NOAA Climate and Global Change Program
Annual Progress Report for
March 16, 2009 through March 15, 2010

Continued Monitoring of Atmospheric Temperature Using Data
from Microwave Sounding Instruments
NOAA Award No. NA08OAR4310674

Principle Investigator: Carl A. Mears, Remote Sensing Systems
Co-Investigator: Frank J. Wentz, Remote Sensing Systems

Introduction.

The work for this award is focused on the following areas.

1. Continued production of Climate Data Records (CDRs) from available microwave sounders.
2. Investigation and removal of any calibration problems that develop in the dataset.
3. Development of CDRs from new instruments that are suitable for merging with the earlier data.
4. Validation of CDRs using adjusted radiosonde and radio occultation measurements.
5. Development of comprehensive error estimates for each data product.
6. Software engineering to improve the reliability, transparency, and efficiency of our data processing system. The end result would be a processing system suitable for converting to operational status.

We have made progress in many of these areas. Work in each area is discussed in the sections below.

1. Continued production of CDRs from the available microwave sounders.

We update the merged MSU/AMSU datasets for TLT, TMT, TTS, and TLS each month. The data are made available to the research community via our website (www.remss.com/msu). This year, the data will be included in several chapters of the State of the Climate Report, published in the Bulletin of the American Meteorological Society (BAMS).

2. Investigation and removal of any calibration problems that develop in the dataset

We continue to monitor data from all operational AMSU instruments for new calibration problems.

3. Development of CDRs that include data from new instruments.

We have completed calibration studies that allow us to include data from the AMSU instruments in AQUA, NOAA-18, METOP-A, and NOAA-19 in our merged datasets. We have completed merging the results from these satellites into the existing MSU/AMSU data and this new version is now ready for release. The new version (Version 3.3) will be released to the scientific community for the May 2010 update.

Version 3.3 is very similar to version 3.2, which only included data from the NOAA-15 satellite. In Fig. 1, we show globally averaged time series for 1979-2008 for versions 3.2 and 3.3 for each channel. The addition of the additional satellites only makes very small changes in the globally averaged temperature trends. Small changes in the monthly temperatures can be seen after mid 2002, when data from AQUA begins to be included.

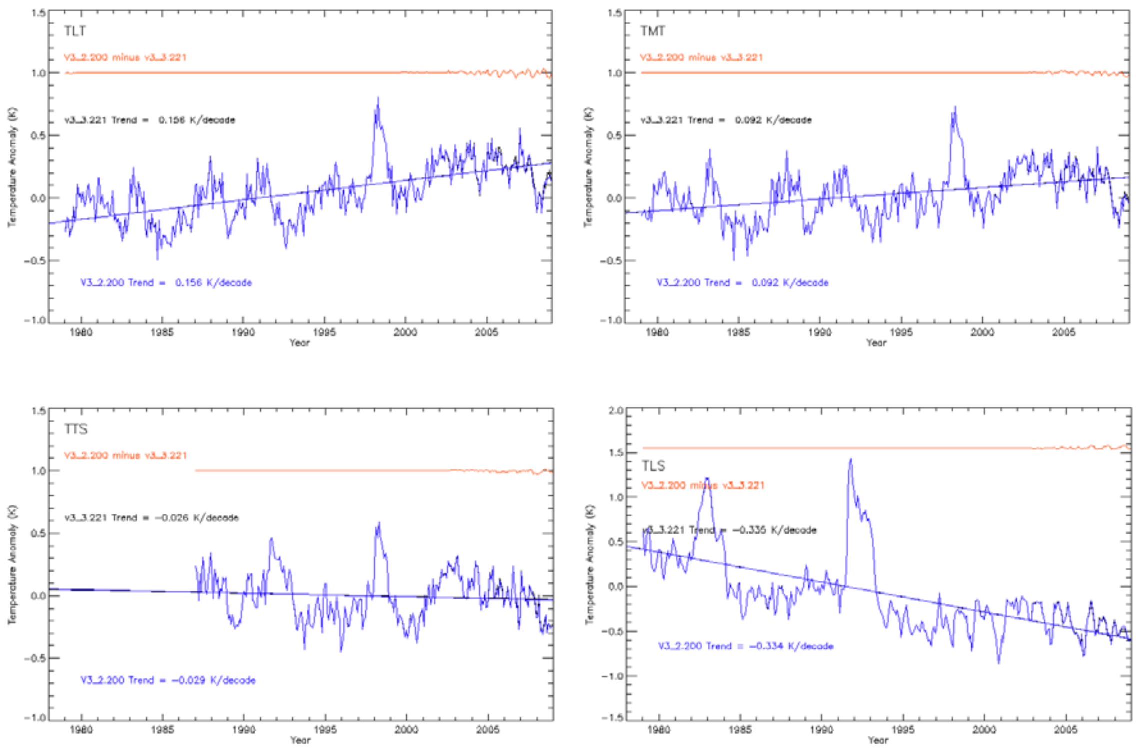

Figure 1. Time series of globally averaged temperature of each CDR. Version 3.2 is the current version that only includes data from NOAA-15. Version 3.3 will be released on May 1, 2010, and includes data from AQUA, NOAA-18, METOP-A and NOAA-19. The addition of the new satellites only makes small changes to the time series, as can be seen in the difference time series shown in red, offset by 1.0.

The fact that only small changes occur when the additional satellites are included is very encouraging. On smaller spatial scales, the differences can be larger. In Figure 2, we show a map of the difference in trends (1979-2008) for TLS between version 3.2 and version 3.3.

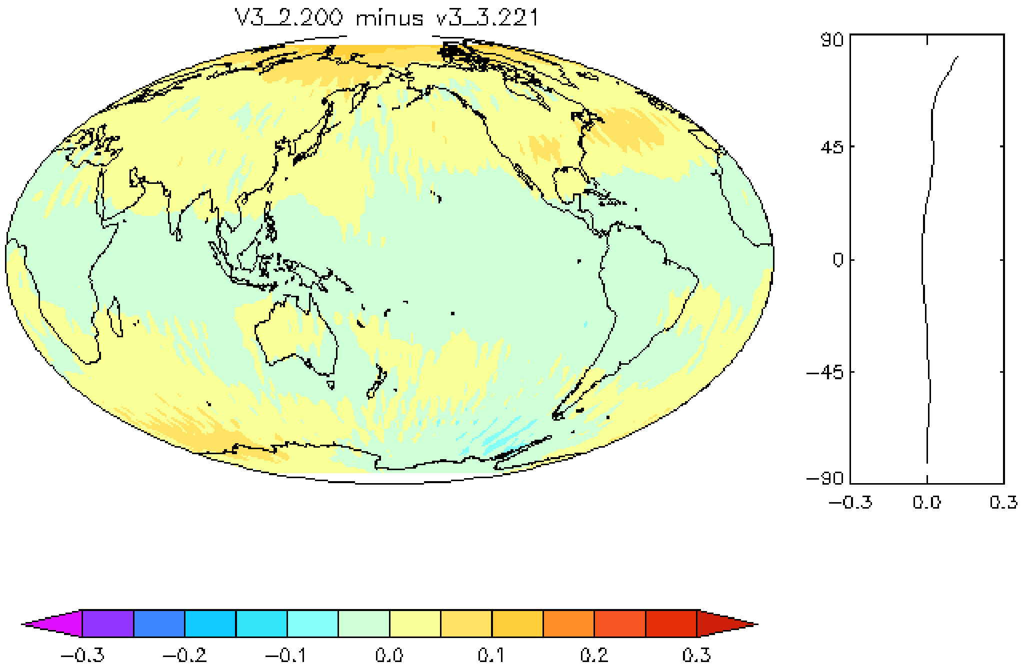

Figure 2. Trend difference V3.2 minus V3.3 for TLS.

4. Validation of CDRs using radiosonde and GPS RO measurements of atmospheric temperature.

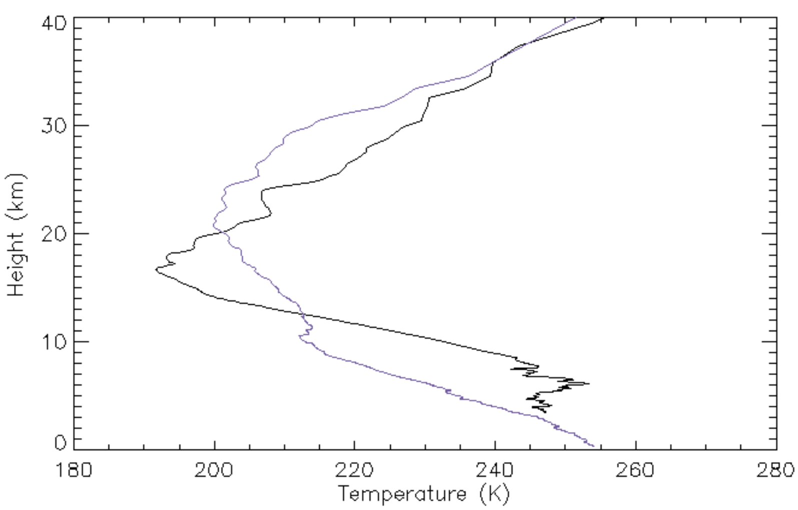

Figure 3. Two typical GPS-RO temperature soundings from the CHAMP instrument. The black curve is a tropical sounding, and the blue curve is from near the north pole.

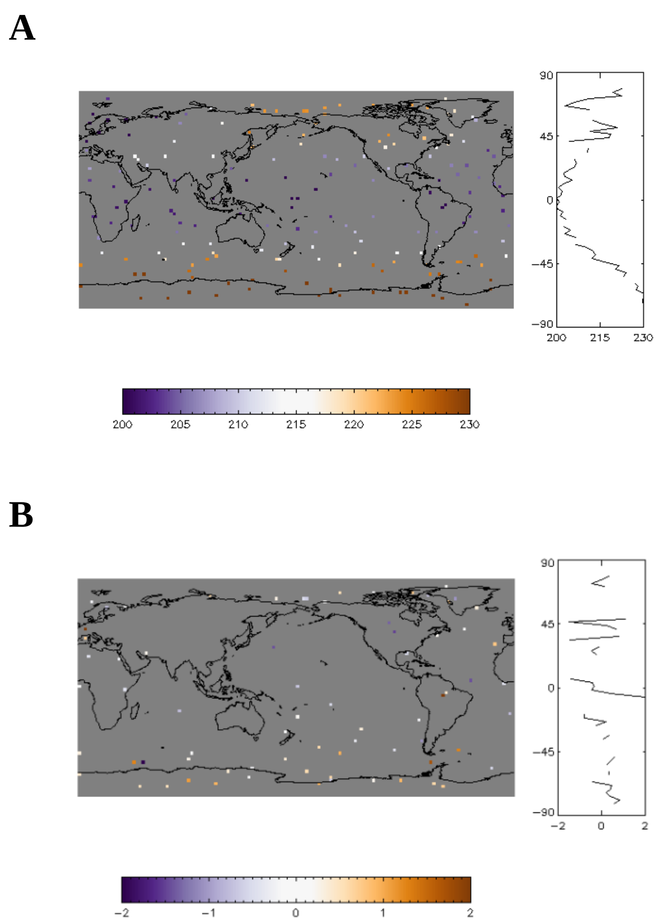

Fig. 4. A) All CHAMP soundings for January 1, 2004 (177 soundings). AMSU 9 equivalent brightness temperatures calculated from each soundings are shown by the color. B) Those soundings that occurred within 3 hours of the NOAA-15 AMSU channel 9 measurement (74 soundings). The difference between the AMSU measurement and the equivalent AMSU 9 brightness temperature calculated from the CHAMP soundings are shown by the color.

During the past year, we have focused our comparison effort on comparing MSU/AMSU data with *in situ* data on GPS RO measurements from the CHAMP (Wickert et al., 2001) and COSMIC (Rocken et al., 2010)programs. Previous intercomparisons have used monthly averaged data(Ho et al., 2007), which contains considerable sampling noise due to the relative sparse sampling for the GPS RO data. We have downloaded all available GPS RO atmospheric temperature soundings from these programs. We then process each sounding to generate MSU and AMSU equivalent brightness temperatures (T_b's) for MSU channel 4 and AMSU channel 9. In Fig. 3, we show two typical soundings from CHAMP, one from the tropics, and one near the poles. In the tropics, data below about 10 km is affected by the presence of water vapor. Near the poles, the valid temperature data extends much lower in the atmosphere. In both regions, the valid data extends low enough to calculate an accurate equivalent temperature for MSU channel 4 or AMSU channel 9.

The CHAMP program only included one receiving satellite, so there are a limited number of soundings per day. In the top panel of Fig. 4A we show all CHAMP soundings for a typical day, and in Fig 4B, we show those soundings the occur within 3 hours of a NOAA-15 measurement for channel 9. After the beginning of the COSMIC mission, the number of collocations increases by about a factor of 10.

In most locations, the CHAMP measurements and the AMSU measurements agree well. This is not always the case for polar latitudes. In Fig. 5, we show a scatter plot of AMSU Channel 9 T_b's as a function of CHAMP equivalent T_b for June 2007. For latitudes north of 50S, the agreement is quite good, with a mean bias of (AMSU channel 9 minus CHAMP) of –0.17K. For latitudes south of 50S, there is a significant bias in the CHAMP Tb's, with a mean bias of 1.43K. Soundings that show this kind of error occur where the temperature profile shows a minimum at a height much higher than normal. In Fig. 6, we show the collocated profiles for June 15, 2007 for the region south of 50S, and the rest of the globe.

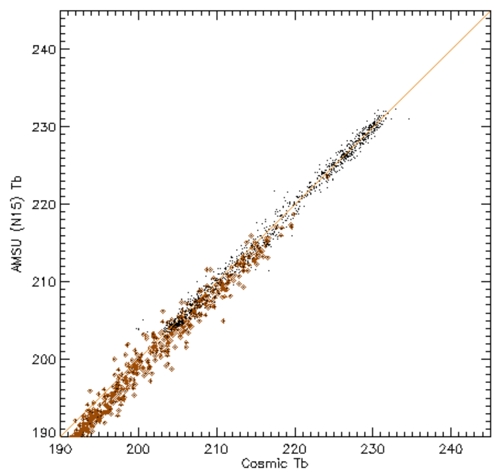

Figure. 5. Scatter plot of AMSU Tb as a function of CHAMP Tb. Soundings south of 50S are shown in brown.

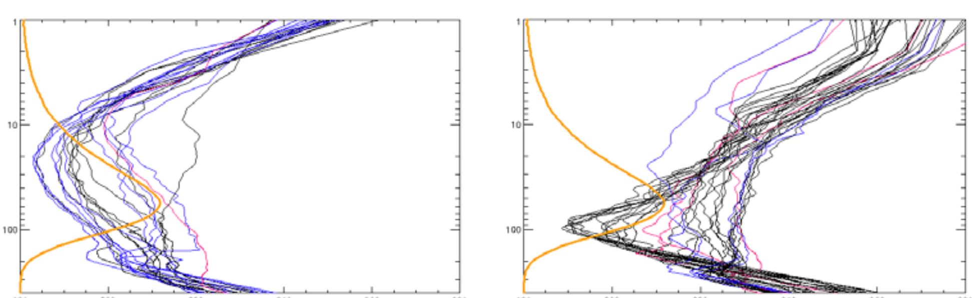

Figure 6. AMSU-collocated CHAMP soundings for June 15, 2007. The left panel shows soundings south of 50S. The right panel shows soundings from the rest of the globe. The AMSU channel 9 weighting function is shown in yellow. Soundings where the CHAMP T_b is too low by at least 1.0K are shown in blue, while those where the CHAMP T_b is to high by 1.0K or more are shown in red.

This error pattern suggests that the temperature weighting function we use to calculate effect temperatures from GPS-RO data peaks too low in the atmosphere. This could be due to insufficient Oxygen absorption in the radiative transfer model (Rosenkranz, 1993) we use to calculate the AMSU-equivalent brightness temperatures from the GPS-RO sounding. We are investigating other oxygen absorption models to try and resolve this problem.

5. Comprehensive uncertainty estimates for each CDR.

We have developed a comprehensive, Monte-Carlo based approach to estimate the uncertainty in our final CDR's. This estimate takes the form of a set of 400 realizations of the estimated uncertainty in a given CDR. Thus associated with each CDR (gridded monthly temperature averages) there are 400 gridded monthly realizations of the estimated error. We are in the final stages of preparing a manuscript that describes this effort. A draft version of the manuscript is attached at the end of this report.

6. Software engineering to improve the reliability, transparency, and efficiency of our data processing system.

We are developing a number of automatically updated web-based tools to make it easier to ensure that the processing system is operating normally and to detect data

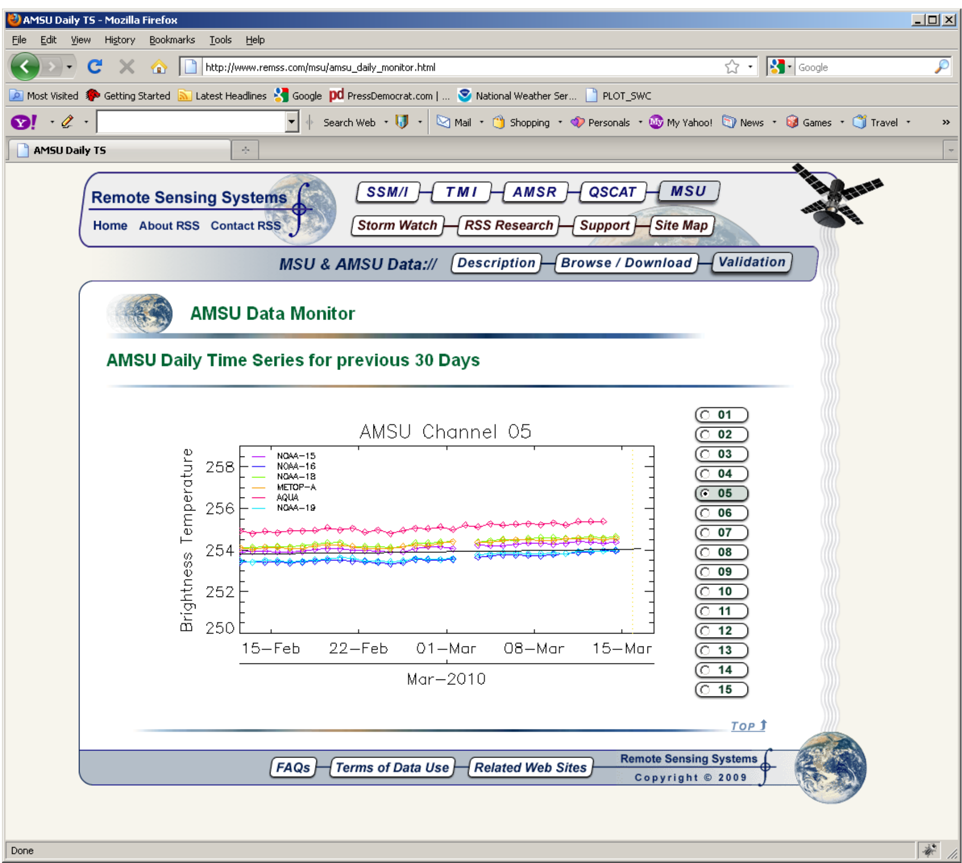

Figure 8. Web-based data monitoring system for AMSU.

anomalies before they have a chance to be included in the merged dataset. We find it useful to plot the daily global average brightness temperature for each AMSU channel in near real time. In Fig. 7, we show a screen shot of the time series page in our web-based AMSU/AQUA monitoring system. The page shows time series of global averages for the previous 30 days, for the channel selected using the radio buttons on the right side of the page. In the case shown (AMSU channel 5), it is easy to see that there are several days of missing data for the NOAA AMSU instruments at the beginning of March. The causes of this gap will be investigated as part of routine monitoring of the AMSU data stream.

7. Planned Work for 3/2010 – 3/2011.

During the next year, we plan work in each of the focus areas for this proposal. An outline of our planned work is presented below.

- Continued production of Climate Data Records (CDRs)
- Complete manuscript on uncertainty estimation.
- Release Version 3.3.
- Finish validation of TLS results vs. radio occultation data.
- Continue to improve automatic monitoring system.

Reference List

Ho, S. P., Y. H. Kuo, Z. Zeng, and T. C. Peterson, 2007: A Comparison of Lower Stratosphere Temperature From Microwave Measurements With CHAMP GPS RO Data. *Geophysical Research Letters*, 34, L15701, doi:10.1029/2007GL030202.

Rocken, C., Y. H. Kuo, W. Schreiner, D. Hunt, and S. Sokolovskiy, 2010: COSMIC System Description. *Terrestrial, Atmospheric and OCeanic Science*, 11, 21-52.

Rosenkranz, P. W., 1993: Absorption of microwaves by atmospheric gases. *Atmospheric Remote Sensing by Microwave Radiometry*, Michael A. Janssen, Ed.,John Wiley and Sons, 37-90.

Wickert, J., C. Reigber, G. Beyerle, R. Konig, C. Marquardt, T. Schmidt, L. Grunwaldt, R. Galas, T. K. Meehan, W. G. Melbourne, and K. Gocke, 2001: Atmosphere Sounding by GPS Radio Coocultation: First Results From CHAMP. *Geophysical Research Letters*, 28, 3263-3266.

Long-term atmospheric temperature records from the MSU and AMSU microwave sounders: uncertainty estimates and comparison to adjusted radiosonde datasets

Carl A. Mears
Frank J. Wentz
Remote Sensing Systems

Peter Thorne
Dan Bernie
U.K. Met Office

Corresponding Author:
Carl A. Mears
Remote Sensing Systems
438 First Street, Suite 200
Santa Rosa, CA 95401

mears@remss.com

Abstract. Measurements made by microwave sounding instruments, including the Microwave Sounding Unit (MSU) and the Advanced Microwave Sounding Unit (AMSU) provide a multi-decadal record of atmospheric temperature change. Calibration methods have been developed to merge data from three sets of MSU/AMSU channels to produce long-term temperature records of thick layers of the atmosphere centered in the lower troposphere, middle troposphere, near the tropopause, and in the lower stratosphere. In this work, we present an internal uncertainty estimate made using a Monte Carlo approach that includes contributions to the total uncertainty from sampling error, pre-merge adjustments, and the merging procedure. The results of this calculation are a set of 400 realizations of the estimated errors in the gridded monthly datasets for each MSU/AMSU channel. These realizations are then used to interpret the results of a comparison between satellite-derived atmospheric temperature, and temperatures from homogenized radiosonde datasets.

1. Introduction

Temperature sounding microwave radiometers flown on polar-orbiting weather satellites provide an important record of upper-atmosphere temperatures beginning with the Microwave Sounding Unit (MSU) on *TIROS-N* in late 1978. In the following years, a series of eight additional MSU instruments provided a continuous record up to the present, with the MSU on *NOAA-14* still in operation. Beginning in 1998, the first of a follow-on series of instruments, the Advanced Microwave Sounding Units (AMSUs) was launched. In order to provide continuous records of atmospheric temperatures, data from the AMSU instruments was merged with data for the previous MSU series of instruments. In two previous papers, we described the methods we used to merge data from the MSU and AMSU satellites together to form the TLT (temperature lower troposphere), TMT (temperature middle troposphere), TTS (temperature troposphere stratosphere), and TLS (temperature lower stratosphere) datasets. The temperature weighting function for each channel is plotted in Fig. 1. In this work, we estimate the uncertainty in these datasets by combining the contributions from various sources of error. These uncertainty estimates are important both when performing comparisons of atmospheric temperature estimates from different measurement sources, and when undertaking scientific studies using these data.

There has been substantial previous discussion of uncertainty in merged microwave sounder datasets (Christy et al., 2000; Christy et al., 2003; Grody et al., 2004; Mears et al., 2003; Mears and Wentz, 2005; Zou et al., 2006). These studies have primarily focused on uncertainty in trends in global scale averages. In this work, we extend these earlier estimates by describing uncertainty on a number of spatial and temporal scales, and more fully accounting for the uncertainty introduced by sampling error and the diurnal adjustment procedures. Since much of the focus of current research is on changes in temperature associated with global climate change, we focus our attention on uncertainty in temperature anomalies, rather than bias in the absolute temperature measurements. These estimates apply to the RSS Version 3.2 MSU/AMSU datasets (Mears and Wentz, 2009a; Mears and Wentz, 2009b), which are available online at www.remss.com. These datasets are monthly averages of temperature, gridded on a 2.5 degree by 2.5-degree grid.

It is also important to note that the trend errors reported for MSU/AMSU data in the major assessment reports (Lanzante et al., 2006; Solomon et al., 2007) are typically generated using information about how well the reported time series fit a linear trend. This type of error provide information about how much the trend might differ if the earth had undergone a different set of short time scale variations, such as those caused by ENSO. They do not provide information about the uncertainty internal to the dataset due to measurement and construction error.

In Section 2, we describe and estimate the magnitude of the various sources of internal uncertainty, and in Section 3, we compare the results from our dataset to results from other approaches used to estimate changes in atmospheric temperature. In Section 4, we discuss the implications of the spatial and temporal structure of the uncertainties for some common methods used to validate satellite data.

2. Estimate of Uncertainty

There are a number of sources of uncertainty in the merged MSU/ASMU temperature datasets. In this section, we describe those sources that we have identified to be important. These are radiometer noise, sampling error, uncertainty in the diurnal adjustment used to account for the effects of drifting local measurement time, and uncertainty in the merging parameters. We note that there may be other unknown sources of uncertainty that we have not yet discovered and thus have not described in this work. The level of consistency between different satellites, and between the satellite data and radiosonde data leads us to conclude that these hypothetical sources of error are likely to be smaller than those that we do describe.

It is difficult to determine the exact structure of the uncertainty in the absence of any reference dataset that is known to be free from error. We therefore provide an informed estimate of the uncertainty, often based on an analysis of intersatellite differences and our knowledge of the merging procedures we used to produce the dataset. In the following sections, we outline the overall approach and the contribution of each source of error to the total error.

Different uncertainty sources are important for different spatial and temporal scales. For example, radiometer noise is important only for short time scales and small spatial scales because its effects are rapidly diminished by averaging multiple observations together. Sampling uncertainty tends to dominate other uncertainty sources for a single monthly average over a 2.5 x 2.5 degree cell (the smallest temporal and spatial resolution we consider), but decreases rapidly when larger temporal or spatial averages are considered. Diurnal and merging parameter uncertainties are spatially and temporally correlated and thus do not decrease rapidly with averaging, and thus become dominant for largest spatial and longest temporal scales.

In this work, we use Monte-Carlo methods to produce a large number of instances of estimated error in the final dataset. These instances are constructed so that they have spatial and temporal correlations that are consistent with those we expect to be present in the final dataset, and thus can be interrogated to produce estimates of the estimated error on a variety of spatial or temporal scales. To generate each instance, we start with a gridded monthly dataset for each satellite that is set to zero for each month that this given satellite has valid data, and is invalid otherwise. We then add to this dataset estimated realizations of the sampling and diurnal adjustment uncertainty. Both these estimates are constructed so that their ensemble averages have zero mean. (Our methods for the construction of these uncertainties are described in the following sections.) The datasets with uncertainty added, which we refer to as "noise datasets", are then analyzed using merging procedures that are identical to those used for the real data. Each noise dataset results in a set of noisy merging parameters (intersatellite offsets and target factors). Since the noise-free dataset is constructed with all zeros, we know that these merging parameters should be zero, and thus any differences from zero are due to the influence the underlying uncertainties, i.e. sampling and diurnal adjustment. Thus, our approach describes the total effect of the underlying uncertainties of the final merged product, including both their direct effect and their indirect effect via uncertainty in the merging parameters.

a. Radiometer Noise

Each measurement made by the MSU or AMSU radiometer has random noise associated with the receiver electronics, which can be characterized by the noise equivalent delta temperature (NEDT). For MSU, the NEDT was specified to be less than 0.3K, and for AMSU, the specified value is 0.25K. For MSU, each monthly 2.5 x 2.5 degree cell contains more than 40 measurements, so the uncertainty associated with radiometer noise is less than 0.04K. For AMSU, there are typically more than 300 measurements, so the corresponding uncertainty for AMSU monthly averages is less than 0.015K. These values are significantly less than the uncertainty from any other source.

b. Sampling Errors

Satellites make measurements at discrete times. When forming a monthly average, all measurements made during a given month that fall within a given 2.5 by 2.5 degree cell were averaged together. Sampling errors occur when this average of discrete measurements does not accurately represent the true monthly average. Fig. 2 shows the simulated hourly TMT brightness temperature from the CCM3 atmospheric model (Kiehl et al., 1996) for a point in the North Pacific (50°N, 170°W). The symbols on the line represent the times at which the temperature would be sampled by over-flights of two co-orbiting satellites. The monthly means obtained by averaging the temperature at these sampling times can be quite different than the true monthly mean obtained by averaging all the hourly temperatures. Note that the mean diurnal cycle is very small for this location, and thus is not an important source of error.

The lower tropospheric dataset (TLT) is formed by calculating a weighted difference of different fields of view (FOVs) (Spencer and Christy, 1992). For the left side of the MSU swath, this difference is given by

$$T_{2LT-MSU} = 2.0\left(T_3 + T_4\right) - 1.5\left(T_1 + T_2\right), \qquad (1)$$

where T_n is the temperature measured by the nth FOV, with T_1 denoting the left-hand, near-limb view. A more complex, but similar differencing procedure is used for AMSU (Mears and Wentz, 2009b). On average, this difference has the effect of extrapolating the measurement lower in the troposphere. However, since each individual view makes a measurement at a different place, an undesirable spatial derivative is also included in the extrapolated value. We consider the effects of this derivative to be part of the sampling error. To estimate this error, we used the hourly CCM3 results to calculate the simulated MSU channel 2 brightness temperature at each FOV for a simulated satellite orbit, and then used the weighted difference of values to deduce a satellite-equivalent TLT that includes the spatial derivative effect. This is the compared to the true average (without spatial derivative or temporal sampling effects) of the modeled TLT-equivalent brightness temperature at the location in question. We find that the spatial derivative effect significantly increases the sampling noise for the TLT dataset. As noted in (Mears and Wentz, 2009b), the extrapolation procedure also results in a location-dependent bias. This bias can be quite large near the poles due to a net north-south spatial derivative in each monthly average. Since the focus of this work is errors in temperature trends and anomalies, we do not evaluate this bias in detail here.

We formulated a model-based estimate of sampling noise for each channel, and for each satellite type (MSU and AMSU) by calculating the differences between the monthly sampled mean and the true monthly mean for 3 different sampling patterns (derived from sampling patterns in the observed data) , for 5 years of hourly model output. Thus, for any given month of the year, there are 15 different possible realizations of the sampling error. In Fig. 3, we show an example of the simulated sampling error for each MSU channel for the month of January and the standard deviation of the MSU sampling error averaged for all months. Sampling errors are largest in the mid latitudes, where the effects of large day-to-day variability and gaps in temporal sampling are combined together. The gaps in temporal sampling are even larger in the tropics, but there is usually much less day-to-day variability so that the magnitude of the sampling error is less. Sampling errors in the tropospheric channels tend to be largest in the winter hemisphere, due to the presence of more intense mid-latitude cyclones. Fig. 3 also shows that the errors at a given grid point tend to be strongly correlated with those in other nearby grid points due to similar temporal sampling patterns.

c. Diurnal adjustment errors.

During the lifetime of most of the MSU and AMSU instruments, the orbit of each instrument's satellite platform slowly changed as a function of time, leading to drifts in local equator crossing time (see Fig. 4). These drifts cause the diurnal cycle to be aliased into the long-term records unless their effects are characterized and removed. As discussed in Mears and Wentz (2008a; 2008b), we used model-based diurnal cycle to make adjustments for drifting local measurement times. We estimate the uncertainty in this diurnal adjustment by evaluating the diurnal adjustments derived from different climate models. These are used to generate plausible realizations of the diurnal adjustment error D, given by

$$D = \sum_{i=1}^{N} a_i \left(D_i - \bar{D} \right).$$

Here, D_i is the adjustment calculated using the *ith* model, $\bar{D}$ is the mean adjustment, and the $a_i\,'s$ are normally distributed random numbers with zero mean and unit variance.

For the data products derived from MSU channel 2 and AMSU channel 5, two models are available for calculating the diurnal adjustment, the NCAR Community Climate Model-3 (CCM3), (Kiehl et al., 1996), and the Hadley Centre Global Environmental Model (HADGEM1), (Martin et al., 2006). In our previous work, the CCM3 model was used to formulate our diurnal correction. For the TTS and TLS data products, which have temperature weighting functions centered higher in the atmosphere, we also considered a diurnal adjustment derived from the Canadian Middle Atmosphere Model (CMAM, Beagley et al., 2000). Our version of the data from this model lacks a surface temperature, making it inappropriate for TLT and TMT, which have a substantial contribution due to surface emission.

The limited number of models available is not ideal. We note that the small number of models makes tenuous our implicit assumption that the differences between the specific models we have available span the range of possible error. Despite this limitation, we proceed with our analysis, since it provides a significant step forward from earlier attempts to characterize the uncertainty in the diurnal adjustment. In Fig. 5, we show a typical realization of the diurnal adjustment and error in the diurnal adjustment for each channel. In Fig. 6, we show global averages of the diurnal adjustments applied to the NOAA-14 instrument. On a global scale, the HADGEM1 model tends to results in larger diurnal adjustments, and a larger seasonal cycle in these adjustments, while the CMAM model tends to result in smaller adjustments

Note that there are large spatial correlations in the diurnal adjustment errors, particularly for the TLT and TMT channels. For these channels, the HADGEM1 diurnal cycle is significantly larger in arid land regions than the CCM3 diurnal cycle. We do not know the reason for this difference, though we suspect it is influenced by differences between the land surface parameterizations used by the two models.

These realizations of the diurnal uncertainty are added to the realizations of the sampling uncertainty we calculated in the previous sections, yielding a set of 400 "noise realizations" that are consistent the estimated combined uncertainty from both sources.

d. Uncertainty in merging parameters

Earlier work found that global averages of simultaneous measurements made by co-orbiting MSU instruments differ by both a time-invariant intersatellite offset and an additional term that is strongly correlated with the variations in temperature of the hot calibration target for each satellite (Christy et al., 2000). To describe these differences, we use an empirical error model for brightness temperature incorporating the target temperature and scene temperature correlation (Mears and Wentz, 2009a),

$$T_{MEAS,i} = T_0 + A_i + \alpha_i T_{TARGET,i} + \beta_i T_{SCENE} + \varepsilon_i \quad (6)$$

where $T_{MEAS,i}$ is the brightness temperature measured by the i-th instrument, T_0 is the true brightness temperature, A_i is the temperature offset for the i-th instrument, α_i is a small multiplicative "target factor" describing the correlation of the measured antenna temperature with the temperature anomalies of the hot calibration target, $T_{TARGET,i}$. The parameter β_i describes the correlation of the calibration error with the scene temperature anomaly T_{SCENE}, and ε_i is an error term that contains additional uncorrelated, zero-mean errors due to instrumental noise and sampling effects. The merging parameters used for the measured datasets were found using a regression procedure that minimizes intersatellite differences between monthly averages. Since these monthly averages contain errors from various sources (dominated by the sampling and diurnal sources mentioned above) there is uncertainty in the calculated merging parameters. This uncertainty leads to additional uncertainty in the final merged dataset. In our analysis we found that the effect of the uncertainty in the β_i's was negligible compared to other error sources and is ignored from this point forward to simplify the analysis.

The α_i's (target factors) are the most important parameters for long term behavior of the merged dataset since they are difficult to determine and they multiply the target temperatures, which often show large long-term changes. Figure 7 shows the standard deviation of the fitted α_i's for each satellite and channel calculated using the 400 noisy realizations from the section above. For TLT and TMT, we find that the largest uncertainty is in the value of $\alpha_{NOAA-09}$, consistent with our earlier findings (Mears et al., 2003). For TLS, the uncertainty in $\alpha_{NOAA-09}$ is less because period during which MSU channel 4 was functioning for both NOAA-09 and NOAA-10 was longer. For TTS, satellites before NOAA-10 are not used because both NOAA-09 and NOAA-06 have large, unexplained drifts in the data for MSU channel 3.

Once the α_i's are determined, we then find the latitude dependent offsets using a regression procedure for each 2.5 degree-wide latitude band. These offsets also vary due to the sampling noise, diurnal adjustment error, and target factors error determined in previous steps. The diurnal adjustment, and thus any error in the diurnal adjustment is typically much larger over land, the errors in these offsets tend to be dominated by errors in the land diurnal cycle. We note that since we use a single offset for each latitude band for both land and ocean regions, any land-caused error can affect the merged dataset in ocean regions at the same latitude.

The MSU and AMSU measurement bands for the corresponding channels for the two instruments differ, leading to small differences in vertical weighting, which in turn lead to small differences in brightness temperature that depend on location and time of year. We remove these differences empirically by using location and time-of-year dependent offsets to adjust the AMSU measurements so that they match the MSU measurements during the 1999-2004 period. These offsets also contain error, which is modeled by our Monte-Carlo process.

e. MSU/AMSU drift for TMT

Examination of the differences between TMT from MSU channel 2 on NOAA-14 and AMSU channel 5 on NOAA-15 shows a long-term trend difference, with NOAA-15 cooling at a rate of 0.2 K per decade relative to NOAA-14 over the July 1998 to December 2004 period of overlap. This trend difference is not present for the other channel pairs, including, to our surprise, TLT. The cause of this trend difference is not known, and could be a drift in calibration in one or both of the satellites. Since we do not know which satellite is closer to being correct, we treat this drift as an additional source of uncertainty.

To estimate the effect of this source of uncertainty, we add in artificial ascending (descending) location independent trends of 0.2K per decade to the MSU (AMSU) noise realizations during the overlap period. Then, we pick 2-year portions of this overlap from a set of four possible periods (1999-2000, 2000-2001, 2001-2002, and 2002-2003) at random. MSU data from after the end of the period, and AMSU data from before the beginning of the period is ignored.

f. Results

At the end of the Monte Carlo merging process, we have 400 realizations of the expected error in our merged MSU/AMSU datasets. In Fig. 8, we show maps of a realization of the gridded error for a typical month, along with maps of the standard deviation of the gridded errors calculated across the 400 realizations. For TLT and to a lesser extent, TMT, the errors are dominated by errors in the diurnal adjustment over land. Errors in TTS are significantly less than for the other channels, due to both the smaller diurnal adjustment, and to the relatively small sampling error.

Figure 9 shows the range of global time series for each channel. The filled region is plus/minus one standard deviation for the globally averaged value for each month, while the two blue lines are the global time series with the largest and smallest overall trend. The red filled region is the plus/minus one standard deviation range without including the uncertainty due to the diurnal adjustment. Table 2 shows the 2-s estimated error in the global and tropical (20S to 20N) trends, with and without including the uncertainty due to the diurnal adjustment. For TMT, the MSU-AMSU drift is also excluded at the same time as the uncertainty in the diurnal adjustments.

3. Comparison with complementary long-term atmospheric temperature datasets.

Our set of gridded estimated error realizations will allow for the first time a rigorous analysis of the uncertainty inherent in a number of dataset intercomparisons and climate model validation activities. These include comparisons of microwave satellite data with in situ datasets, comparisons with other satellite-derived temperature datasets such as those from COSMIC, and comparison with reanalysis and climate model output, including fingerprinting studies. Here we provide an example of this type of analysis by briefly describing the intercomparison of our data with four homogenized radiosonde datasets.

We chose four of the most recent homogenized datasets (HadAT, RAOBCORE, RICH, and IUK) (Thorne et al., 2005), (Haimberger, 2007; Lanzante et al., 2003) (Haimberger et al., 2008) (Sherwood et al., 2008). The datasets were constructed using automated methods to find and estimate the size of "breakpoints" in the time series for radiosonde station which are then used to create adjusted versions of the radiosonde data with the effects of the detected breakpoints removed. A short summary of the features of each dataset is available in (Mears and Wentz, 2009b). These datasets are either available as gridded measurements vertically-weighted to correspond to each channel, or contain enough information so that it is possible for us to construct such a gridded dataset by weighting individual levels. (We do not consider the RATPAC dataset because after 1997, a homogenized version of the dataset is only available as regional averages.)

We have already completed an intercomparison of TLT with these radiosonde datasets (Mears and Wentz, 2009b), and thus this work only serves to add error estimates to this previous work. For TMT, TTS and TLS we follow the earlier analysis method exactly. For each month, the satellite data is sampled at the location of the available radiosonde stations in each dataset. This allows us to make a direct comparison with the radiosonde products, without needing to worry about whether or not the radiosonde sampling is dense enough to faithfully represent a global average. The appearance and disappearance of radiosonde stations over time is also automatically taken into account. Note that this procedure results in a separate sampled satellite dataset for each radiosonde dataset. We then construct area-weighted global and regional time series from each dataset for intercomparison. An example set of time series, for globally averaged TLS RAOBCORE data is shown in Fig. 10. RAOBCORE was chosen for this plot because it agrees better with the satellite based TLS measurements than the other three radiosonde datasets, which all show more stratospheric cooling than the satellite data. In Fig. 10, we show for the satellite data both the true global time series, and the time series calculated using the sampling procedure. The sampling procedure improves the agreement between the radiosonde data and the satellite data on short time scales in almost all cases. Often, agreement in long-term trend is also improved.

Trend error estimates for the radiosonde-sampled RSS satellite data we obtained by constructing error time series by sampling the gridded error realizations at the radiosonde locations, and then calculating the standard deviation of the trends in these time series. In general, the estimated trend error is larger than it is for more spatially complete averages, reflecting the influence of spatial sampling error. This is particularly true for the southern extratropics, where the radiosonde sampling is poor.

Figure 11 shows a summary of trends for each channel, radiosonde dataset and averaging region. We also include sampled trends from the UAH (Christy et al., 2003) and STAR (Zou et al., 2006) satellite datasets, calculated using identical methods. We use the most recent versions of each dataset available in gridded form – versions 5.1 (TMT, TLS) and 5.2 (TLT) for UAH, and 1.3 for STAR. Several general patterns are evident in this figure. First, agreement between radiosonde and satellite datasets is best for TLT. For TLT, both the radiosonde and UAH trends lie within our error bars, except for HADAT in the southern extratropics. All datasets agree the largest warming is in the north, with much less warming in the south, and becomes worse as we move higher in the atmosphere. As we move higher in the atmosphere, the radiosonde-satellite trend differences tend to increase. For TMT, only about 50% of the radiosonde trends lie within our error bars, except in the northern extratropics, where the agreement remains good. For TMT, the STAR trends to be larger than RSS, while the UAH trends tend to be less. For both TLT and TMT, the different satellite datasets agree well in the southern extratropics, where the effects of diurnal drift are small because of the small amount of land. For TTS , the RSS data show less warming than the STAR data, and more than the radiosonde data. Almost all other datasets lie outside the range of our error bars, suggesting that our analysis may be missing some important error sources for this channel. For TLS, the results are similar, except RSS and STAR are in better agreement with most STAR trends lying inside our error bars. The radiosonde data show much more cooling than the RSS satellite data, particularly in the southern extratropics. The increasing discrepancy as we move higher in the atmosphere is likely to be caused at least in part by residual errors in the homogenized radiosonde datasets (Mears et al., 2006; Randel and Wu, 2006).

TLT trends in the tropics have been the subject of much recent controversy (Douglass et al., 2008; Santer et al., 2008). Unfortunately, this is a region where our estimated uncertainty is larger than other regions due to uncertainty in the diurnal adjustment. Our error bars encompass all radiosonde trend estimates as well as trends from the UAH dataset.

It is notable than agreement between all radiosonde and satellite datasets is better in the northern extratropics, where the radiosonde spatial sampling is much more complete. Spatial complete sampling is likely to improve the accuracy of the homogenization methods, which are based on neighbor comparisons.

4. Conclusions

We have performed a comprehensive internal error analysis of our MSU/AMSU based datasets of atmospheric temperature. This work improves upon earlier work in that it provides uncertainty information on a variety of spatial and temporal scales, which is critical for meaningful application of the datasets to the study of climate change on both global and regional scales. The fundamental results of our calculations, a set of 400 realizations of estimated error for each MSU/AMSU channel, is available for download on our website, http://www.remss.com.

Table 1
Long Term Trends using Different Diurnal Models

Channel	Diurnal Model	Global (80S to 80N)	Tropics (30S to30N)
TLT (70S to 80N)	CCM3	0.168	0.147
	HADGEM	0.181	0.195
TMT	CCM3	0.093	0.114
	HADGEM	0.109	0.137
TTS (1987-2008)	CCM3	-0.026	-0.016
	HADGEM	-0.015	-0.004
	CMAM	-0.023	-0.019
TLS	CCM3	-0.335	-0.321
	HADGEM	-0.324	-0.306
	CMAM	-0.348	-0.342

Table 2
Uncertainty estimates (2σ) for trends for each channel (K/decade)

		TLT	TMT	TTS	TLS
Global (75S-75N)	All Errors	0.044	0.042	0.014	0.028
	No Diurnal	0.022	0.012	0.008	0.020
Tropical (20S-20N)	All Errors	0.034	0.038	0.020	0.030
	No Diurnal	0.026	0.012	0.008	0.020

Reference List

Beagley, S. R., C. McLandress, V. I. Fomichev, and W. E. Ward, 2000: The Extended Canadian Middle Atmosphere Model. *Geophysical Research Letters*, **27**, 2529-2532.

Christy, J. R., R. W. Spencer, and W. D. Braswell, 2000: MSU Tropospheric Temperatures: Dataset Construction and Radiosonde Comparisons. *Journal of Atmospheric and Oceanic Technology*, **17**, 1153-1170.

Christy, J. R., R. W. Spencer, W. B. Norris, W. D. Braswell, and D. E. Parker, 2003: Error Estimates of Version 5.0 of MSU-AMSU Bulk Atmospheric Temperatures. *Journal of Atmospheric and Oceanic Technology*, **20**, 613-629.

Douglass, D. H., J. R. Christy, B. D. Pearson, and F. S. Singer, 2008: A Comparison of Tropical Temperature Trends With Model Predictions. *International Journal of Climatology*, **28**, 1693-1701, DOI: 10.1002/joc.1651.

Grody, N. C., K. Y. Vinnikov, M. D. Goldberg, J. T. Sullivan, and J. D. Tarpley, 2004: Calibration of Multi-Satellite Observations for Climatic Studies: Microwave Sounding Unit (MSU). *Journal of Geophysical Research*, **109**, doi:10.1029/2004JD005079.

Haimberger, L., 2007: Homogenization of Radiosonde Temperature Time Series Using Innovation Statistics. *Journal of Climate*, **20**, 1377-1403.

Haimberger, L., C. Tavolato, and S. Sperka, 2008: Towards the Elimination of Warm Bias in Historic Radiosonde Records -- Some New Results From a Comprehensive Intercomparioson of Upper Air Data. *Journal of Climate*, **21**, 4587-4606.

Ho, S. P., Y. H. Kuo, Z. Zeng, and T. C. Peterson, 2007: A Comparison of Lower Stratosphere Temperature From Microwave Measurements With CHAMP GPS RO Data. *Geophysical Research Letters*, **34**, L15701, doi:10.1029/2007GL030202.

Kiehl, J. T., J. J. Hack, G. B. Bonan, B. A. Boville, B. P. Briegleb, D. L. Williamson, and P. J. Rasch, 1996: Description of the NCAR Community Climate Model (CCM3). report TN-420,

Lanzante, J. R., S. Klein, and D. J. Seidel, 2003: Temporal Homogenization of Monthly Radiosonde Temperature Data. Part I: Methodology. *Journal of Climate*, **16**, 224-240.

Lanzante, J. R., T. C. Peterson, F. J. Wentz, and K. Y. Vinnikov, 2006: What do observations indicate about the change in temperatures in the atmosphere and at the surface since the advent of measuring temperatures vertically. *Temperature Trends in the Lower Atmosphere: Steps for Understanding and Reconciling Differences*, Thomas R. Karl, S. J. Hassol, Christopher D. Miller, and W. L. Murray, Eds..

Martin, G. M., M. A. Ringer, V. D. Pope, A. Jones, C. Dearden, and T. J. Hinton, 2006: The Physical Properties of the Atmosphere in the New Hadley Centre Global

Environmental Model, HadGEM1. Part 1: Model Description and Global Climatology. *Journal of Climate*, **19**, 1274-1301.

Mears, C. A., C. E. Forest, R. W. Spencer, R. S. Vose, and R. W. Reynolds, 2006: What is our understanding of the contribution made by observational or methodological uncertainties to the previously reported vertical differences in temperature trends? U.S. Climate Change Science Program and the Subcommittee on Global Change Research Final Report Synthesis and Assessment Product 1.1, Chapter 4, [Available online at http://www.climatescience.gov/Library/sap/sap1-1/finalreport/default.htm.]

Mears, C. A., M. C. Schabel, and F. J. Wentz, 2003: A Reanalysis of the MSU Channel 2 Tropospheric Temperature Record. *Journal of Climate*, **16**, 3650-3664.

Mears, C. A. and F. J. Wentz, 2005: The Effect of Drifting Measurement Time on Satellite-Derived Lower Tropospheric Temperature. *Science*, **309**, 1548-1551.

Mears, C. A. and F. J. Wentz, 2008a: Construction of the Remote Sensing Systems V3.2 atmospheric temperature records from the MSU and AMSU microwave sounders.

Mears, C. A. and F. J. Wentz, 2008b: Contruction of the Remote Sensing Systems V3.2 lower tropospheric temperature records from the MSU and AMSU microwave sounders.

Mears, C. A. and F. J. Wentz, 2009a: Construction of the Remote Sensing Systems V3.2 Atmospheric Temperature Records From the MSU and AMSU Microwave Sounders. *Journal of Atmospheric and Oceanic Technology*, **26**, 1040-1056.

Mears, C. A. and F. J. Wentz, 2009b: Construction of the RSS V3.2 Lower Tropospheric Dataset From the MSU and AMSU Microwave Sounders. *Journal of Atmospheric and Oceanic Technology*, **26**, 1493-1509.

Randel, W. J. and F. Wu, 2006: Biases in Stratospheric and Tropospheric Temperature Trends Derived From Historical Radiosonde Data. *Journal of Climate*, **19**, 2094-2104.

Rocken, C., Y. H. Kuo, W. Schreiner, D. Hunt, and S. Sokolovskiy, 2010: COSMIC System Description. *Terrestrial, Atmospheric and OCeanic Science*, **11**, 21-52.

Rosenkranz, P. W., 1993: Absorption of microwaves by atmospheric gases. *Atmospheric Remote Sensing by Microwave Radiometry*, Michael A. Janssen, Ed.,John Wiley and Sons, 37-90.

Santer, B. D., P. W. Thorne, L. Haimberger, K. E. Taylor, T. M. L. Wigley, J. R. Lanzante, S. Solomon, M. Free, P. J. Gleckler, and P. D. Jones, 2008: Consistency of Modelled and Observed Temperature Trends in the Tropical Troposphere. *International Journal of Climatology*, **28**, 1703-1722, doi:10.1002/joc.1756.

Sherwood, S. C., C. L. Meyer, R. J. Allen, and H. A. Titcher, 10-1-2008: Robust Tropospheric Warming Revealed by Iteratively Homogenized Radiosonde Data. *Journal of Climate*, **21**, 5336-5352.

Solomon, S. D., D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, 2007: Climate Change 2007: The Physical Scientific Basis: Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. 996pp.

Spencer, R. W. and J. R. Christy, 1992: Precision and Radiosonde Validation of Satellite Gridpoint Temperature Anomalies. Part II: A Tropospheric Retrieval and Trends During 1979-1990. *Journal of Climate*, **5**, 858-866.

Thorne, P. W., D. E. Parker, S. F. B. Tett, P. D. Jones, M. McCarthy, H. Coleman, and P. Brohan, 2005: Revisiting Radiosonde Upper-Air Temperatures From 1958 to 2002. *Journal of Geophysical Research*, **110**, D18105, doi:10.1029/2004JD005753.

Wickert, J., C. Reigber, G. Beyerle, R. Konig, C. Marquardt, T. Schmidt, L. Grunwaldt, R. Galas, T. K. Meehan, W. G. Melbourne, and K. Gocke, 2001: Atmosphere Sounding by GPS Radio Coocultation: First Results From CHAMP. *Geophysical Research Letters*, **28**, 3263-3266.

Zou, C. Z., M. D. Goldberg, Z. Cheng, N. C. Grody, J. T. Sullivan, C. Cao, and J. D. Tarpley, 2006: Recalibration of Microwave Sounding Unit for Climate Studies Using Simultaneous Nadir Overpasses. *Journal of Geophysical Research*, **111**, D19114, doi:10.1029/2005JD006798.