the different MSU channels on their ftp-server based on the US Standard Atmosphere (Remote Sensing Systems, 2014). While the exact form of the weighting function depends on the temperature, humidity, and liquid water content of the atmospheric column, the representative weighting function based on the mean state of the atmosphere for MSU4+AMSU9 is given in the file.

As this weighting function $\omega(p)$ is normalized to unity, the weighted means of our temperature time series are given by

$$T_{\text{weighted}} = \int \omega(\ln(p)) \cdot T(\ln(p)) \, \mathrm{d}\ln(p). \tag{7}$$

The uncertainty on the weighted mean is calculated as a weighted average of the measurement uncertainties

$$\sigma_{\text{weighted}} = \int \omega(\ln(p)) \cdot \sigma(\ln(p)) \, \mathrm{d}\ln(p). \tag{8}$$

Vertically integrated values are considered to be valid only if data from all 17 pressure levels are available.

Figure 5 illustrates the calculation of the vertically integrated temperature for the latitude zone 35 to 40° N in May 2002 as an example.

5.3 Break-point analysis of the merged MSU4 and AMSU9 data

Vertically integrating our VRT data set creates a proxy for the merged MSU4+AMSU9 temperature series. This proxy was then compared to the two MSU4+AMSU9 data sets available from the RSS and UAH groups in each 5° latitude zone (see Fig. 6 for an example). Our primary interest is not the absolute difference between our iVRT data set and the MSU4+AMSU9 data sets but rather whether there are any discontinuities in the differences

that might be indicative of steps in the MSU4+AMSU9 data resulting from imperfect merging of the source data. To this end a regression model consisting of an offset term and linear trend term was fitted to the difference time series (panel b of Fig. 6). Both the offset and linear trend fit coefficients were expanded in three Fourier terms to account for seasonality. In this way the regression model accounts for any systematic biases between the vertically integrated database and the merged MSU4+AMSU9 databases. Any steps in the residuals from those regression model fits would be indicative of problems either with the merging of the MSU4 and AMSU9 time series or with the construction of our iVRT data set.

Since it was not possible to determine definitively when the switch from using MSU to using AMSU data in the combined databases happened, the Standard Normal Homogeneity Test (SNHT) (Alexandersson and Moberg, 1997) was used to statistically test for any break-points in the residuals from the start of 2005 to the end of 2007 in all 36 latitude zones. Two variants of the SNHT were applied. In the first variant, the standard deviation (SD) of the time series is assumed to be the same before and after a potential break. Critical values from Khaliq and Ouarda (2007) at a 95 % level of significance were used (linearly interpolated to the exact number of data points). In the second variant, the SDs within the two parts of the series, before and after a possible break, were allowed to differ from each other. For this test, critical values were taken from the original Alexandersson and Moberg (1997) publication.

Assuming a constant SD, the SNHT did not detect any break-points in either the RSS or UAH data sets. However, when the SD is allowed to change after the break-point, ~~the three latitude zones shown in Fig. **??**~~ five latitude zones exhibited a break-point in the residuals for the ~~UAH data set (~~RSS data set indicated by the vertical dashed lines ~~).~~in Fig. 7, and seven for the UAH data set shown in Fig. 8.

While there are no discernible steps in the residuals for any of the ~~three~~ depicted zones, the SDs change after the break-point. For ~~the zone from 60 to 65N, the SD appears to decrease after~~ RSS, the SDs appear to decrease in all of the zones after the break-points. This is also the case for most of the ~~break-point in September 2006 (Fig. **??**b). However, the critical value of 16.88 is only barely exceeded by the test value of 17.54 at a 5level of~~

significance (Table **??**). In the absence of break-points across a number of neighbouring latitude zones , this break-point can be considered a statistical artefact.

For zones compared to UAH except for the zones close to the poles (Fig. **??**a and c), 8a and g)where the SDs increase considerablyin 2005. As the number of measurements close to the poles is generally small, an increase in the variability does not necessarily indicate an anomaly with the merging process. Unlike the UAH database, the RSS merged data set does not report any monthly mean temperatures for either polar zone.

Three of the break-points in comparison to RSS are also detected relative to UAH, namely 70–75° S (Figs. . 7a and . 8b), 30–35° S (Figs. . 7b and . 8d), and 70–75° N (Figs. . 7e and . 8e). Comparing the month at which the break-points occur for these three common zones between RSS and UAH (Tables 1 and 2) reveals differences of up to three months. As the SNHT only gives the most probable month at which a break-point occurs, those differences can be considered to be within the sensitivity of the test.

While most of the break-points occur around the end of 2005 / beginning of 2006, break-points occur throughout the whole two years search period (the earliest in January 2005 and the latest in November 2006). If break-points were due to a problem with the merging of the MSU4 and AMSU9 time series, all break-points would be expected to be around the same time (when MSU4 is switched to AMSU9).

To establish whether the detected increase changes in variability of the residuals comes from the UAH or iVRT data set, the break-point analysis was repeated, but the iVRT data were averaged by month over all years (i.e. the same climatology was used for all years). Compared to this repeated climatology, the residuals of the UAH data set showed no increase in variability in the polar zoneswhere most of the previously detected break-points were detected previously indicating that the change in the SDs observed in Fig. **??**a and c disappear with three exceptions: for RSS a break-point is still detected in the 70–75° S latitude zone, and for UAH in the 70–75° S and 85–90° S zones. All of these break-points relative to the repeated climatology appear in November 2006. While this is consistent within the SHNT's sensitivity to the previously detected break-points in the 70–75° S zone, the break-point in the 85–90° S relative to UAH was in December 2005 (see Table 2). How-

ever, there is a secondary break-point in comparison to UAH in the 85–90° S zone in the in December 2006 with a test value of 30.5 which is almost as high as the test value of 30.78 reported in Table 2. This secondary break-point is within a month of the November 2006 break-point of the repeated climatology. This would indicate that there is some detectable change in variability both in the RSS and UAH database around November 2006 in the southern-most latitude zones. All other changes in variability of the residuals in Figs. 7 and 8 are likely caused by the iVRT data set.

As there are no steps found in the residuals we can conclude that the merged MSU4+AMSU9 temperature series of both the RSS and UAH group are homogeneous in time and provide climate data records devoid of discontinuities that are suitable for long-term trend detection. ~~The~~ Most of the changes in the SDs of the residuals ~~in the polar zones for the UAH database~~ for the RSS and UAH databases are likely caused by an increase in variability in the iVRT data set ~~.~~ except for those in November 2006 in the 70–75° S and 85–90° S zones that indicate some change in variability both in the RSS and UAH database

6 Discussion and conclusions

A new database of monthly mean zonal mean temperatures spanning the upper troposphere and stratosphere was created by merging monthly mean zonal mean temperature measurements from the mid-infrared spectrometers ACE-FTS and MIPAS, the microwave sounder SMR, and the RO experiments GRACE, CHAMP and TSX. The new temperature data record is aggregated in 5° latitudinal zones and is vertically resolved on pressure levels from 300 to 7 hPa.

Systematic biases between different instruments were corrected by statistically modelling the differences allowing for an offset and drift as well as temporal periodic fluctuations. The merging process was initiated with the CHAMP temperature time series as RO measurements are known to provide an effective climate benchmark. As the merging algorithm mainly removes systematic biases relative to the initial data set, it is crucial to choose this

initial benchmark data set carefully. The sensitivity to the initial benchmark data set was tested by using the same merging algorithm, but starting the process with MIPAS, GRACE and TSX, respectively, rather than CHAMP. A comparison of these latter three temperature time series to that initiated with CHAMP shows all resultant merged data sets being in very good agreement. The main differences between the different data sets is an offset of similar magnitude to the offset seen in the right hand column of Fig. 1. The merged temperature set starting with CHAMP typically lies somewhere in the middle of the three comparative data sets. However, for some latitude zones and pressure levels, especially in the Southern Hemisphere, the other three temperature time series show a trend relative to the one started with CHAMP. A trend in the differences is more problematic because one intended use of our merged data set is to detect relatively small warming or cooling trends in the atmosphere. This trend is relatively small for TSX, but bigger for MIPAS and GRACE. Nevertheless, the comparison shows that CHAMP remains the best choice as an initial benchmark data set in the merging process.

The merged VRT climate record was validated against three different databases. The first validation was done against RATPAC-A which provides annual mean temperature anomalies from radiosonde measurements aggregated over seven large geographical regions. The inter-annual variability in temperatures is well represented by our VRT data set over most of the upper troposphere and stratosphere at the ± 0.5 K level, but can exceed 1 K at the 250 and 300 hPa levels. The comparison with the COSMIC RO database showed more divergent results. Temperature differences were typically up to ~~1~~1.5 K at most pressure levels (≥ 50 hPa), but could reach up to ~~4~~5 K at lower pressure levels. There were also clear spatial differences between tropical and extra-tropical zones depending on altitude. The temperatures from our VRT database were consistently lower in the tropics than in the extra-tropics above 200 hPa, but this pattern reversed below 200 hPa. The third validation was performed against NCEPCFSR reanalyses. Differences were typically less than 2 K over most pressure levels, but could reach 5 K at the highest altitudes (≤ 10 hPa). Seasonal variations, on the other hand, were more pronounced than in the other two validations.

As a major application, the VRT temperature record was used to verify the quality of the merging process of the MSU4 and AMSU9 channels of both the RSS and UAH groups. After removing systematic biases, the residuals relative to our iVRT data set were examined for statistically significant break-points. No statistically significant steps were found in the residuals around the switch from MSU4 to AMSU9, confirming that both groups made appropriate adjustments in the merging process to assure a continuous temperature time series that is not affected by calibration errors between the two types of instruments or other systematic biases due to satellite drifts. ~~Only in the two polar 5latitude zonesdid our~~ In various latitude zones, the residuals of the iVRT temperature record ~~show an increase in variability in 2005 relative to the UAH temperature series~~with respect to both the RSS and UAH temperature time series show a change in variability A break-point analysis with repeated climatologies showed that these changes are mostly due to the iVRT data set, but there is some evidence that changes in variability in November 2006 at latitudes South of 70° S, are present in both the RSS and UAH merged (A)MSU records.

The VRT climate record can also be seen as a contribution to establish long-term vertically resolved temperature records to enable updated knowledge of long-term changes in temperatures across different altitude ranges. To date, the microwave sounding measurements have been a de-facto standard for temperature climate data records. With the increased use of satellite infrared-sounders as well as highly accurate RO measurements, more finely vertically-resolved data records are available. Our VRT record is, to the best of our knowledge, the first that merged temperature measurements from selected ESA/ESA-TPM missions with RO measurements. ~~It has been shown that the~~ As the number of instruments used in the merging increases, the uncertainty on the monthly mean zonal mean temperatures decreases~~with an increased number of instruments used in the merging~~. The VRT climate record thus provides a useful data set for long-term temperature trend analyses.

Acknowledgements. This work was funded in large part through the Support to Science Element ESA (European Space Agency) SPARC (Stratosphere–troposphere Processes And their Role in Climate) Initiative (ESA Contract No. 4000105291/12/I-NB) and we take this opportunity to acknowl-

edge ESA for this support. We acknowledge the individual instrument teams (GRACE, CHAMP, TSX, ACE-FTS, MIPAS, and SMR) and the respective space agencies for making their measurements available, and the ESA SPARC Initiative (SPIN) for organizing and coordinating the compilation of the monthly mean zonal mean temperature data sets used in this work. ACE is a Canadian-led mission mainly supported by the Canadian Space Agency (CSA). Development of the ACE-FTS temperature climatology was supported by a grant from the CSA. Special thanks go to Bernd Funke and Kaley Walker for helpful comments on an early draft of this paper. We also acknowledge NOAA (National Oceanic and Atmospheric Administration) for full and open data access to RATPAC temperature anomalies through their National Climatic Data Center (NCDC). We would like to thank Torsten Schmidt from GeoForschungsZentrum (GFZ) Potsdam for providing the COSMIC radio occultation data, and Chi-Fan Shih from the National Center for Atmospheric Research in Boulder, CO, for providing the NCEPCFSR reanalyses. Last but not least, we would like to acknowledge Joachim Urban, not only for his work on the SMR data set used in this article, but also his general contribution to the scientific community. Sadly, he passed away in August 2014.

References

Alexander, P., de la Torre, A., Llamedo, P., and Hierro, R.: Precision estimation in temperature and refractivity profiles retrieved by GPS radio occultations, J. Geophys. Res.-Atmos., 119, 8624–8638, doi:10.1002/2013JD021016, 2014.

Alexandersson, H. and Moberg, A.: Homogenization of Swedish temperature data, Part I: Homogeneity test for linear trends, Int. J. Climatol., 17, 25–34, doi:10.1002/(SICI)1097-0088(199701)17:1<25::AID-JOC103>3.0.CO;2-J, 1997.

Anthes, R. A.: Exploring Earth's atmosphere with radio occultation: contributions to weather, climate and space weather, Atmos. Meas. Tech., 4, 1077–1103, doi:10.5194/amt-4-1077-2011, 2011.

Anthes, R. A., Ector, D., Hunt, D. C., Kuo, Y.-H., Rocken, C., Schreiner, W. S., Sokolovskiy, S. V., Syndergaard, S., Wee, T.-K., Zeng, Z., Bernhardt, P. A., Dymond, K. F., Chen, Y., Liu, H., Manning, K., Randel, W. J., Trenberth, K. E., Cucurull, L., Healy, S. B., Ho, S.-P., McCormick, C., Meehan, T. K., Thompson, D. C., and Yen, N. L.: The COSMIC/FORMOSAT-3 mission: early results, B. Am. Meteorol. Soc., 89, 313–333, doi:10.1175/BAMS-89-3-313, 2008.

Bernath, P. F., McElroy, C. T., Abrams, M. C., Boone, C. D., Butler, M., Camy-Peyret, C., Carleer, M., Clerbaux, C., Coheur, P.-F., Colin, R., DeCola, P., DeMazière, M., Drummond, J. R., Dufour, D.,

Evans, W. F. J., Fast, H., Fussen, D., Gilbert, K., Jennings, D. E., Llewellyn, E. J., Lowe, R. P., Mahieu, E., McConnell, J. C., McHugh, M., McLeod, S. D., Michaud, R., Midwinter, C., Nassar, R., Nichitiu, F., Nowlan, C., Rinsland, C. P., Rochon, Y. J., Rowlands, N., Semeniuk, K., Simon, P., Skelton, R., Sloan, J. J., Soucy, M.-A., Strong, K., Tremblay, P., Turnbull, D., Walker, K. A., Walkty, I., Wardle, D. A., Wehrle, V., Zander, R., and Zou, J.: Atmospheric Chemistry Experiment (ACE): mission overview, Geophys. Res. Lett., 32, L15S01, doi:10.1029/2005GL022386, 2005.

Beyerle, G., Schmidt, T., Michalak, G., Heise, S., Wickert, J., and Reigber, C.: GPS radio occultation with GRACE: Atmospheric profiling utilizing the zero difference technique, Geophys. Res. Lett., 32, L13806, doi:10.1029/2005GL023109, 2005.

Beyerle, G., Grunwaldt, L., Heise, S., Köhler, W., König, R., Michalak, G., Rothacher, M., Schmidt, T., Wickert, J., Tapley, B. D., and Giesinger, B.: First results from the GPS atmosphere sounding experiment TOR aboard the TerraSAR-X satellite, Atmos. Chem. Phys., 11, 6687–6699, doi:10.5194/acp-11-6687-2011, 2011.

Boone, C. D., Walker, K. A., and Bernath, P. F.: Version 3 retrievals for the Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS), in: The Atmospheric Chemistry Experiment ACE at 10: a Solar Occultation Anthology, edited by: Bernath, P. F., A. Deepak Publishing, Hampton, Virginia, USA, 103–127, 2013.

Borsche, M., Kirchengast, G., and Foelsche, U.: Tropical tropopause climatology as observed with radio occultation measurements from CHAMP compared to ECMWF and NCEP analyses, Geophys. Res. Lett., 34, L03702, doi:10.1029/2006GL027918, 2007.

Christy, J. R. and Spencer, R. W.: Correcting temperature data sets, Science, 310, 972–973, doi:10.1126/science.310.5750.972, 2005.

Christy, J. R., Spencer, R. W., Norris, W. B., Braswell, W. D., and Parker, D. E.: Error estimates of version 5.0 of MSU–AMSU bulk atmospheric temperatures, J. Atmos. Ocean. Tech., 20, 613–629, doi:10.1175/1520-0426(2003)20<613:EEOVOM>2.0.CO;2, 2003.

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.

Free, M., Seidel, D. J., Angell, J. K., Lanzante, J., Durre, I., and Peterson, T. C.: Radiosonde Atmospheric Temperature Products for Assessing Climate (RATPAC): a new data set of large-area anomaly time series, J. Geophys. Res.-Atmos., 110, doi:10.1029/2005JD006169, 2005.

Hegglin, M. I. and Tegtmeier, S. (Eds.): SPARC Data Initiative, SPARC report on the evaluation of trace gas and aerosol climatologies from satellite limb sounders, in preparation, 2015.

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.

Karl, T. R., Hassol, S. J., Miller, C. D., and Murray, W. L. (Eds.): U.S. Climate Change Science Program and the Subcommittee on Global Change Research, Temperature Trends in the Lower Atmosphere: Steps for Understanding and Reconciling Differences, UNT Digital Library, Washington, DC, USA, available at: http://digital.library.unt.edu/ark:/67531/metadc12017/, 2006.

Khaliq, M. N. and Ouarda, T. B. M. J.: On the critical values of the standard normal homogeneity test (SNHT), Int. J. Climatol., 27, 681–687, doi:10.1002/joc.1438, 2007.

Kuo, Y.-H., Wee, T.-K., Sokolovskiy, S., Rocken, C., Schreiner, W., Hunt, D., and Anthes, R.: Inversion and Error Estimation of GPS Radio Occultation Data:, J. Meteorol. Soc. Jpn., 82, 507–531, doi:10.2151/jmsj.2004.507, 2004.

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.

Mears, C. A. and Wentz, F. J.: The effect of diurnal correction on satellite-derived lower tropospheric temperature, Science, 309, 1548–1551, doi:10.1126/science.1114772, 2005.

Mears, C. A. and Wentz, F. J.: Construction of the Remote Sensing Systems V3.2 Atmospheric Temperature Records from the MSU and AMSU Microwave Sounders, J. Atmos. Ocean. Tech., 26, 1040–1056, doi:10.1175/2008JTECHA1176.1, 2009a.

Mears, C. A. and Wentz, F. J.: Construction of the RSS V3.2 lower-tropospheric temperature dataset from the MSU and AMSU microwave sounders, J. Atmos. Ocean. Tech., 26, 1493–1509, doi:10.1175/2009JTECHA1237.1, 2009b.

Randel, W. J., Shine, K. P., Austin, J., Barnett, J., Claud, C., Gillett, N. P., Keckhut, P., Langematz, U., Lin, R., Long, C., Mears, C., Miller, A., Nash, J., Seidel, D. J., Thompson, D. W. J., Wu, F., and Yoden, S.: An update of observed stratospheric temperature trends, J. Geophys. Res.-Atmos., 114, D02107, doi:10.1029/2008JD010421, 2009.

Remote Sensing Systems: Weighting Function Channel 4, available at: ftp://ftp.ssmi.com/msu/weighting_functions/std_atmosphere_wt_function_chan_4.txt (last access: 6 October 2014), 2014.

Rodgers, C. D. and Connor, B. J.: Intercomparison of remote sounding instruments, J. Geophys. Res., 108, 4116, doi:10.1029/2002JD002299, 2003.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y.-T., Chuang, H.-Y., Juang, H.-M. H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Van Delst, P., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., Van Den Dool, H., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.-K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C.-Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R. W., Rutledge, G., and Goldberg, M.: The NCEP climate forecast system reanalysis, B. Am. Meteorol. Soc., 91, 1015–1057, doi:10.1175/2010BAMS3001.1, 2010.

Sheese, P. E., Boone, C. D., and Walker, K. A.: Technical note: Detecting outliers in satellite-based atmospheric measurements, Atmos. Meas. Tech. Discuss., 7, 8393–8414, doi:10.5194/amtd-7-8393-2014, 2014.

Sica, R. J., Izawa, M. R. M., Walker, K. A., Boone, C., Petelina, S. V., Argall, P. S., Bernath, P., Burns, G. B., Catoire, V., Collins, R. L., Daffer, W. H., De Clercq, C., Fan, Z. Y., Firanski, B. J., French, W. J. R., Gerard, P., Gerding, M., Granville, J., Innis, J. L., Keckhut, P., Kerzenmacher, T., Klekociuk, A. R., Kyrö, E., Lambert, J. C., Llewellyn, E. J., Manney, G. L., McDermid, I. S., Mizutani, K., Murayama, Y., Piccolo, C., Raspollini, P., Ridolfi, M., Robert, C., Steinbrecht, W., Strawbridge, K. B., Strong, K., Stübi, R., and Thurairajah, B.: Validation of the Atmospheric Chemistry Experiment (ACE) version 2.2 temperature using ground-based and space-borne measurements, Atmos. Chem. Phys., 8, 35–62, doi:10.5194/acp-8-35-2008, 2008.

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.

Stiller, G. P., Kiefer, M., Eckert, E., von Clarmann, T., Kellmann, S., García-Comas, M., Funke, B., Leblanc, T., Fetzer, E., Froidevaux, L., Gomez, M., Hall, E., Hurst, D., Jordan, A., Kämpfer, N., Lambert, A., McDermid, I. S., McGee, T., Miloshevich, L., Nedoluha, G., Read, W., Schneider, M., Schwartz, M., Straub, C., Toon, G., Twigg, L. W., Walker, K., and Whiteman, D. N.: Validation of MIPAS IMK/IAA temperature, water vapor, and ozone profiles with MOHAVE-2009 campaign measurements, Atmos. Meas. Tech., 5, 289–320, doi:10.5194/amt-5-289-2012, 2012.

Thompson, D. W. J. and Solomon, S.: Recent stratospheric climate trends as evidenced in radiosonde data: global structure and tropospheric linkages, J. Climate, 18, 4785–4795, doi:10.1175/JCLI3585.1, 2005.

Thompson, D. W. J., Seidel, D. J., Randel, W. J., Zou, C.-Z., Butler, A. H., Mears, C., Osso, A., Long, C., and Lin, R.: The mystery of recent stratospheric temperature trends, Nature, 491, 692–697, doi:10.1038/nature11579, 2012.

von Clarmann, T., Glatthor, N., Grabowski, U., Höpfner, M., Kellmann, S., Kiefer, M., Linden, A., Tsidu, G. M., Milz, M., Steck, T., Stiller, G. P., Wang, D. Y., Fischer, H., Funke, B., Gil-López, S., and López-Puertas, M.: Retrieval of temperature and tangent altitude pointing from limb emission spectra recorded from space by the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS), J. Geophys. Res.-Atmos., 108, 4736, doi:10.1029/2003JD003602, 2003.

von Clarmann, T., De Clercq, C., Ridolfi, M., Höpfner, M., and Lambert, J.-C.: The horizontal resolution of MIPAS, Atmos. Meas. Tech., 2, 47–54, doi:10.5194/amt-2-47-2009, 2009a.

von Clarmann, T., Höpfner, M., Kellmann, S., Linden, A., Chauhan, S., Funke, B., Grabowski, U., Glatthor, N., Kiefer, M., Schieferdecker, T., Stiller, G. P., and Versick, S.: Retrieval of temperature, H_2O, O_3, HNO_3, CH_4, N_2O, $ClONO_2$ and ClO from MIPAS reduced resolution nominal mode limb emission measurements, Atmos. Meas. Tech., 2, 159–175, doi:10.5194/amt-2-159-2009, 2009.

Wickert, J., Beyerle, G., König, R., Heise, S., Grunwaldt, L., Michalak, G., Reigber, Ch., and Schmidt, T.: GPS radio occultation with CHAMP and GRACE: A first look at a new and promising satellite configuration for global atmospheric sounding, Ann. Geophys., 23, 653–658, doi:10.5194/angeo-23-653-2005, 2005.

Zou, C.-Z. and Wang, W.: Intersatellite calibration of AMSU-A observations for weather and climate applications, J. Geophys. Res.-Atmos., 116, D23113, doi:10.1029/2011JD016205, 2011.

Table 1. Break-points detected in ~~UAH~~ residuals of iVRT relative to RSS database. The month shown in the table is the last month before the break-point.

Zone	Month	T value	Crit. value
~~85–90~~70–75° S	~~Apr~~ Nov 2006	20.96	16.86
30–35° S	Dec 2005	~~61.54~~ 19.27	16.88
~~60–65~~5–10° N	~~Sep~~ Jan 2005	25.02	16.86
10–15° N	May 2005	24.67	16.86
70–75° N	Jan 2006	~~17.54~~ 27.30	16.86

Table 2. Break-points detected in residuals of iVRT relative to UAH database. The month shown in the table is the last month before the break-point.

Zone	Month	T value	Crit. value
85–90° S	Dec 2005	30.78	16.86
70–75° S	Oct 2006	29.98	16.86
60–65° S	Nov 2005	34.03	16.88
30–35° S	Feb 2006	17.97	16.88
70–75° N	Apr 2006	39.71	16.86
75–80° N	Apr 2005	18.43	16.86
85–90° N	Sep 2005	~~60.10~~ 26.18	~~16.88~~16.85

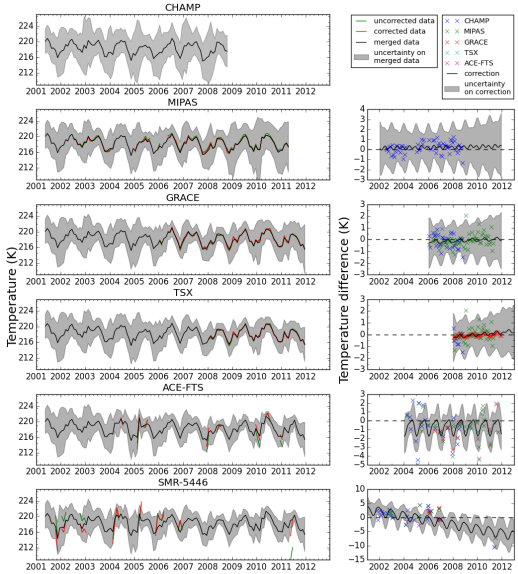

Figure 1. An example of the generation of the VRT database, in this case showing the time series at 100 hPa and between 50 and 55° N. The upper left panel shows the CHAMP monthly means and their 1σ uncertainties which form the initial benchmark data set for the merging. The second row of panels show the uncorrected MIPAS monthly mean temperatures (green) in the leftmost panel and their differences against the CHAMP data in the rightmost panel. The regression model fit to these differences is shown as a black line in the rightmost panel together with the uncertainty on the regression model fit (grey shaded area). The regression model fit is then used to correct the MIPAS data which is then shown as the red line in the leftmost panel. The black line and grey shaded region in the leftmost panel show what a merged CHAMP and (corrected) MIPAS data set would look like. The remaining data sets are then successively folded in as described in the text. The final resultant data set is displayed as a black line and grey shaded region in the bottom left panel.

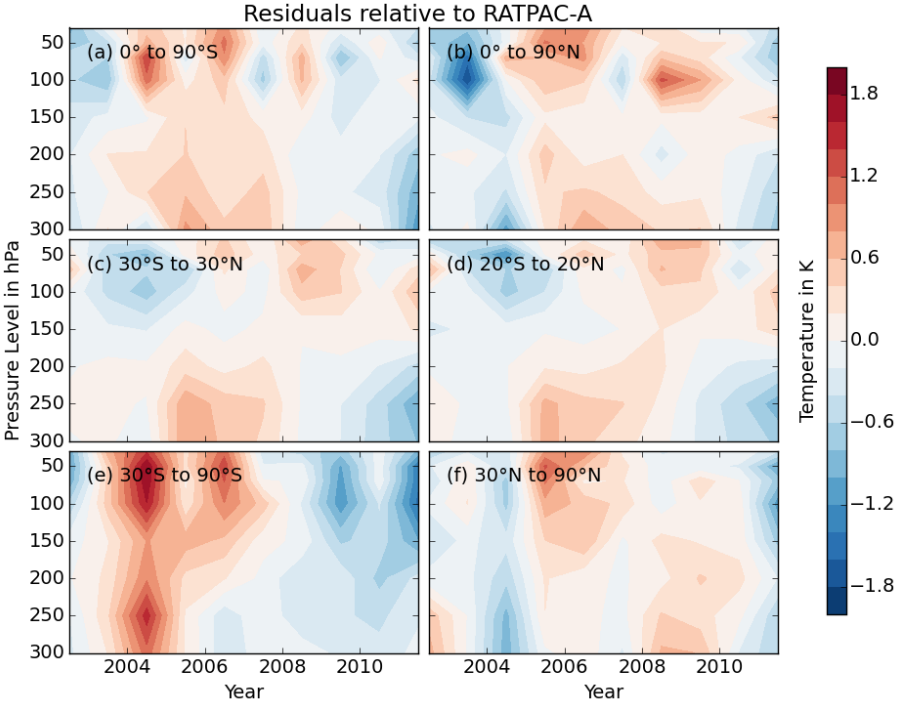

Figure 2. Differences between VRT anomalies and those from the RATPAC-A database for **(a)** the Southern Hemisphere, **(b)** the Northern Hemisphere, **(c)** the tropics from 30° S to 30° N, **(d)** the tropics from 20° S to 20° N, **(e)** the extra-tropics in the Southern Hemisphere, and **(f)** the extra-tropics in the Northern Hemisphere.

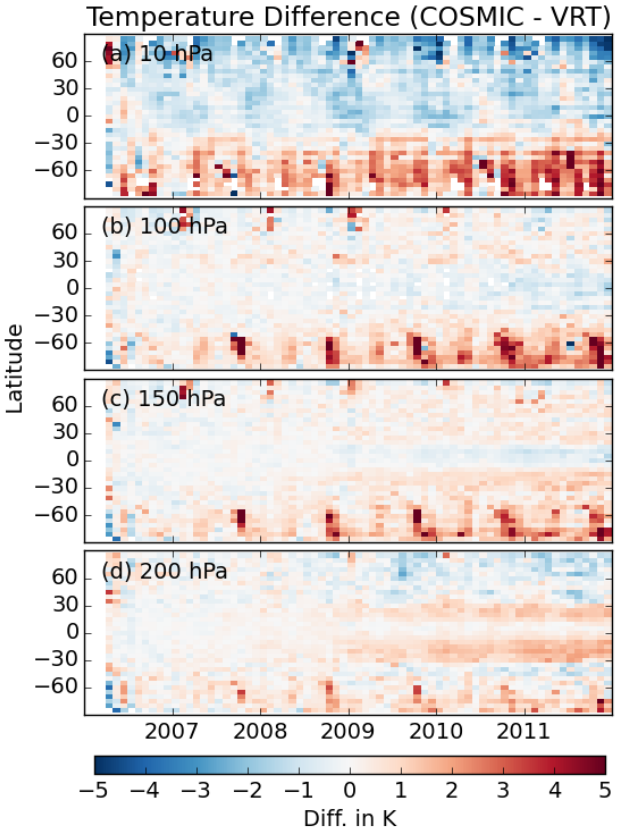

Figure 3. Temperature differences between the COSMIC RO and the VRT data sets at 10, 100, 150 and 200 hPa.

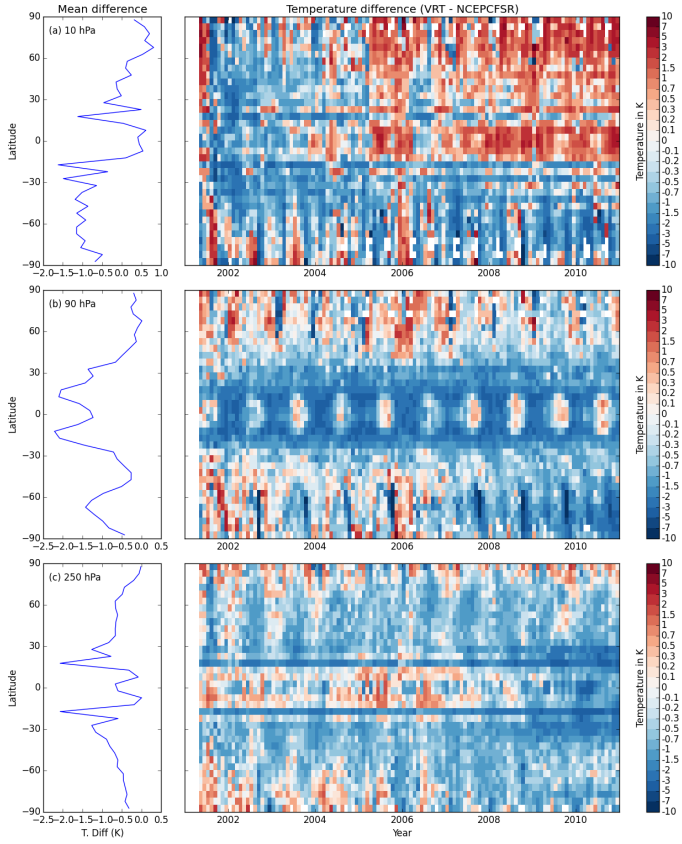

Figure 4. Differences between monthly mean zonal mean (5° zones) temperatures from the VRT data set and from the NCEPCFSR data set. Results are shown for three different pressure levels viz. 10 hPa (upper panel), 90 hPa (middle panel) and 250 hPa (lower panel). The left hand column shows the mean differences for each latitude zone averaged over all months.

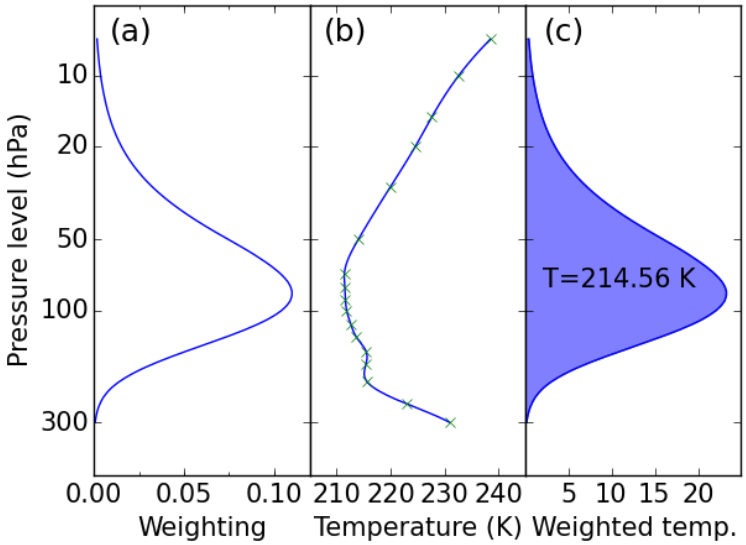

Figure 5. Calculation of weighted temperature in latitude zone 35 to 40° N in May 2002. **(a)** Normalized weighting function of MSU4+AMSU9. **(b)** Temperature profile. **(c)** Weighted temperature profile $\omega(p) \cdot T(p)$ and the integrated temperature T_{weighted} in the shaded region.

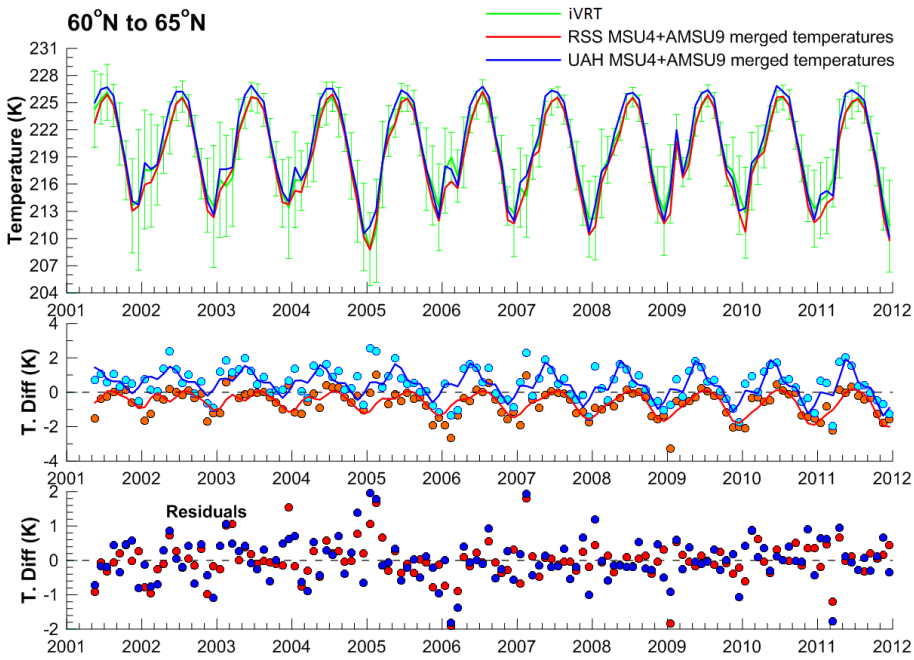

Figure 6. An example of the comparison of the weighted iVRT data set and the two merged MSU4+AMSU9 data sets. **(a)** The original raw monthly mean time series, **(b)** regression model fits (lines) to the differences between the RSS and iVRT time series (cyan dots and blue line showing the regression model fit) and between the UAH and iVRT time series (orange dots and red line showing the regression model fit), **(c)** the residuals from the regression model fits shown in panel **(b)**.

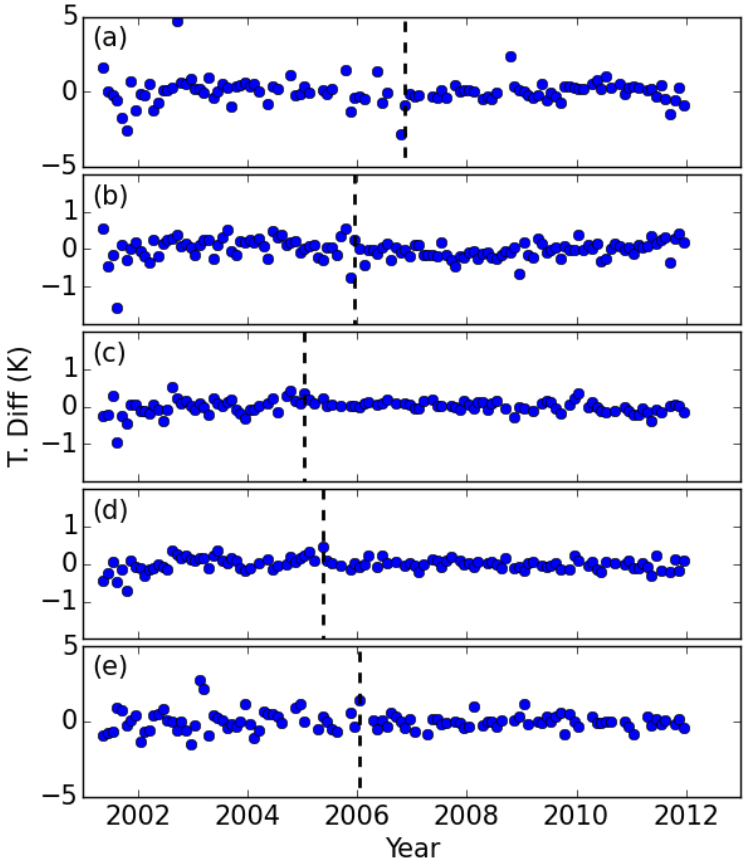

Figure 7. Latitudinal zones of the ~~UAH~~ RSS merged MSU4 and AMSU9 database with break-points in the residuals relative to the iVRT data set. The residuals from the regression model for the zones from **(a)** ~~85~~ 70 to ~~90~~75° S, **(b)** ~~60~~ 30 to ~~65~~35° S, **(c)** 5 to 10° N, **(d)** 10 to 15° N, and ~~**(c)** 85~~ **(e)** 70 to ~~90~~75° N are shown. The vertical dashed lines indicate where the SNHT detected break-points.

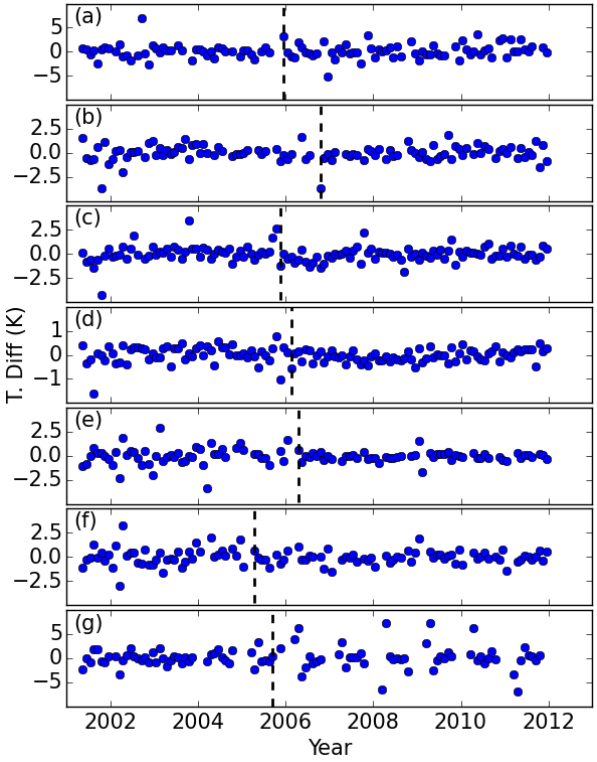

Figure 8. Latitudinal zones of the UAH merged MSU4 and AMSU9 database with break-points in the residuals relative to the iVRT data set. The residuals from the regression model for the zones from **(a)** 85 to 90° S, **(b)** 70 to 75° S, **(c)** 60 to 65° S, **(d)** 30 to 35° S, **(e)** 70 to 75° N, **(f)** 75 to 80° N, and **(g)** 85 to 90° N are shown. The vertical dashed lines indicate where the SNHT detected break-points.