NOAA SDS project:

Generating consistent MSU/AMSU/SSU radiance FCDRs and deep-layer atmospheric temperature TCDRs from the MSU/AMSU/SSU temperature-sounding channels

Year 1 Annual Report

June 2009 – June 2010
NOAA SDS Technical Officer: **Dr. Jeff Privette**

Principal Investigator:

Cheng-Zhi Zou
NOAA/NESDIS/Center for Satellite Applications and Research
5200 Auth Road, RM 712
Camp Springs, MD 20746
Phone: 301-763-8042 ext 156, Fax: 301-763-8108, email: cheng-zhi.zou@noaa.gov

Co-Investigators:

Sid Boukabara
NOAA/NESDIS/STAR; Phone: 301-763-8136; email: Sid.Boukabara@noaa.gov

Lidia Cucurull
NOAA/NESDIS/STAR; Phone: 301-763-8000; email: Lidia.Cucurull@noaa.gov

Dick Dee
ECMWF, Phone: (+44) 118 949 9352; email: Dick.Dee@ecmwf.int

Qiang Fu
University of Washington; Phone: 206-685-2070; email: qfu@atmos.washington.edu

Thomas Kleespies
NOAA/NESDIS/STAR; Phone: 301-763-8136; email: Thomas.J.Kleespies@noaa.gov

Carl Mears
Remote Sensing Systems; Phone: 707-545-2904; email: mears@sonic.net

John Woollen
NOAA/NWS/NCEP; Phone: 301-763-8000; email: Jack.Woollen@noaa.gov

Table of Contents

Purpose of this Document 2
Project Objectives 2
Summary of Accomplishments 2
MSU radiance FCDRs and atmospheric temperature TCDRs 2
AMSU-A radiance FCDR 4
Merged MSU and AMSU atmospheric temperature TCDR 5
SSU radiance FCDR and stratospheric temperature TCDR 6
Applications of Recalibrated MSU/AMSU Data 7
Validating Climate Trends from Model Simulations 7
NOAA Workshop on CDRs from Satellite Microwave Radiometers 8
Dataset Availability 9
Publications and Selected Meeting Presentations 9

1. Purpose of this Document

This document is a report of activities that took place during the first year of the NOAA Scientific Data Stewardship (SDS) project "Generating consistent MSU/AMSU/SSU radiance SDRs and deep-layer atmospheric temperature TCDRs from the MSU/AMSU/SSU temperature-sounding channels". This is a three year project. Four different agencies/institutions are involved: the NOAA/NESDIS/Center for Satellite Applications and Research (STAR) in Camp Springs, MD; the University of Washington (UW) in Seattle, Washington; Remote Sensing Systems (RSS) in Santa Rosa, CA; and European Centre for Medium-Range Weather Forecasts (ECMWF) in Reading, United Kingdom. This report summarizes work done by STAR and UW which receive funding from the SDS program. RSS will report directly to the SDS program manager.

2. Objectives

The primary objective of the project is to develop consistent, inter-calibrated MSU/AMSU/SSU radiance fundamental climate data records (FCDRs) and their subsequent deep-layer atmospheric temperature thematic climate data records (TCDRs). This will include MSU-, AMSU-, and SSU-only radiance FCDRs and temperature TCDRs as well as merged MSU/AMSU and AMSU/SSU temperature TCDRs. The project will extend the NESDIS recalibration effort by including more products with extended time period. A second objective of the project is to have the MSU/AMSU/SSU data production teams and climate application community to work closely for an optimal use of the recalibrated dataset.

3. Summary of Accomplishments

The accomplishments are segmented into development and merging of FCDRs and TCDRs from different instruments. Application and comparison of the FCDRs and TCDRs with other data sources are also briefly summarized.

3.1 MSU radiance FCDRs and atmospheric temperature TCDRs

The entire 28-year MSU observations (1979-2006) have been inter-calibrated using simultaneous nadir overpass (SNO) method. An improved 28-year MSU level-1c FCDR was generated from the inter-calibration for reanalysis data assimilation that accounts for solar heating induced instrument temperature variability. Figure 1 shows how the SNO calibration improves the level-1c radiance dataset. As shown, inter-satellite brightness temperature differences of the SNO calibrated are significantly reduced compared to pre-launch calibration.

Version 1.2 of the 28-year MSU atmospheric temperature TCDR was generated using the SNO calibrated FCDRs. The TCDR includes gridded monthly and pentad global deep-layer atmospheric temperatures of mid-troposphere (TMT, MSU channel 2), tropopause and stratosphere (TTS, MSU channel 3), and lower-stratosphere (TLS, MSU channel 4) with spatial resolution of 2.5^0 longitude by 2.5^0 latitude. Merging procedure includes correction of instrument temperature variability, diurnal-drift correction, incident angle adjustment, residual bias correction, and quality control. Figure 2 shows the global-mean pentad temperature

anomalies of the three products for the nine MSU satellites from TIROS-N through NOAA-14. As shown, different satellite agrees with each other on a point-by-point basis, demonstrating high quality of the merged products.

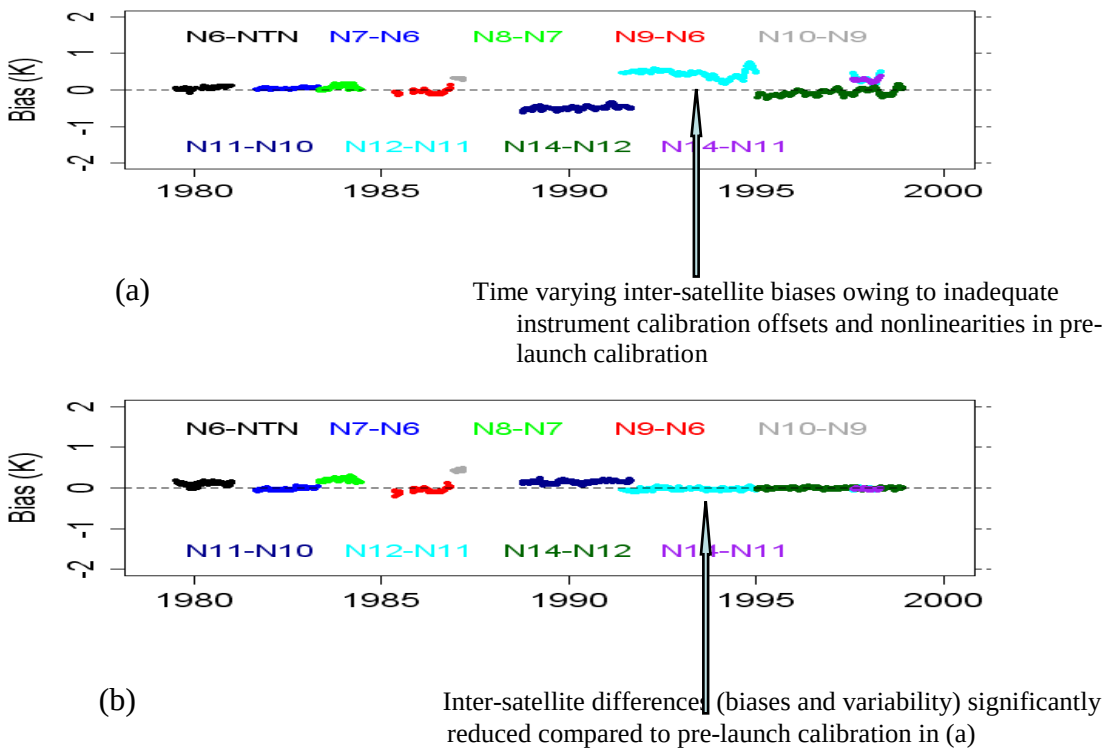

Figure 1 Global ocean-mean inter-satellite brightness temperature difference time series for MSU channel 2 onboard TIROS-N through NOAA-14, which characterizes level-1c radiance biases between satellite pairs. (a) Pre-launch calibration; (b) SNO calibration.

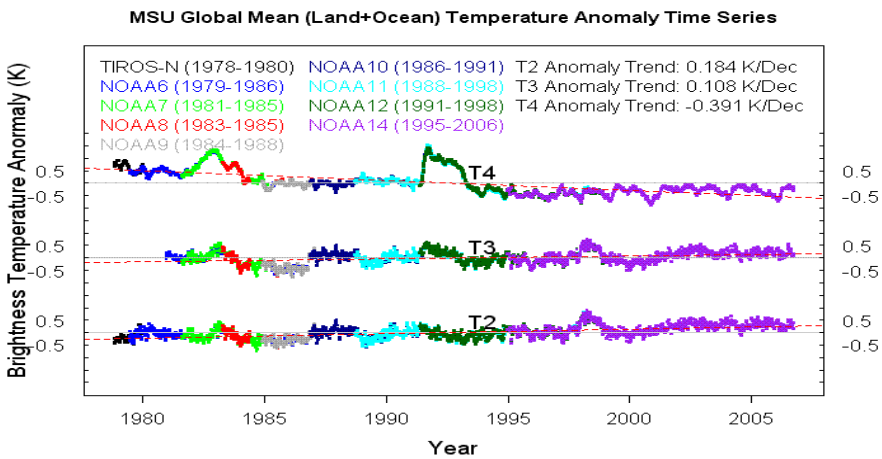

Figure 2 SNO-recalibrated anomaly time series and trends for the 5-day and global mean T_2 (TMT), T_3 (TTS) and T_4 (TLS) from 1979 to 2006. Because of small inter-satellite biases, different satellites agree with each other on a point-by-point basis during overlapping observations, indicating no relative drifts between satellite pairs.

3.2 AMSU-A radiance FCDR

Similar to the MSU instrument, the AMSU-A level-1c radiances onboard NOAA-15 through NOAA-18 and MetOp-A were inter-calibrated using the SNO method. Major achievements in the recalibration included removing of a long-term bias drift in the NOAA-16 observations and solar heating induced instrument temperature variability in the NOAA-15 channel 6 observations. Time varying calibration coefficients were introduced as needed in resolving the bias drift problems. Figure 3 shows AMSU-A channel 6 inter-satellite brightness temperature difference time series before and after the SNO calibration. As shown, the instrument temperature variability and bias drift as seen in the pre-launch calibration were both removed after the SNO calibration.

During the first project year, recalibrations of AMSU-A channels 5 to 10 were completed. Recalibrated level-1c FCDRs for these channels is available through the designated website for the project.

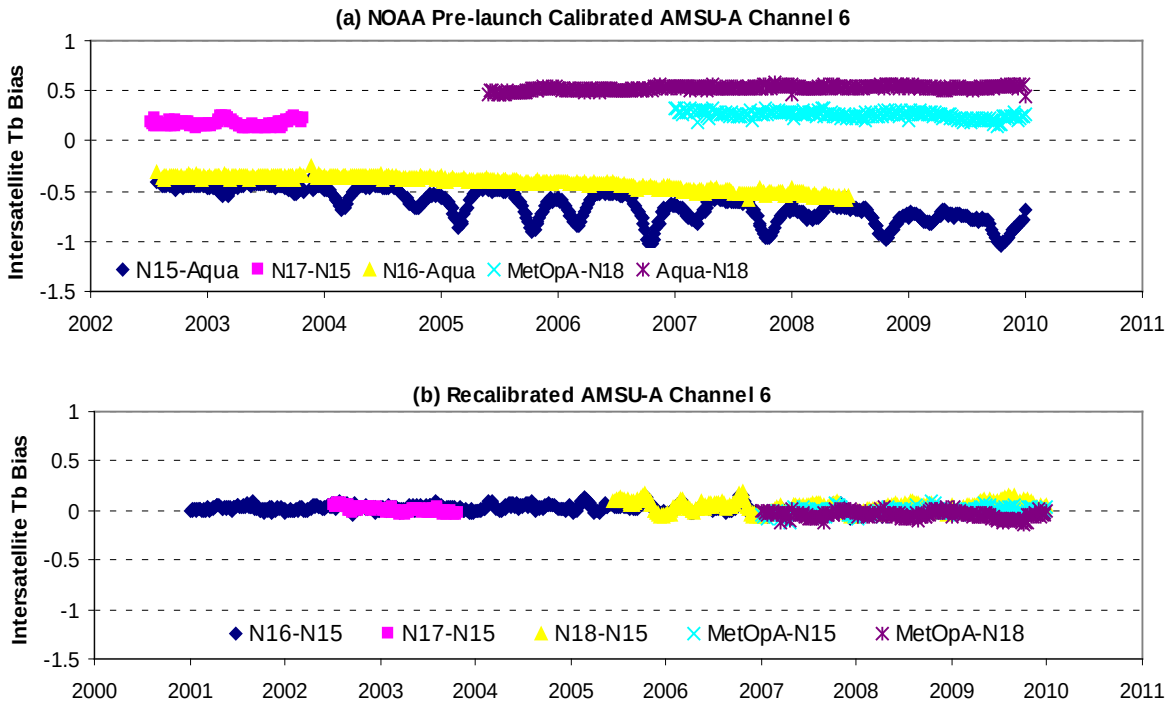

Figure 3 Oceanic-mean inter-satellite brightness temperature difference time series for AMSU-A channel 6 onboard NOAA-15 through NOAA-18 and MetOp-A. (a) Pre-launch NOAA operational calibration; (b) post-launch SNO calibration. Post-launch calibration is not possible with AQUA since its raw counts data are unavailable.

3.3 Merged MSU and AMSU atmospheric temperature TCDR

Version 2.0 of a 31-year (1979-2009) merged MSU and AMSU atmospheric temperature TCDR was generated using the SNO calibrated MSU and AMSU FCDRs. Similar to Version 1.2, the TCDR includes gridded monthly and pentad global TMT (MSU channel 2 and AMSU channel 5), TTS (MSU channel 3 and AMSU channel 7), and TLS (MSU channel 4 and AMSU channel 9) products with grid resolution of 2.5^0 longitude by 2.5^0 latitude. In addition to the bias correction procedure applied in Version 1.2, other bias corrections owing to channel frequency and scanning geometry differences between MSU and AMSU were also investigated and implemented in the Version 2.0 TCDR. Figure 4 shows the temperature anomaly time series of the three products for the 14 NOAA and MetOp-A polar orbiting satellites. Again, different satellites agree with each other on a point-by-point basis, demonstrating high quality of the merged products.

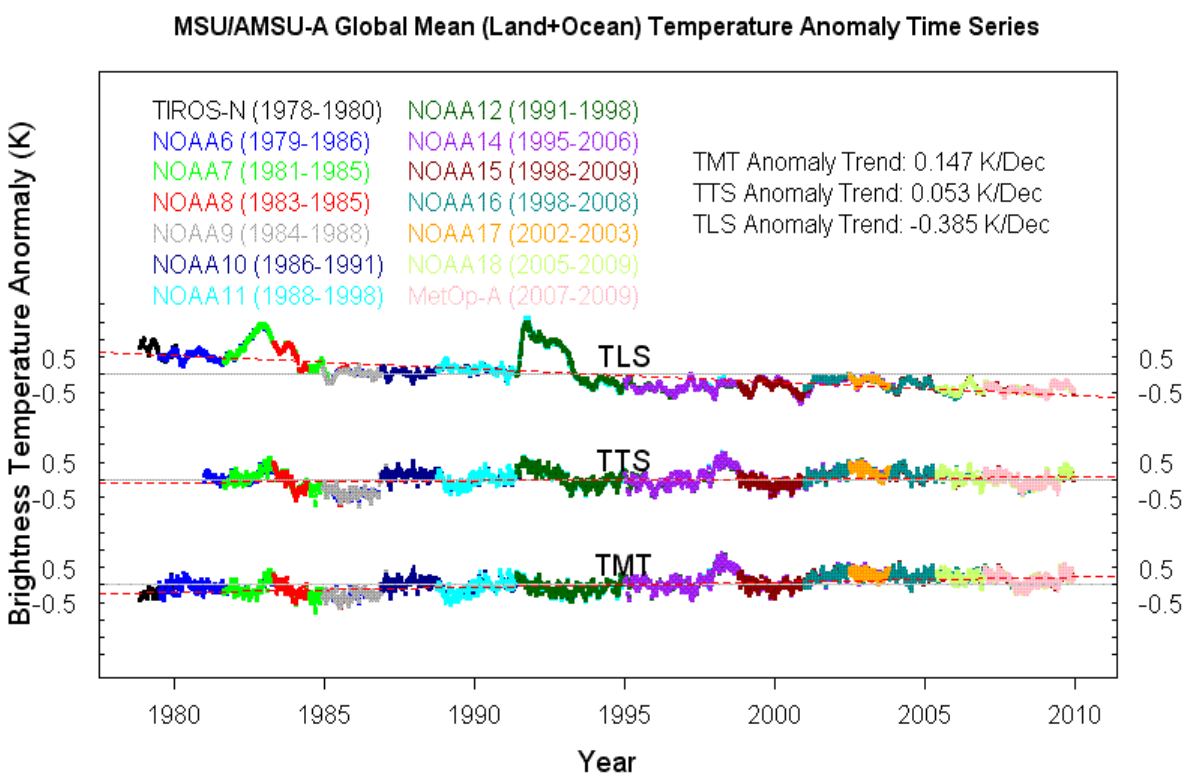

Figure 4 Pentad global mean anomaly time series for Temperature Mid-Troposphere (TMT), Temperature Tropopause and Stratosphere (TTS) and Temperature Lower-Stratosphere (TLS) for different satellites after all calibrations and bias corrections are made.

3.4 SSU radiance FCDR and temperature TCDR

Through extensive research within the project team, key steps for generating homogeneous and consistent SSU FCDR and TCDR have been identified. These include limb correction, instrument CO_2 cell pressure correction, atmospheric CO_2 concentration corrections, diurnal sampling corrections, SNO calibration, and statistical merging. During the first project year, the limb correction and instrument CO_2 cell pressure correction have been completed. Figure 5 shows the SSU time series after these corrections. Two facts in the time series before and after the bias correction clearly shows the improvements from the recalibration 1) inter-satellite biases between satellites have been greatly reduced (e.g. NOAA-11 and NOAA-14); and 2) the drift caused by the leaking instrument CO_2 cell has been corrected (e.g., channel 2 of NOAA-7).

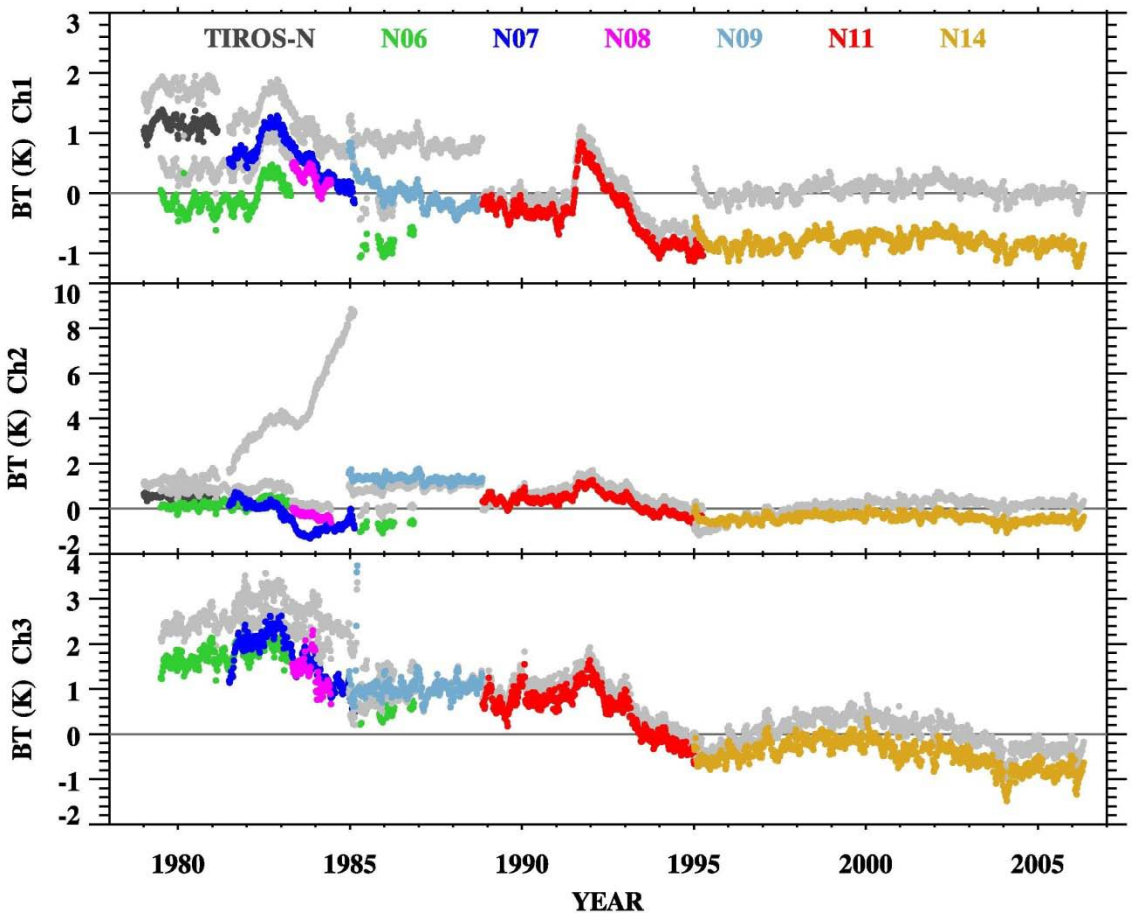

Figure 5 SSU brightness temperature anomalies before (gray) and after (color) the recalibration. Note that each point represents a five-day global mean. The time series clearly show the stratospheric warming after volcano El Chichón and Pinatubo eruption in 1982 and 1991. After recalibration, the cooling trend shows up in Channels 1 and 3 and much clearer than original time series.

3.5 Applications of Recalibrated MSU/AMSU Data

The recalibrated MSU/AMSU data have been used for climate change research and monitoring by investigators from both within and external to NOAA. Major applications of the data in the first project year include i) the atmospheric temperature TCDRs were used to support the NOAA/NCDC annual climate change assessment report "state of the climate in 2008"; ii) 20-year (1987-2006) of the recalibrated MSU radiance FCDR has been assimilated in the new generation of NCEP CFSR and NASA MERRA reanalysis systems to test recalibration impact on reanalysis development. Preliminary analysis of the reanalysis bias corrections shows that the inter-calibrated MSU data have reached their performance expectations in reanalysis data assimilations; iii) Spatial patterns of the atmospheric temperature trend during the past 20 years have been analyzed in details by the project team and the results were published in AMS journals; iv) The recalibrated MSU/AMSU data products have been used to compare with observations from other observing systems such as radiosonde and Global Positioning System (GPS) Radio Occultation (RO).

3.6 Validating Climate Trends from Model Simulations

Tropospheric temperature trends derived from the NOAA MSU/AMSU version 2.0 data as well as similar data products from the Remote Sensing Systems (RSS) and University of Alabama at Huntsville (UAH) have been compared to simulations from 21 climate models for the period of 1979-2009. Troposphere/surface warming ratios from the observed MSU/AMSU datasets and model simulations were also examined. Figure 6 shows T_{24} trends in the MSU/AMSU datasets and GCMs for the globe (Fig.6a-c) and Tropics (Fig.6d-f). The observed tropospheric warming is ~0.1-0.2 K/dec for the global average (Fig. 6a), with the largest values from NOAA and the smallest from UAH. The 0.1 K/dec spread in the global averaged trends is primarily due to differences in warming over the ocean (Fig. 6c) rather than trends over land (Fig. 6b). Over land regions, warming in the MSU/AMSU datasets from different groups agrees well with each other (Fig. 6b and e). However, this agreement is likely spurious since land-only trends reflect the differences in adjustment procedures over ocean as well as diurnal corrections. Over the oceans where observed datasets have better accuracy since the diurnal drift correction is negligible, the multi-model ensemble mean T_{24} trends agree best with the NOAA dataset.

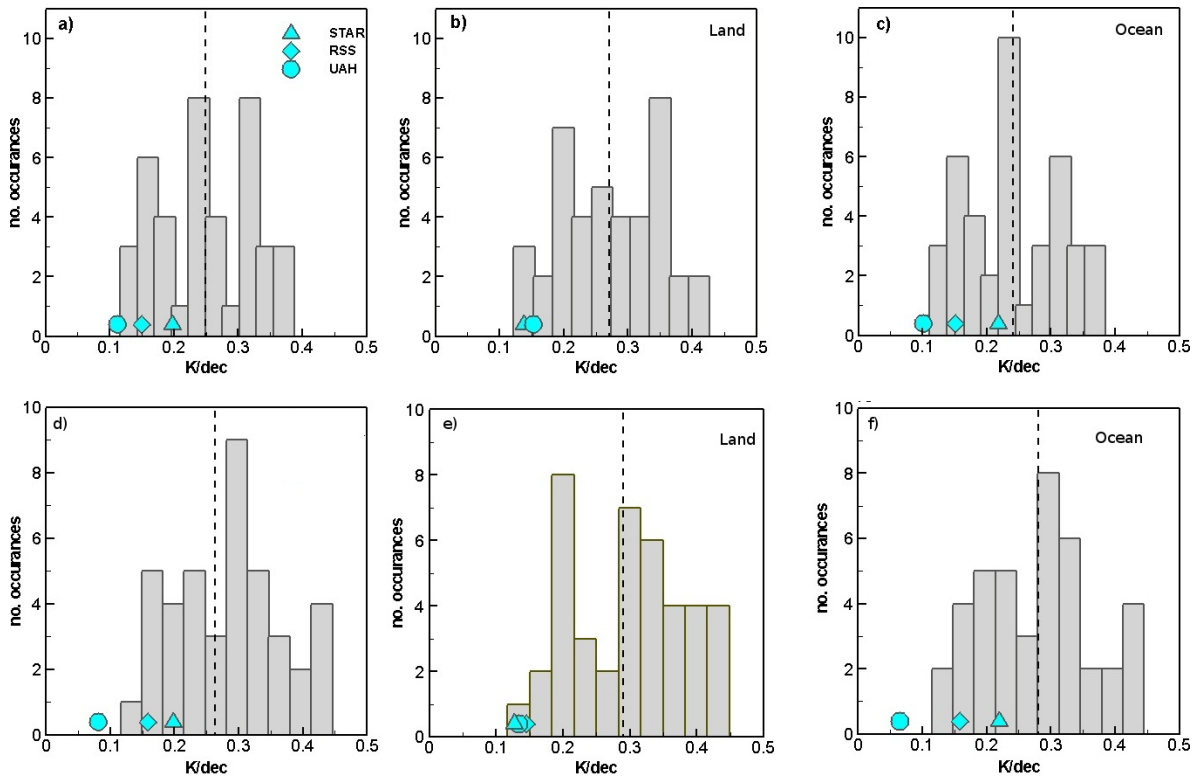

Figure 6. Tropospheric temperature (T_{24}) trends for 1979-2009 from GCM simulations and observational analyses. Histograms represent GCM ensemble member T_{24} trends that fall into a given range with the multi-model ensemble mean given by the dashed line. Observed trends from NOAA (STAR), RSS and UAH are given by the triangles, diamonds, and circles respectively. Trends are given for the global mean (left), land regions only (middle), and ocean only regions (right). The upper and lower panels are respectively for the globe (a-c) and Tropics (30N-30S) (d-f).

3.7 NOAA Workshop on CDR from Satellite Microwave Radiometers

The SDS microwave imagery and sounding project teams hosted a NOAA Workshop on Climate Data Record (CDR) from Satellite Microwave Radiometers at NOAA Science Center, Silver Spring, Maryland from March 22-24, 2010. The workshop covered topics on instrument calibration and CDR development from long-term satellite microwave imagery and sounding observations onboard NOAA, NASA, DOD, and EUMETSAT operational polar-orbiting satellite series. The main purpose of the workshop was for the NOAA SDS CDR development teams to get and respond to input from users and other CDR developers on key concepts and concerns to ensure NOAA CDRs are both highly useful and appropriately up-to-date. The workshop also provides a mechanism for running project transparency program to gain community acceptance and credibility by formally and openly describing the project approaches. More than 50 scientists and program mangers attended the workshop, representing more than 20 national/international weather operational centers, research and academic institutes. Presentations from different institutes reviewed their respective progresses on calibration technique development, CDR accuracy evaluation, CDR format requirement, and many other

issues involved in the CDR development. Issues and recommendations toward developing community consensus CDR products were extensively discussed at the workshop.

At the workshop, the SDS MSU/AMSU/SSU CDR development team described in details the NOAA approaches in recalibrating these instruments using the SNO methods and the release of Version 2.0 MSU/AMSU TCDR products. One highlight in the presentations is that the NOAA TLS product has been validated by NCAR using GPSRO calibrated AMSU products. For a 7.5-year global point-by-point inter-comparison on a 10^0 latitude by 10^0 longitude grid resolution, the NOAA TLS product reaches a precision (standard deviation) as high as 0.5 K. This is an extremely encouraging result that demonstrates the good quality of the SNO calibrated MSU/AMSU products.

4. Dataset Availability

All the SNO inter-calibrated MSU/AMSU data can be freely acquired through the designated website for the project with the URL address: http://www.orbit.nesdis.noaa.gov/smcd/emb/mscat/mscatmain.htm. The website contains both the SNO inter-calibrated radiance FCDRs and gridded deep-layer temperature TCDRs for Versions 1.2 and 2.0. Software for reading the dataset is also provided on the website.

5. Publications and Selected Meeting Presentations

Zou, C.-Z., M. Gao, M.Goldberg, 2009, Error structure and atmospheric temperature trends in observations from the microwave sounding unit, *J. Climate*, **22**, 1661-1681, DOI: 10.1175/2008JCLI2233.1

Zou, C.-Z. and W. Wang, 2009, Stability of the MSU-derived atmospheric temperature trend, *J. Atmos. Oceanic. Technol.*, In press

Zou, C.-Z. and W. Wang, 2009, Diurnal drift correction in the NESDIS/STAR MSU/AMSU atmospheric temperature climate data record, Proc. SPIE, Vol. 7456, 745616

Zou, C.-Z., 2010: The NOAA MSU/AMSU/SSU CDR project: Team, Methods, Current Status, and Future plan, *NOAA Workshop on Climate Data Records (CDR) from Satellite Microwave Radiometers*, Silver Spring, Maryland from March 22-24, 2010.

Han, Y., et al., 2010: CRTM's SSU Module that Accounts for Spectral Response Function Variations, *NOAA Workshop on Climate Data Records (CDR) from Satellite Microwave Radiometers*, Silver Spring, Maryland from March 22-24, 2010.

Wang, L., et al., 2010: NOAA's SSU CDR Development, *NOAA Workshop on Climate Data Records (CDR) from Satellite Microwave Radiometers*, Silver Spring, Maryland from March 22-24, 2010.

Zou, C.-Z., et al. 2010: Bias correction of AMSU-A channels in ERA-Interim and its comparison to STAR recalibrated time series, *NOAA Workshop on Climate Data Records (CDR) from Satellite Microwave Radiometers*, Silver Spring, March 22-24, 2010

Wang. W., and C.-Z. Zou, 30-Year Polar Regions Mid-Tropospheric Temperature Trends Using MSU/AMSU-A Observations, *AGU Fall Meeting, San Francisco, CA, December 14-18, 2009*

Johanson, C., and Q. Fu, Tropical warming in MSU/AMSU vs. GCMs: land and ocean differences, *NOAA Workshop on Climate Data Records (CDR) from Satellite Microwave Radiometers*, Silver Spring, Maryland from March 22-24, 2010.

NOAA Climate and Global Change Program
Semi-annual Progress Report for
June 1, 2009 through November 30, 2009

Generating consistent radiance SDRs and deep-layer atmospheric temperature TCDRs from the MSU/AMSU/SSU temperature-sounding channels

NOAA Award No. NA09NES4400015

Science Team Lead: *Cheng-Zhi Zou*
Center for Satellite Applications and Research (StAR)
NOAA/NESDIS
Camp Springs, MD 20746

Science Team Members: *Sid Boukabara, NESDIS/StAR*
Lidia Cucurull, NESDIS/StAR
Dick Dee, ECMWF
Qiang Fu, University of Washington
Thomas Kleespies, NESDIS/StAR
Carl Mears, Remote Sensing Systems
John Woollen, NWS/NCEP

Collaborators: *Ralph Ferraro/NESDIS/StAR*
Ben Ho/UCAR

Introduction

This project is a complex, multi-institutional effort aimed at generating climate quality temperature datasets from satellite-borne atmospheric sounders. This report only covers the part of this work assigned to Remote Sensing Systems, which is to perform comparisons of various satellite products with each other, and with other sources of atmospheric temperature measurements. The purpose of this part of the work is to use such comparisons to help determine the reasons for any discrepancies between satellite datasets. Although the official period of performance for this report is June 1, 2009 – November 30, 2009, funding for this project did not arrive at Remote Sensing Systems until August 1, 2009. Thus, this report only covers 4 months of work.

1. Assembly of datasets for comparison with satellite data.

There are two sources of atmospheric temperature data that are appropriate for comparison with data from satellite sounders. The first is a form of in-situ measurements made by balloon-borne radiosondes. In general, the radiosonde data record is plagued by inhomogeneities caused by (sometimes undocumented) changes in instrumentation recording practices and station location. A number of research groups have undertaken projects to identify and remove the effects of these inhomogeneities as best they can using a variety of methods. We use results from these projects for intercomparison

studies. Fortunately, these data already exist in easy to use formats at Remote Sensing Systems and are updated as appropriate for other purposes. A list of the data sets we use is shown in Table 1.

Radiosonde Dataset	**Source Institution**
HADAT (Thorne et al., 2005)	U. K. Met Office
RATPAC (Free et al., 2005)	NOAA Air Resources Laboratory
RAOBCORE (Haimberger, 2007)	University of Vienna
RICH (Haimberger et al., 2008)	University of Vienna
IUK (Sherwood et al., 2008)	Yale University

Many of the problems with radiosondes are likely to be larger in the stratosphere, where corrections for solar radiation are much larger than in the troposphere. For this reason, we plan to supplement the radiosonde data with radio occultation measurements (Kursinski et al., 1997) made using the CHAMP and COSMIC satellites for comparisons involving the stratospheric channels (MSU channel 4, and AMSU channels 8-14).

We have downloaded the entire CHAMP and COSMIC datasets from the CDAAC website to Remote Sensing Systems (a considerable effort due to the large data volume involved). GPS-RO data is in the form of high vertical resolution temperature profiles at specific locations and times. In order to compare the GPS-RO data with data from satellite sounders, each profile must be converted to an effective brightness temperature that would be observed by the satellite. We have developed code to perform

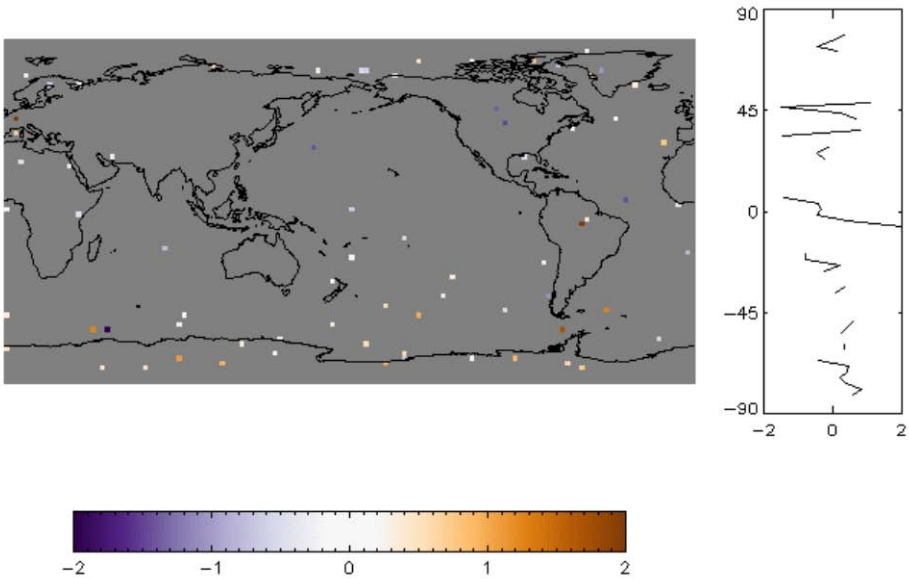

Fig. 1. Typical daily CHAMP data is collocated with AMSU data from NOAA-15. These data are from January 1, 2004, and there are 77 collocations. Each pixel is color coded with the difference between the AMSU channel 9 observation and the AMSU channel 9 equivalent temperature calculated from the CHAMP profile.

this conversion using a radiative transfer model developed by Rosenkranz (1993).

An additional program is used to collocate the GPS-RO observations with daily maps of satellite data. The sounder and GPS-RO measurements were required to be within 3 hours of each other. The use of daily data, rather than monthly averages, significantly reduces the sampling noise and will make it possible to make useful inferences even when using the relatively sparse CHAMP dataset. In the example shown below, there are 77 collocations with a standard deviation of ~0.9K. If these errors were random, then the standard deviation of the mean difference would be about 0.1K, good enough to check calibration accuracy.

2. Preliminary intercomparisons of satellite data.

We have begun preliminary comparisons of satellite data from different research groups with homogenized radiosonde data. We have found that plotting time series of the difference between the various satellite datasets (RSS, UAH, and STAR) and the mean of the radiosonde datasets is especially useful, since differences from the radiosonde record and differences between the satellite datasets are easy to see. Our mean radiosonde dataset includes global averages of the HADAT, RICH, IUK, RAOBCORE 1.4, and RATPAC-B datasets.

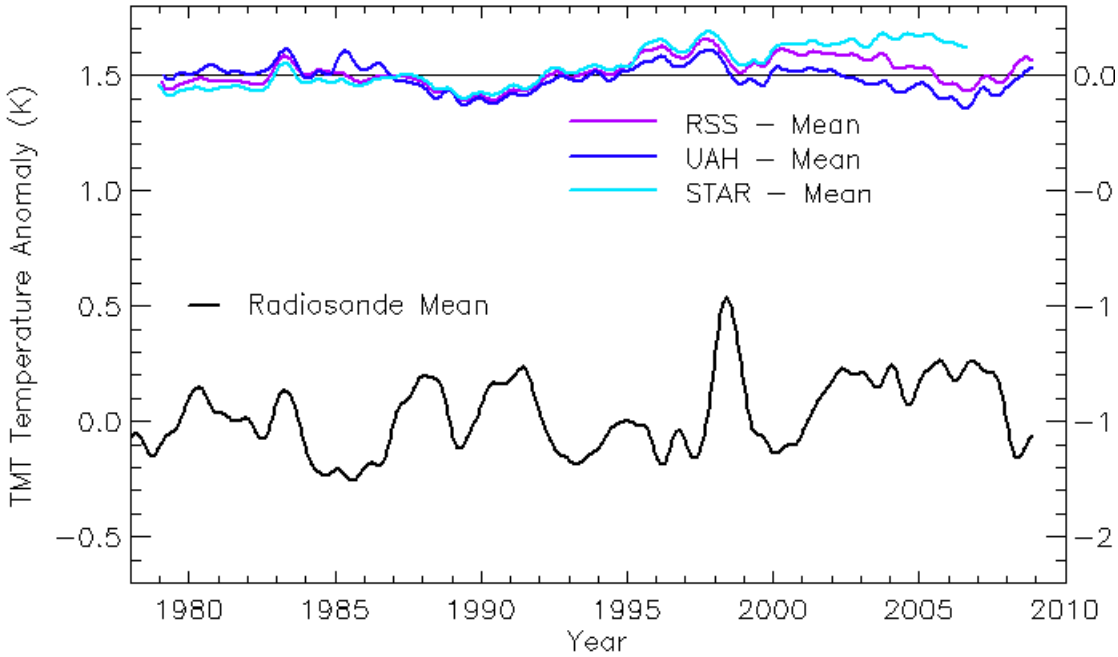

Fig. 2. Globally averaged difference time series between the satellite datasets and the mean radiosonde datasets. The data are smoothed to remove variability on time scales shorter than 1 year.

In Fig. 2, it is apparent that the 3 satellite datasets are more similar to each other than they are to the mean radiosonde (or any single radiosonde dataset – not shown). In particular, the STAR and RSS data track each other closely before 2002. After 2002, the STAR data trend higher than the other datasets. This is likely to be because the STAR data do not yet include AMSU data. We have found that the MSU Channel 2 data trend towards warmer temperatures than AMSU channel 5 data during this period. It is interesting that the STAR data shows a smaller trend relative to the radiosonde data than the other two satellite datasets during this period.

In Fig. 3, we show the difference (STAR – RSS) time series for two versions of the RSS data – one that includes the AMSU data, and one that does not. The agreement between the STAR data and the RSS data is better after 2002 for the version that does not include the AMSU data.

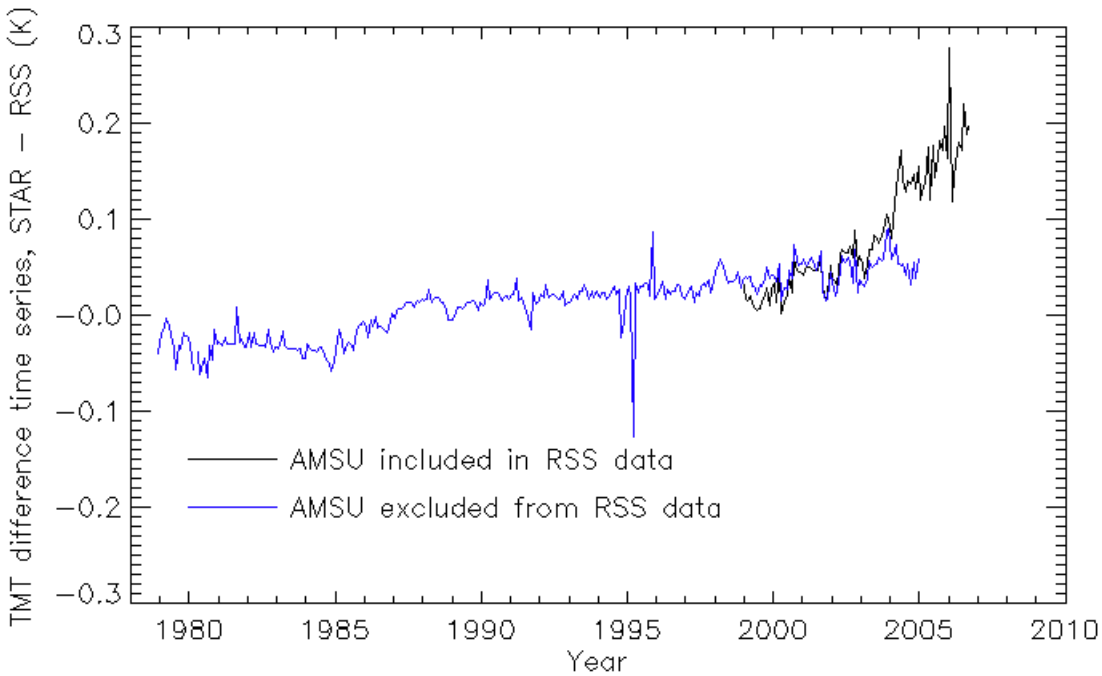

Fig. 3. TMT global average difference time series (STAR – RSS). The agreement between the RSS and STAR dataset is better when the AMSU data is excluded from the RSS data. Also note the ramp in the difference from 1985-1987. This is during the NOAA-09 overlap period, when the calibration adjustments are poorly defined.

3. Work Planned for the next six months

A. Continued comparison of the GPS-RO data with AMSU channel 9 and MSU channel 4. Investigate the possibility of calibration drifts in these channels
B. Investigate the use of GPS-RO data for validation of MSU channel 3 and AMSU channel 7. Since these channels are centered lower in the atmosphere, the GPS-RO data is more likely to be corrupted by the presence of water vapor, especially in the tropics.

C. Continued analysis of the differences between the various MSU/AMSU datasets and the homogenized radiosonde datasets.
D. Participation in workshop on MSU / AMSU / SSU Climate Data Records, at the NOAA Science Center, Camp Springs MD.

References

Free, M., D. J. Seidel, J. K. Angell, J. R. Lanzante, I. Durre and T. C. Peterson, Radiosonde Atmospheric Temperature Products for Assessing Climate (RATPAC): A new dataset of large-area anomaly time series, *Journal of Geophysical Research, 110*, D22101, 2005.

Haimberger, L., Homogenization of radiosonde temperature time series using innovation statistics, *Journal of Climate, 20*(7), 1377-1403, 2007.

Haimberger, L., C. Tavolato and S. Sperka, Towards the elimination of warm bias in historic radiosonde records -- some new results from a comprehensive intercomparison of upper air data, *Journal of Climate, 21*, 4587-4606, 2008.

Kursinski, E. R., G. A. Hajj, J. T. Schofield, R. P. Linfield and K. R. Hardy, Observing Earth's atmosphere with radio occultation measurements using the Global Positioning System, *Journal of Geophysical Research, 102*(D19), 23429-23466, 1997.

Rosenkranz, P. W., Absorption of microwaves by atmospheric gases, in Atmospheric Remote Sensing by Microwave Radiometry, edited by M. A. Janssen, pp. 37-90, John Wiley and Sons, New York, 1993.

Sherwood, S. C., C. L. Meyer, R. J. Allen and H. A. Titcher, Robust tropospheric warming revealed by iteratively homogenized radiosonde data, *Journal of Climate, 21*(20), 5336-5352, 2008.

Thorne, P. W., D. E. Parker, S. F. B. Tett, P. D. Jones, M. McCarthy, H. Coleman and P. Brohan, Revisiting radiosonde upper-air temperatures from 1958 to 2002, *Journal of Geophysical Research, 110*, 2005.

NOAA Climate and Global Change Program
Semi-annual Progress Report for
December 1, 2009 through May 31, 2010

Generating consistent radiance SDRs and deep-layer atmospheric temperature TCDRs from the MSU/AMSU/SSU temperature-sounding channels

NOAA Award No. NA09NES4400015

Science Team Lead: *Cheng-Zhi Zou*
Center for Satellite Applications and Research (StAR)
NOAA/NESDIS
Camp Springs, MD 20746

Science Team Members: *Sid Boukabara, NESDIS/StAR*
Lidia Cucurull, NESDIS/StAR
Dick Dee, ECMWF
Qiang Fu, University of Washington
Thomas Kleespies, NESDIS/StAR
Carl Mears, Remote Sensing Systems
John Woollen, NWS/NCEP

Collaborators: *Ralph Ferraro/NESDIS/StAR*
Ben Ho/UCAR

Introduction

This project is a complex, multi-institutional effort aimed at generating climate quality temperature datasets from satellite-borne atmospheric sounders. This report only covers the part of this work assigned to Remote Sensing Systems, which is to perform comparisons of various satellite products with each other, and with other sources of atmospheric temperature measurements. The purpose of this part of the work is to use such comparisons to help determine the reasons for any discrepancies between satellite datasets. This report covers 6 months of work, from December 1, 2009 to May 31, 2010.

Introduction.

One of the most important goals of this part of the work is to accurately describe the causes for the differences between the RSS and STAR results for each channel. Our initial focus is on the middle troposphere channels, MSU channel 2 and AMSU channel 5. It is easy to assume that all differences are due to what appears to be the main difference between the two datasets, i.e. that RSS used the UAH-developed empirical calibration method (Mears and Wentz 2009; Mears et al. 2003; Christy et al. 2003), while STAR uses the simultaneous nadir overpass (SNO) method to adjust the root level calibration equations (Zou et al. 2006, 2009). There are a number of other important differences in the methods used by the two groups to produce the final results, and it is

important to characterize the importance of these differences on the final result. These differences include different quality control methods, which result in different sets of months being valid for each satellite, and different methods for determining the intersatellite offsets to be removed, and different assumptions about the effective temperature of the cold space view. In the work performed over the past six months, we have focused on performing a number of experiments to help understand the effects of various calibration and merging methods on the final results. The experiments completed so far are:

1. Simple Case Comparisons.
2. Processing STAR L1C brightness temperatures using RSS merging.
3. Adjusting RSS merging procedures to more closely match those used by STAR.

Experiment #1: Simple Case Comparisons. In this experiment, we focus on the part of the MSU dataset that is most free of uncertainty. Uncertainty in MSU channel 2 largely comes from two sources. First, the poorly defined calibration of NOAA-09. NOAA-09 only overlaps with the satellites that precede it and follow it for short periods. This leads to significant uncertainty in the calibration coefficients (target factors for the RSS/UAH style merge, or recalibrated Level 1 calibration coefficients for the SNO calibration) for this satellite. The second major source of uncertainty is uncertainty in the diurnal cycle used to adjust for drifts in local measurement time. To minimize the effects of both of these problems, we exclude data from satellites before NOAA-10, and only uses ocean data where the diurnal cycle is minimal. This study was coordinated with STAR personnel, so that both RSS and STAR produced a version of the TMT data that only used data from NOAA-10 through NOAA-14, and only considered measurements made over the ocean. In Figure 1, we plot the results of this study, along with an ocean-

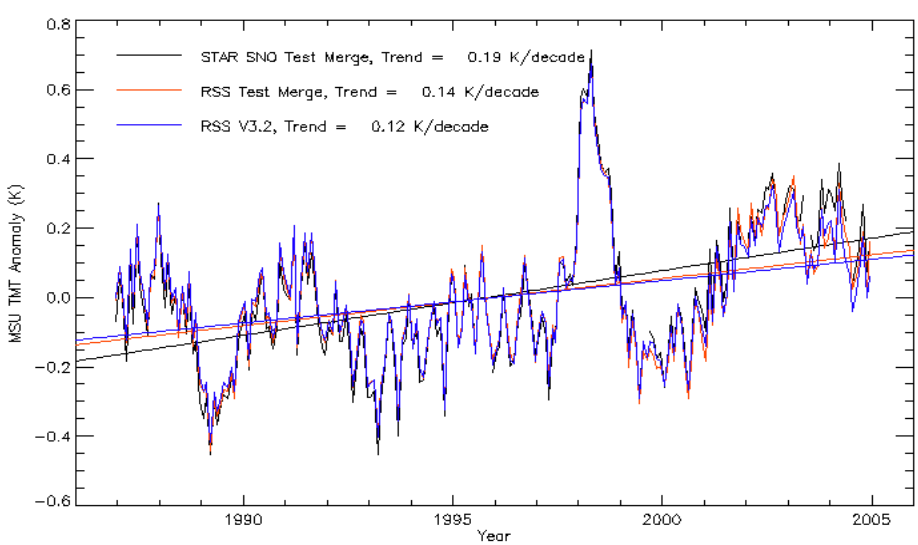

Fig. 1. TMT time series from the ocean-only test cases discussed in Experiment #1.

only time series derived from the standard RSS V3.2 product over the same period. As is the case for the standard STAR product (V2.0), the test case show a significantly larger trend (0.19 K/decade) than either the RSS test case, or the standard RSS product. The use of the ocean-only merge does increase the RSS trend slightly. This is likely an effect of removing the effects of the land diurnal cycle (and possible errors from the adjustment of it) from the RSS merge. The conclusion of this study is that most of the differences in global-scale trends between the RSS and STAR results are due to difference in calibration, i.e. SNO vs. empirical target factor, as opposed to differences in the merging methods applied after the non-linear calibration step was performed.

Experiment #2: Processing STAR L1C brightness temperatures using RSS merging algorithms. In this experiment, we use RSS merging algorithms to process the L1C brightness temperatures produced by the STAR group. The STAR group produces two sets of LIC data, the "NESDIS Operational" calibrated version, and the SNO calibrated version. The NESDIS operational calibration uses the non-linear calibration coefficients determined by NOAA from pre-launch calibration procedures, while the SNO calibration uses an updated set of calibration coefficients produced by analyzing the simultaneous nadir overpass (SNO) matchup dataset produced by STAR. We use both sets of data in this experiment. We downloaded the MSU L1C swath data from the STAR ftp site, which was then averaged into mean monthly maps of brightness temperature using exactly the same code (except for read routines) as we use to produce the RSS version of the data.

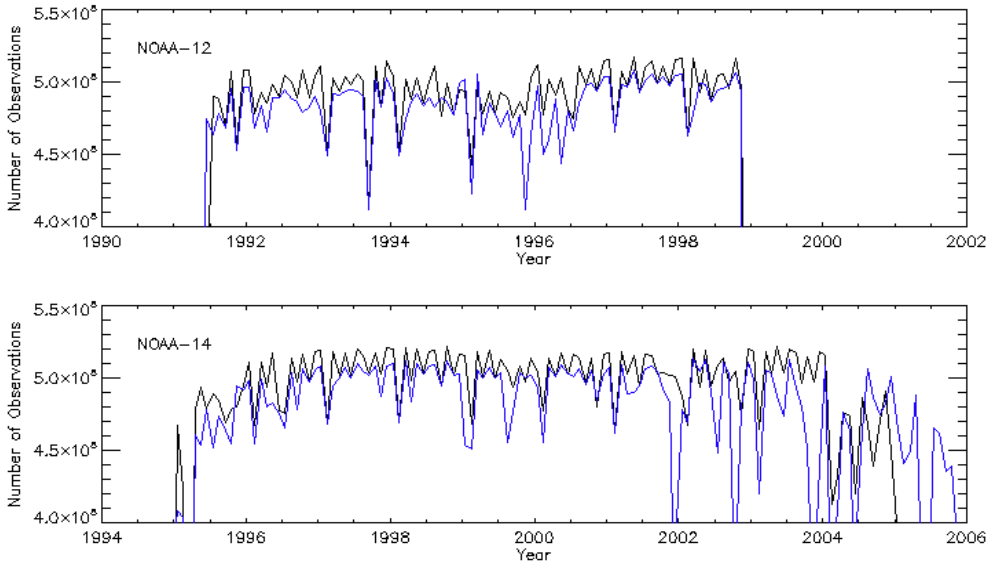

Fig. 2 Number of observations in each globalmonthly average for the RSS (black) and STAR (blue) datasets. In general, the RSS dataset contains a few percent more observations. Results for NOAA-12 and NOAA-14 are shown. Results for other satellites are similar.

The first tests we performed were aimed at determining the degree to which the RSS and STAR orbit data archives contain the same data. In Fig. 2. we plot the number of observations in each global monthly average from the RSS and STAR datasets for a representative satellite, NOAA-14. For a typical month, the RSS dataset contains a few percent more observations. This may be due to differences in quality control procedures, or to missing data in the STAR dataset. During the latter part of the NOAA-14 mission, the reverse is true, with the STAR dataset containing more data than the RSS dataset. This is not surprising, because RSS stopped using NOAA-14 data after December 2003 due to the increasing incidence of data gaps in the data feed from NESDIS. In any event, this difference needs to investigated more thoroughly is order to determine the root cause.

There is also an overall calibration difference between the RSS and STAR data due to different assumptions about the effective temperature of the cold space view. RSS uses 2.73 K for the deep space view, while STAR uses a value of 4.77 K, to account for the antenna spillover for the space view, which causes stray radiation from the spacecraft and the earth to increase the measured cold space radiance. The RSS brightness temperature values can easily be adjusted to account for the effects of the STAR cold space temperature using the expression

$$T_{adj} = 4.77 + (T_{tar} - 4.77)(T_b - 2.73)/(T_{tar} - 2.73) \quad ,$$

where T_{adj} is the adjusted brightness temperature, T_{tar} is the temperature of the warm calibration target, and T_b is the original brightness temperature. This transformation has only a very small effect on long term means.

The gridded, monthly datasets from the two STAR L1C datasets are then processed using the RSS merging software to produce new versions of the RSS merged data based on STAR L1C brightness temperatures. The goal of this part of the project is to try and estimate to what degree the differences between the RSS and STAR results are due to calibration differences, and what part is due to other parts of the merging procedure. Here we present the results of 3 merging calculations with the STAR datasets. For the SNO data, we produce two versions of the data. The first version, which we will call STAR_SNO_ZERO, no further calibration is performed except for the application of small offsets to match the absolute calibration of the satellites together. For the second version of the SNO data (STAR_SNO_ADJ), we apply our empirical non-linearity adjustment to the SNO calibrated data to see if it has any additional effect. We also perform our empirical adjustment to the NESDIS operational dataset (STAR_OP_ADJ). In the empirical method, instead of attempting to determine the physical source of the calibration errors unambiguously, we use an empirical error model for brightness temperature incorporating the target temperature and scene temperature correlation,

$$T_{meas} = T_0 + A_i + \alpha_i T_{tar,i} + \beta_i T_{scene} + \epsilon_i$$

where 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, $T_{tar.i}$. The parameter β_i describes the correlation of the calibration error with

Table I

L1C data source	Target Factors	Global Trend (1979-2004) (K/decade)
RSS	Ocean-only Merge	0.107
RSS_ADJ	Ocean-only Merge	0.104
STAR_OP_ADJ	Ocean-only Merge	0.102
STAR SNO_ZERO	All set to zero	0.171
STAR SNO_ADJ	Ocean-only Merge	0.119

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 parameters are determined using a least squares fitting procedure that minimizes the intersatellite differences. A detailed discussion of these methods is provided in (Mears and Wentz 2009).

These three versions of the STAR data are then compared with RSS version of the merged data with (RSS_ADJ) and without (RSS) the adjustment for the different cold space radiance assumptions. In all cases (except where they are set to 0.0), the target factors are determined using ocean only data, which is different from the approach we

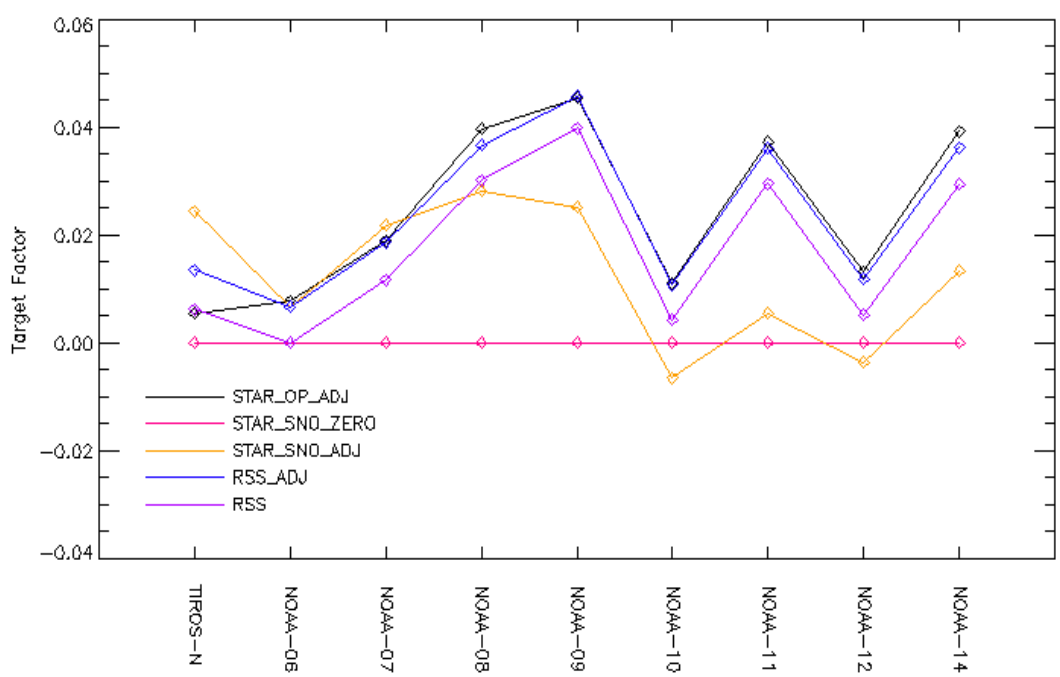

Fig. 3. Target factors α_i for each satellite, for each of the test datasets discussed in the text.

use for the RSS V3.2 data set which uses land and ocean data to determine the target factors. We are planning on adopting this approach for future versions of the RSS data since it is less likely to be affected by errors in the diurnal adjustment. In Table 1, we compare decadal trends in the global average for the data sets that results from each test, and in Fig. 3, we plot the target factors α_i used in each test.

There are some encouraging results in Table 1. First, despite the differences in some details between the RSS and STAR operationally calibrated datasets, the final results are very similar, indicating that the exact dataset used is not of critical importance. Second, the choice of the effective temperature for the space view does not make a large difference in the globally averaged trends (though it does make a significant difference in the absolute calibration and the globally averaged seasonal cycle – not shown). It is interesting to note that the SNO calibrated brightness temperatures result in a much larger trend when no further calibration steps are applied. However, when we apply the target factor calibration to the SNO calibrated brightness temperature, the trends are reduced so that it is only 0.01K/decade larger than the operationally calibrated case.

One way to assess the success of the calibration procedure is to examine the post-calibration residual differences between satellite pairs. In Fig. 4, we plot the standard deviation of the post-calibration differences for a number of representative satellite pairs. The plot reveals a number of interesting features. First, the two RSS cases (RSS and RSS_ADJ) have the lowest standard deviations. This is likely due to the more complete sampling in the RSS datasets, and to the manual removal of a few problem months. For the less complete sampling in the STAR dataset, we needed to relax the threshold mean number of observations per month in order to keep enough months to perform the merge. This allowed a number of satellite months through the quality control procedure that we

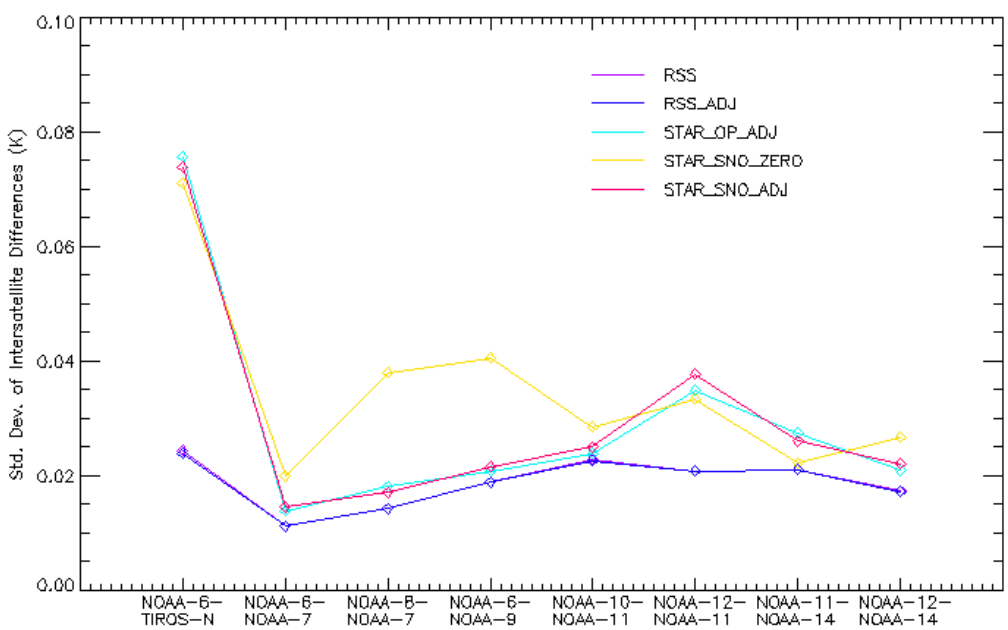

Fig. 4. Standard deviation of the post-calibration intersatellite differences for some representative pairs of satellites.

excluded in the RSS cases. An example of this is the last 12 months of the TIROS-N dataset, and is probably the cause of the increased standard deviation in the TIROS-N minus NOAA-06 differences. A more interesting effect is the reduction in standard deviation for a number of satellite pairs when the empirical target-factor calibration step is applied to the SNO calibrated brightness temperatures. The suggests the empirical calibration may remove some types of error that are missed by the SNO calibration. Thus we find that the empirical calibration may be complementary to the SNO calibration for MSU channel 2, confirming the earlier findings of (Zou and Wang 2010).

Experiment #3: Alternative merging algorithms. Another important difference between RSS and STAR processing is the procedure used to calculate and remove the residual intersatellite offsets. These are small offsets (usually on the order of a few tenths of a Kelvin) that remain after the previous adjustment and calibration steps are applied. Currently, RSS uses a separate offset for each satellite, channel, and 2.5 degree latitude bin, but does not include any east-west variability in these offsets. In contrast, STAR calculates an offset value for each satellite, channel, and 2.5 by 2.5 degree grid cell. In this section, we study the effects of two new offset removal methods on the RSS data. In one method (RSS_SEP_OFFSET), the offsets for land and ocean parts of each 2.5 degree latitude band are computed separately. The motivation for this approach is that at least part of the offsets is due to errors in the adjustment for the local diurnal cycle. Since the diurnal cycle, and thus errors in the adjustment, are much larger over land, it is likely that errors in the land diurnal cycle cause errors in the ocean offsets when a single offset is used for both land and ocean. Use of separate offsets from land and ocean avoids this problem. RSS_SEP_OFFSET can also be viewed as an intermediate step between the RSS standard processing (RSS_baseline), and the next experiment, RSS_GRID_OFFSET, where offsets are computed separately for each 2.5 x 2.5 degree grid cell. We (RSS) originally rejected this approach because we concluded that too much sampling error was aliased into the final results.

In Fig. 5, we plot maps of the difference in decadal trends between RSS_baseline, and other datasets. In the top panel, we show the difference between RSS_baseline and the STAR V2.0 TMT merged dataset. The map has 3 important features. First, there is a lot of vertical banding in the differences, particularly in the mid-latitudes. This is very likely to be due to spatial sampling noise affecting the gridded offsets in the STAR dataset. There also appear to be a land-ocean contrast, with the RSS baseline trends being greater than the STAR trends for land, while the opposite is true (on average) for ocean. Third, there is an overall latitude dependence in the difference. In the middle panel, we plot trends from RSS_baseline minus RSS_SEP_OFFSET. The map appears to exhibit some of the land-ocean contrast we see in the RSS_baseline minus STAR map, though with a slightly different magnitude. In the lower panel, we plot the difference between RSS_baseline and RSS_GRID_OFFSET. This map show both the land-ocean contrast, and mid-latitude vertical banding very similar to that observed in the RSS_baseline minus STAR 2.0 case.

These results strongly suggest that much of the difference in the spatial pattern of trends is due to the different methods use to calculate the intersatellite offsets, while the differences in overall trend, and the slowly varying north-south pattern is due to the differences between the SNO and the empirical calibrations methods.

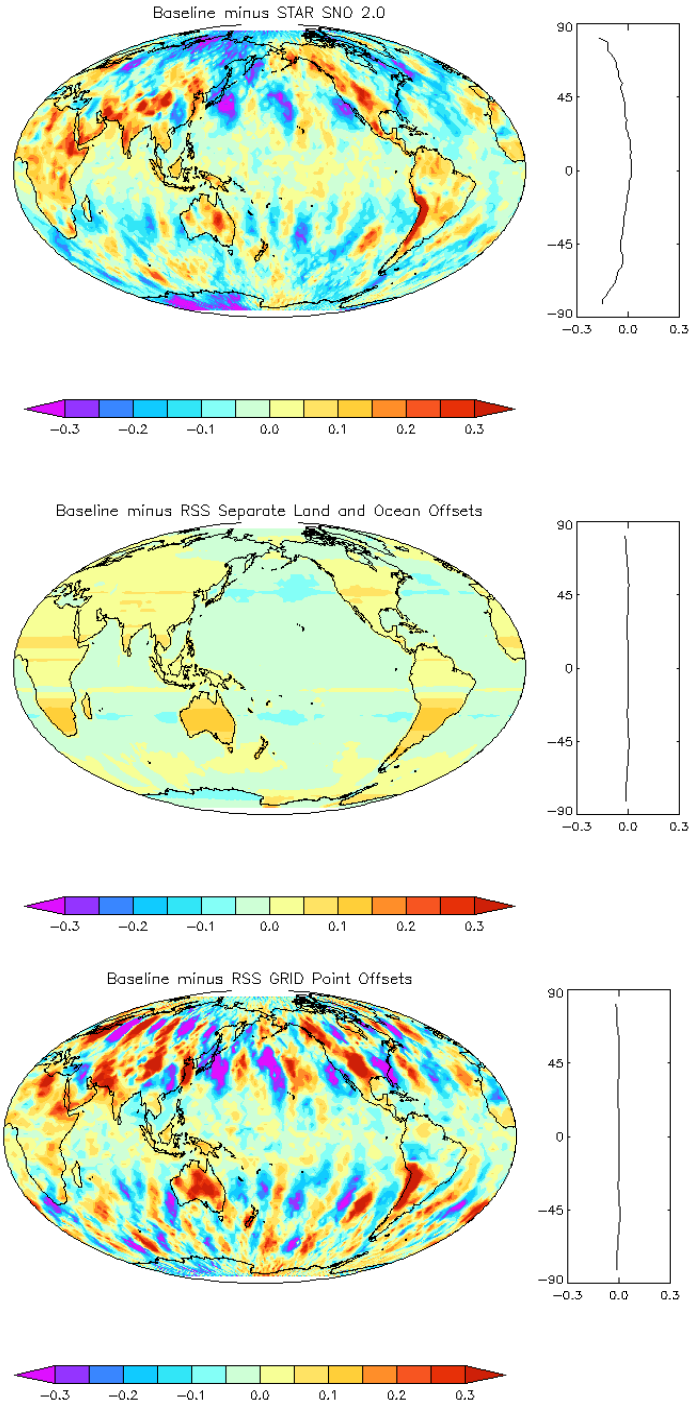

Fig. 5. Maps of differences between decadal trends (1979-2004) between the RSS baseline dataset (an MSU-only version of RSS V3.2) and various other datasets. The top panel is the difference between the RSS baseline and the STAR SNO V2.0 merged dataset, and represents the difference we are trying (at least in part) to explain in experiment #3. The middle panel is the difference between the RSS baseline, and a version of the RSS_SEP_OFFSET. The bottom panel is a similar difference between RSS baseline, and RSS_GRID_OFFSET

Activity planned for the next 6 months.

1. Continue investigating sampling difference between the two datsets.
2. Investigate the impact of varying the swath width of the data included on the RSS data. Currently, RSS uses the central 5 fields of view, while STAR uses the central 7 fields of view.
3. Expand the analysis to include MSU channels 3 and 4.
4. Expand the analysis to include the effects of AMSU data.

References

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.

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, 2009: 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.

Zou, C. Z., M. Gao, and M. D. Goldberg, 2009: Error structure and atmospheric temperature trends in observations from the Microwave Sounding Unit. *Journal of Climate*, **22**, 1661-1681.

Zou, C. Z., M. D. Goldberg, Z. Cheng, N. C. Grody, J. T. Sullivan, C. Cao, and J. Tarpley, 2006: Recalibration of microwave sounding unit for climate studies using simultaneous nadir overpasses. *Journal of Geophysical Research*, **111**. file://X:\DOCUMENT\Journal_Papers\Zou_JGR_2006.pdf.

Zou, C., and W. Wang, 2010: Stability of the MSU-derived atmospheric temperature trend. *Journal of Atmospheric and Oceanic Technology*,

NOAA Climate and Global Change Program
Semi-annual Progress Report for
June 1, 2010 through November 30, 2010

Generating consistent radiance SDRs and deep-layer atmospheric temperature TCDRs from the MSU/AMSU/SSU temperature-sounding channels

NOAA Award No. NA09NES4400015

Science Team Lead: *Cheng-Zhi Zou*
Center for Satellite Applications and Research (StAR)
NOAA/NESDIS
Camp Springs, MD 20746

Science Team Members: *Sid Boukabara, NESDIS/StAR*
Lidia Cucurull, NESDIS/StAR
Dick Dee, ECMWF
Qiang Fu, University of Washington
Thomas Kleespies, NESDIS/StAR
Carl Mears, Remote Sensing Systems
John Woollen, NWS/NCEP

Collaborators: *Ralph Ferraro/NESDIS/StAR*
Ben Ho/UCAR

Introduction

This project is a complex, multi-institutional effort aimed at generating climate quality temperature datasets from satellite-borne atmospheric sounders. This report only covers the part of this work assigned to **Remote Sensing Systems**, which is to perform comparisons of various satellite products with each other, and with other sources of atmospheric temperature measurements. The purpose of this part of the work is to use such comparisons to help determine the reasons for, and significance of ,any discrepancies between satellite datasets. This report covers 6 months of work, from June 1, 2010 to November 30, 2010.

One of the most important goals of this part of the work is to intecompare results from the satellite-based datasets with those from other data sources, such as radiosondes or radio occultation. In this report, we report results we have obtained by comparing middle tropospheric temperatures from the three currently active satellite-based datasets (STAR, RSS, and UAH) with results from homogenized radiosonde datasets. Our report is divided into three sections.

1. Description of radiosonde datasets
2. Simple time series comparisons
3. Comparisons of short-term trends

Description of radiosonde datasets.

Before radiosonde measurements can be used to describe long-term changes, these inhomogeneities need to be characterized and removed to the extent possible. A number of groups have produced homogenized radiosonde datasets. For comparison with our satellite measurements of TLT, we chose four of the most recent homogenized datasets. These datasets are either available as gridded TLT

measurements, or contain enough information so that it is possible for us to construct a gridded TLT dataset. These datasets were constructed using automated methods to find and estimate the size of "breakpoints" in each radiosonde's time series which are then used to create adjusted versions of the radiosonde data with the effects of the detected breakpoints removed. We list each radiosonde dataset below.

i. HadAT. HadAT is constructed at the Hadley Centre by analyzing the difference between the monthly time series from a given radiosonde and a composite of neighboring radiosondes (Thorne et al. 2005). The earlier, human-analyzed Lanzante-Klein-Seidel (LKS) (Lanzante, Klein, and Seidel 2003) dataset is used as a starting point, and LKS is assumed to contain no large inhomogeneities.

ii. RAOBCORE. The Radiosonde Observation Correction using Reanalysis (RAOBCORE) (Haimberger 2007) dataset is constructed at the University of Vienna by evaluating the difference between individual (daily 00Z or 12Z) radiosonde observations and background forecast fields calculated during the construction of the ERA-40 reanalysis. Since ERA-40 uses measurements from MSU and AMSU during its construction, RAOBCORE cannot be considered to be completely independent of the satellite measurements. RAOBCORE uses the background fields both to detect and to estimate the size of breakpoints in the satellite record.

iii. *RICH.* In contrast, the Radiosonde Innovation Composite Homogenization (RICH)(Haimberger, Tavolato, and Sperka 2008) dataset uses the breakpoints identified by the RAOBCORE analysis, but estimates the size of any jumps using a composite of neighboring stations. Thus, the amount of adjustment performed for each breakpoint is independent of ERA-40 and satellite measurements.

iv. *IUK.* The radiosonde Iterative Universal Kriging (IUK) (Sherwood et al. 2008) dataset was constructed at Yale University using an analysis method that simultaneously evaluates long-term variability and observing biases using an iterative kridging approach (Sherwood 2007). This method also depends on the analysis of neighboring stations, but it is designed to be robust with respect to errors due to inhomogeneities in these neighbors. This dataset is not being updated.

v. *RATPAC.* The Radiosonde Atmospheric Temperature Products for Assessing Climate (Free et al. 2005) is produced at the NOAA Air Resources Laboratory. The version of the dataset we use here, RATPAC-B, is an extension the LKS dataset (Lanzante, Klein, and Seidel 2003) to the present, though no adjustments are made to the individual radiosondes after 1997. We are also investigating a subset of this dataset defined by Randel and Wu (2006), RATPAC_RW, which is intended to be free of long-term inhomogeneities.

In order to compare the results of these datasets with satellite-derived TLT datasets, TLT equivalent versions of the radiosonde datasets need to be calculated. This is done by calculating weighted vertical averages of the discrete-level radiosonde data using weights calculated so that the effective weighting function of the resulting average closely matches the TLT weighting functions. For HadAT and RAOBCORE, this is already performed by the originating institutions; both these datasets are available as gridded TLT equivalent temperatures. RICH data are available as gridded products at 12 pressure levels from 850 hPa to 30 hPa. Using these temperatures, along with surface temperatures from HadCRUT3, we calculated equivalent TLT temperatures for RICH.

The IUK dataset is available as adjusted monthly profiles at each radiosonde location. We binned each monthly measurement into a 10 degree by 10 degree grid using "bucket" sampling, and then calculated TLT equivalent temperatures using HadCRUT3 surface temperatures.

Simple Time Series Comparions.

We compare the satellite and homogenized radiosonde temperature anomalies by comparing time series and trends of large spatial scale averages. Large spatial scale averages are more useful because of the significant uncertainties in the measurements from isolated radiosondes and in single satellite grid points. Over larger spatial scales, these uncertainties are reduced by the averaging procedure. We choose to focus on global averages (75S to 75N), the southern extratropics (SXT, 75S to 30S), the tropics (30S to 30N), and the northern extratropics (NXT, 30N to 75N). In all regions, but especially the tropics and the SXT, radiosonde spatial coverage is far from complete. Figure 1 shows a typical radiosonde sampling pattern.

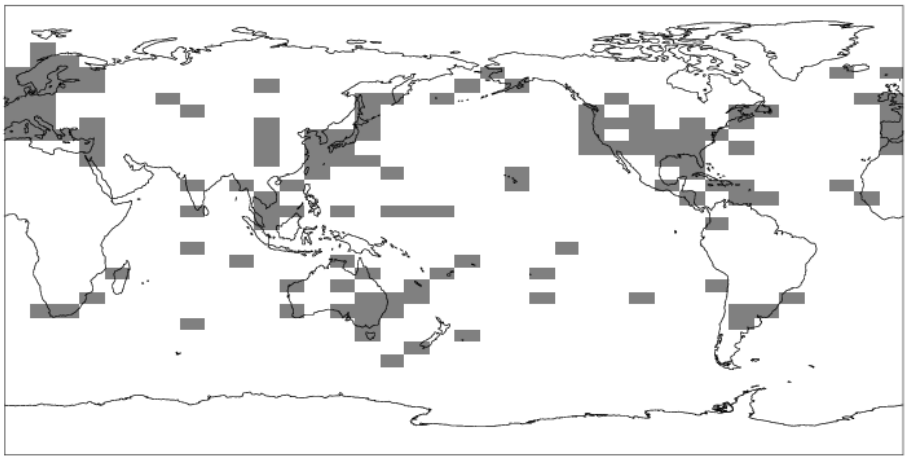

Fig. 1. Sampling pattern for the HadAT dataset for January, 2003. Grey boxes are valid data, and white regions are missing data.

A simple way to construct time series would be to use area-weighted means of the available data. Comparing time series of simple area-weighted (AW) global averages of radiosonde data with the area-weighted means of the spatially complete satellite data will lead to substantial discrepancies, due to the large unsampled areas in the SXT where the trends are much lower than in the well-sampled NXT. Additional discrepancies occur because of the sampling pattern in the tropics which excludes the eastern tropical Pacific Ocean, where the ENSO signal is often the strongest. In (Mears and Wentz 2009) we established that the best approach to resolving these discrepancies is to sample the satellite data at the actual radiosonde sampling for each month, and then computed an area-weighted average from the sub-sampled data for each month to produce a "global" average. This is intended to modify the satellite means so that they closely match the area-weighted radiosonde means, and automatically take into account the presence or absence of a radiosonde measurement for a given location and month. We refer to the sampled satellite means as "sampled at radiosonde locations" (SRL).

Because the simple area weighting approach is still often used by other workers to perform radiosonde/satellite intercomparisons, we show results for both methods. In Fig. 2, we show global-scale TMT difference time series for each satellite dataset relative to each of the adjusted radiosonde datasets. For each radiosonde dataset, we show 2 sets of plots. On the left, we show the difference between AW global averages, and on the right, we show differences between SRL satellite means, and AW radiosonde means. In all cases, the AW time series show significantly larger differences, confirming the need to use the SRL satellite averages for comparison purposes.

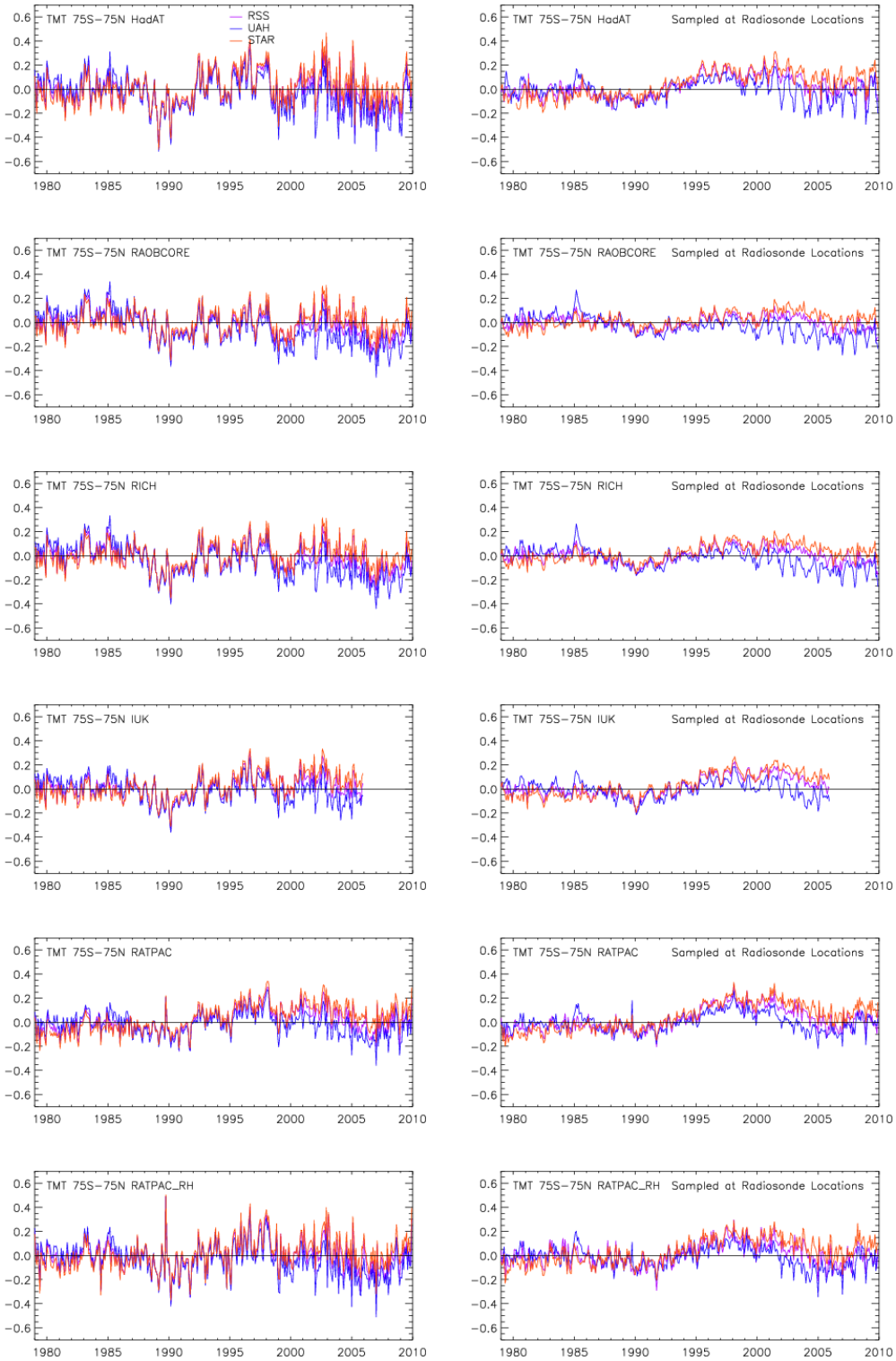

Fig. 2. Difference time series between the three satellite datasets (RSS, UAH, and STAR), and 6 radiosonde datasets. The plots on the left side use area-weighted (AW) averages for the satellite data, and the plots on the right use averages that are sampled at the radiosonde locations (SRL). The averages are near-global in extent, from 75S to 75N.

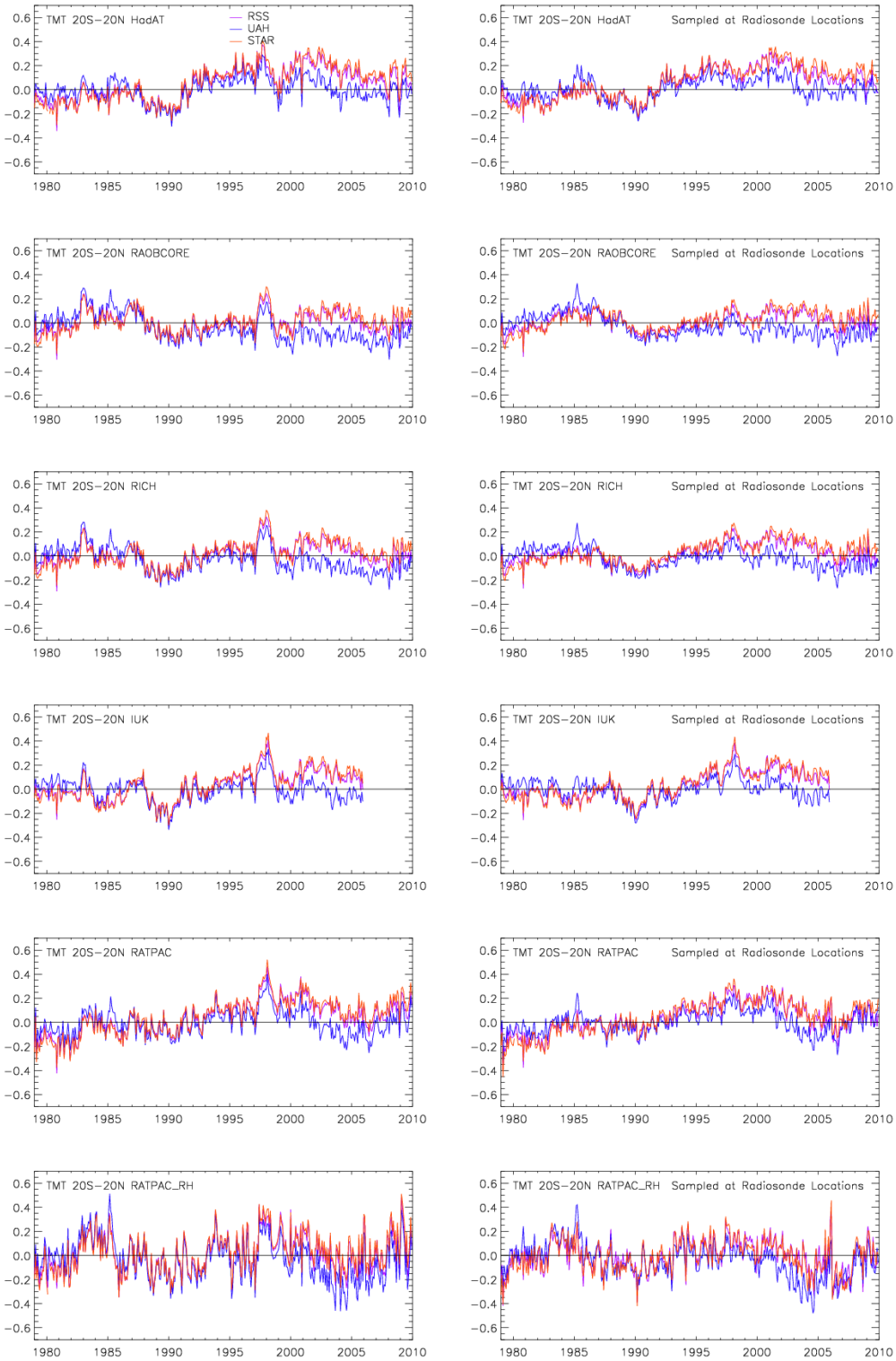

Fig. 3. Same as Fig. 2, except the area considered is the deep tropics (20S to 20N)

Turning our attention to the SRL time series, we note some common features. Perhaps the most obvious is that all satellite datasets warm relative to all radiosonde datasets from about 1990 to about 1998. In all cases, the UAH dataset warms a bit less than the RSS or STAR datasets. This feature has been previously discussed in the literature as it occurs during the period where the RSS warms relative to the UAH datasets (Christy et al. 2007; Christy, John R. and Norris, William B. 2009; Randall and Herman 2008). The better agreement between UAH and the radiosondes has been used by these authors to argue that the RSS dataset contains a warming bias during this period. These assertions provide motivation for the current work, which takes a broader view than these earlier works, since we include more versions of the radiosonde dataset, and do not restrict the focus of our work to this time period. We will revisit this feature later, when we consider short-term trend analysis. It is interesting to note that the RAOBCORE sonde analysis, which is often asserted to be corrupted by anomalous warming during this period (Christy, et al. 2010), is the dataset that agrees best with <u>all</u> satellite datasets. Other features of similar scale include the cooling in the satellite data relative to the radiosonde from about 1985 to about 1989. This feature is strongest in the UAH data. The satellites also tend to cool over the 2000-2005 period, with the cooling appearing to be strongest in the UAH data. Also notable is the appearance of a strong seasonal cycle in the differences involving the UAH data – this appears to start after the advent of the AMSU series of instruments in 1998, suggesting a problem with merging the MSU and AMSU temperatures in the UAH dataset.

The two important points to take away from this part of the analysis are as follows.

1. Despite the earlier focus on the 1990-1998 warming of the satellite (in particular RSS) relative to the radiosondes, there are many other time periods with significant relative warming or cooling events.
2. The 3 satellite datasets are more similar to each other than they are to any of the radiosonde datasets. The satellite sonde comparisons show many common features across both satellite and radiosonde datasets.

In Fig. 3 is a similar plot, except averaged over the deep tropics (20S to 20N) instead of the entire globe. This set of plots shares many of the features of the global plots, such as the relative warming in the satellite dataset in the mid 1990's, and the relative cooling in the late 1980's and early in the 21st century. One important difference is that there appears to be more differences between the radiosonde dataset in the 1990's, with the satellites showing strong warming relative to radisondes in IUK and RATPAC, with considerably less warming relative to RAOBCORE and to a lesser extent, RICH. For HadAT, most of the warming is shifted to a short time period in the early 1990's. It is not surprising that the differences are getting larger because the number of radiosonde stations in the sample has decreased, leading to more variability. Also note that the use of the SRL averaging removed the bump in the AW curves associated with the larger ANSO event in 1997-1998.

Comparisons of short-term trends

One of the dangers inherent in making comparisons with radiosonde trends is that it is likely that the radiosonde datasets contain residual time-dependent biases that were not removed by the homogenization procedures. The problems are often in the form of undetected jumps in the radiosonde time series caused by changes in instrumentation or measurement practices. One approach that has been used is the analysis of short term trends, with the hope that the problems in the radiosonde datasets are located at different times than the problems with the satellite datasets. Randall and Herman (2008) used the Randel and Wu (2006) subset of the RATPAC radiosondes to analyze 5 and 10 year trends in UAH and RSS data. While the approach of using short-term changes in radiosondes data can yield useful information, as we shall see below, there are some problems with the Randall and Herman

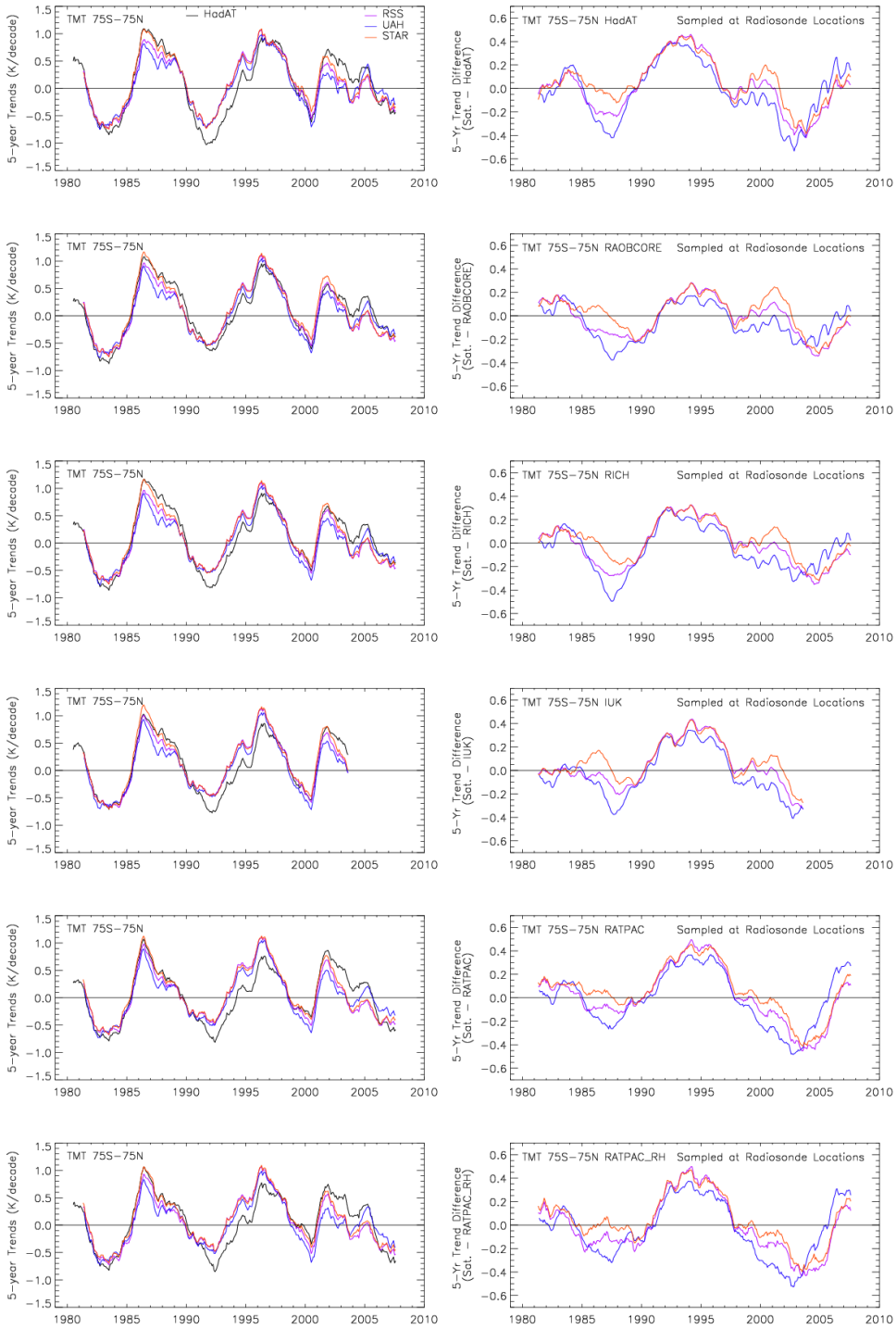

Fig. 4. Short-term trend analysis of global-scale TMT averages. The plots on the left side are time series of 5 year trends in radiosonde and SRL averages for each satellite dataset. The plots on the right side are differences between the 5 year satellite trends and the sonde trends.

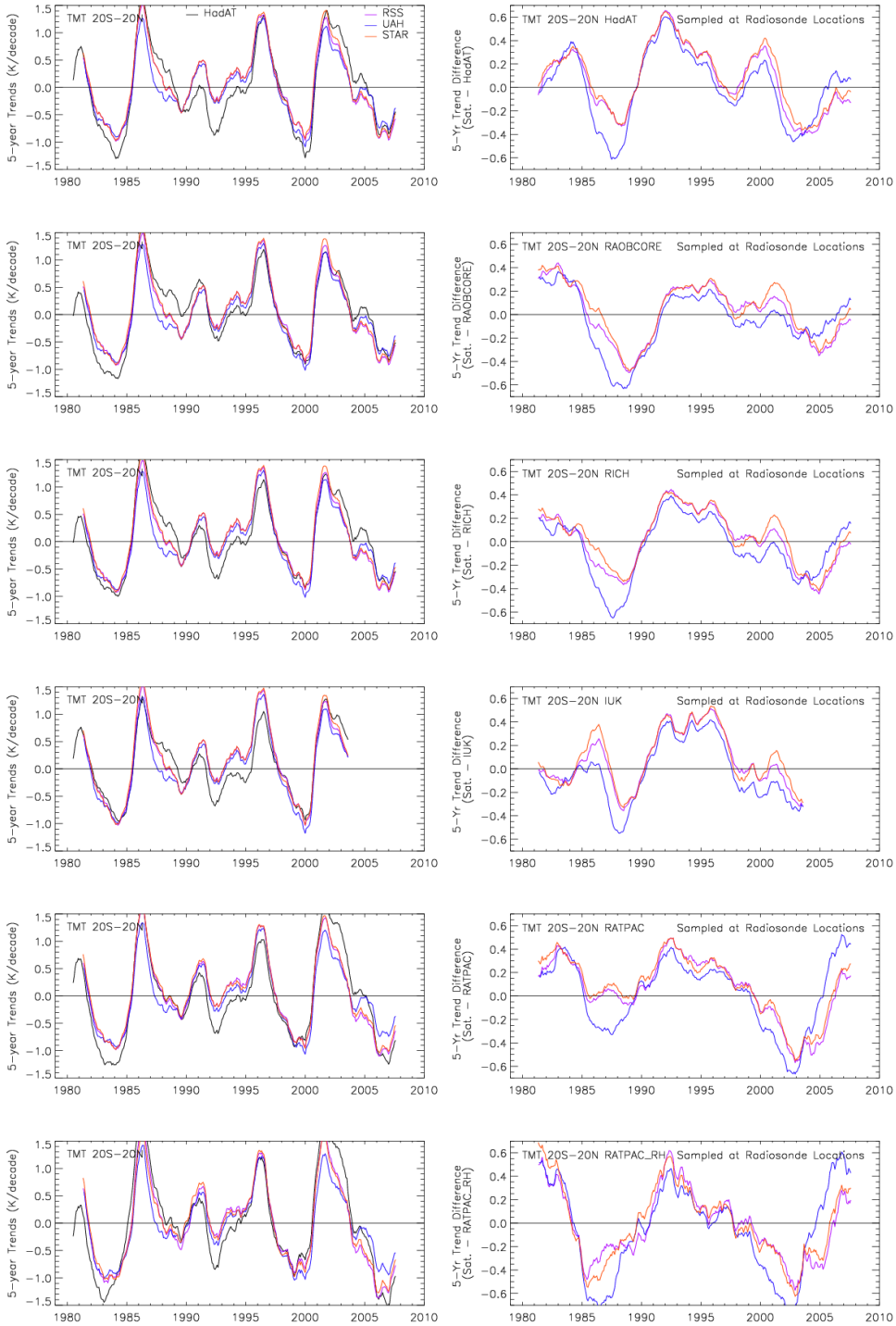

Fig. 5. Same as Figure 4, except for the deep tropics (20S-20N)

approach.

1. They did not subsample the satellite data at the satellite locations. From the plots presented above, we can see that this is an important step for improving the agreement between the two types of data.
2. They focus their attention on one time period, and do not present results from other time periods when the radiosondes and satellite datasets tend to disagree.
3. They only consider 1 homogenized radiosonde dataset, RATPAC. It is not known whether the results are similar for the other radiosonde datasets.
4. They focused a lot of attention on the difference between lower tropospheric and middle tropospheric trends. This differencing procedure tends to add additional noise to the results, making them more difficult to interpret accurately.
5. The idea of using short-term trends is probably best for analyzing comparisons between individual radiosonde stations and the satellite data. When an average over a large number of stations is used (in the Randall and Herman case, 47), it is likely that the signals from the inhomogeneities in single stations is lost due to averaging.

Given these problems, we conclude that using Randall and Herman (2008) as an arbiter of satellite data quality is of limited use. Problem number 5 is probably very important, but we will ignore this issue for the time being. Instead, we will extend the RH analysis using methods intended to fix the first 4 problems.

1. We will sub-sample the satellite data at the radiosonde locations
2. We will present short-term trend difference results for the entire time period
3. We will consider results for 5 radiosonde datasets.
4. We will focus on the trends from each channel separately.

In Fig. 4, we plot results from a short-term trend analysis using global-scale SRL averages for TMT. The plots on the left side of the figure are trends over 5 year periods, with the trend value plotted at the time at the center of the trend period. The plots on the right hand side of the figure are the differences between the 5 year satellite trends and the radiosonde trends. Much of the information present in these figures simply reinforces the conclusions we reached above when considering the time series differences. The plots are very similar for different satellite/radiosonde pairs, with satellites cooling relative to radiosondes in the late 1980's, warming relative to radiosondes from 1990 to 1998, and cooling relative to radiosondes from 2000-2005. This indicates that:

1. The satellite datasets have more in common with each other than they do with any radiosonde dataset.
2. Radiosonde datasets are closer to each other than they are to any satellite dataset.

The use of the short term time series does make it a little easier to pin-point the periods during which the satellite time series are most different from each other. We conclude that there are 4 periods where these differences are the largest. First, from 1986 to 1989, where STAR warms relative to UAH, and to a lessor extent RSS. Second from 1993 to 1997, where both RSS and STAR warm relative to UAH, and 1998-2003, where STAR warms relative to UAH, and to a lessor extent relative to RSS. Beginning in 2004, this pattern reverses itself, and STAR and RSS cool relative to UAH. The importance of the various periods to the overall shift in the time series relative to each other is roughly equal to the area between the curves, which can most easily be seen in the right hand series of plots. Given this criteria, the most important periods are 1986-1989 and 1998-2004. When considering trends, the 1993-1997

period gains some importance due to its position near the center of the time series. This is in contrast to findings in (Randall and Herman 2008; Christy et al. 2007; Christy, John R. and Norris, William B. 2009)

Now that we have established these time periods, an interesting question to ask is which satellite

Table 1.
Area between 5-year Trend Curves, 75S - 75N

Radiosonde Dataset	Time Period	RSS	UAH	STAR
HadAT	1985-1989	-0.068	-0.102	**-0.007**
RAOBCORE	1985-1989	-0.079	-0.117	**-0.020**
RICH	1985-1989	-0.100	-0.138	**-0.041**
IUK	1985-1989	-0.051	-0.094	**0.006**
RATPAC	1985-1989	-0.042	-0.070	**0.004**
RATPAC_RH	1985-1989	-0.074	-0.095	**-0.018**
HadAT	1992-1997	0.159	**0.118**	0.150
RAOBCORE	1992-1997	0.095	**0.054**	0.091
RICH	1992-1997	0.131	**0.091**	0.128
IUK	1992-1997	0.172	**0.130**	0.166
RATPAC	1992-1997	0.202	**0.163**	0.205
RATPAC_RH	1992-1997	0.195	**0.151**	0.195
HadAT	1998-2003	-0.075	-0.143	**-0.033**
RAOBCORE	1998-2003	**-0.013**	-0.078	0.026
RICH	1998-2003	-0.051	-0.115	**-0.011**
RATPAC	1998-2003	-0.108	-0.163	**-0.065**
RATPAC_RH	1998-2003	-0.118	-0.191	**-0.080**
HadAT	2004-2006	-0.044	**0.001**	-0.034
RAOBCORE	2004-2006	-0.074	**-0.034**	-0.067
RICH	2004-2006	-0.071	**-0.031**	-0.064
RATPAC	2004-2006	-0.069	**-0.002**	-0.057
RATPAC_RH	2004-2006	-0.060	**0.007**	-0.045

dataset is closer to the radiosondes during these periods. In Table 1, we present the area between the satellite trends and the radiosonde trends for each period, and for each radiosonde-satellite pair. The units are degrees K, and the number roughly corresponds* to the offset between time series that each time period causes. In each case, the number corresponding to the satellite that is closest to the radiosondes is shown in bold. For the first period (1985-1990), STAR is much closer to all radiosonde datasets, with RSS next closest. For the second period, UAH is closer to the radiosondes than either RSS or STAR, which are roughly equidistant. This confirms the conclusions of Randall and Herman (2008), but it is important to note that all satellites disagree markedly with radiosonde over this period, with the distance between UAH and radiosondes is usually twice as large as the distance between

* If the plots were instantaneous slope differences, the correspondance would be exact, since the integral of a derivative returns the original function plus an arbitrary constant offset. Because we are using 5-year trends, the integral is related to the change in offset in a smoothed time series.

Table 2.

Area between 5-year Trend Curves, 75S - 75N

Radiosonde Dataset	Time Period	RSS	UAH	STAR
HadAT	1985-1990	-0.057	-0.149	**-0.0**5
RAOBCORE	1985-1990	-0.121	-0.207	**-0.089**
RICH	1985-1990	-0.111	-0.197	**-0.080**
IUK	1985-1990	-0.028	-0.113	**0.000**
RATPAC	1985-1990	**-0.001**	-0.087	0.024
RATPAC_RH	1985-1990	**-0.127**	-0.221	-0.139
HadAT	1992-1998	0.161	**0.116**	0.162
RAOBCORE	1992-1998	0.120	**0.071**	0.118
RICH	1992-1998	0.170	**0.121**	0.168
IUK	1992-1998	0.221	**0.170**	0.219
RATPAC	1992-1998	0.185	**0.132**	0.175
RATPAC_RH	1992-1998	0.146	**0.096**	0.131
HadAT	1998-2004	**-0.005**	-0.073	0.026
RAOBCORE	1998-2004	**0.017**	-0.047	0.043
RICH	1998-2004	-0.034	-0.098	**-0.008**
RATPAC	1998-2004	-0.118	-0.188	**-0.108**
RATPAC_RH	1998-2004	**-0.137**	-0.250	-0.158
HadAT	2004-2007	-0.070	**-0.019**	-0.060
RAOBCORE	2004-2007	-0.073	**-0.026**	-0.062
RICH	2004-2007	-0.085	**-0.038**	-0.074
RATPAC	2004-2007	-0.063	**0.038**	-0.043
RATPAC_RH	2004-2007	-0.049	0.080	**-0.022**

satellite datasets. An exception to this occurs for RAOBCORE, which is significantly closer to satellites over this period. This is an interesting finding in light of Christy's dismissal of the RAOBCORE dataset as flawed (Christy, et al. 2010). For 1998-2003, STAR is the closest to the radiosondes (except RAOBCORE), with RSS a close second. For 2004-2006, UAH is closer to the radiosondes, with STAR a bit closer than RSS. In Table 2, we present a similar set of results for the deep tropics (20S to 20N). Here the story is mostly the same, except that RSS and STAR are more similar over the tropics, and most of the differences between them are negligible.

The above analysis suffers from the limitation that the time periods of interest were chosen subjectively. Another approach could be to evaluate the RMS difference between the 5 year satellite trends and the 5 year radiosonde trends over the entire time period. In Table 3, we present these results for each radiosonde/satellite pair, and for 4 averaging regions (Globe (75S-75N), Southern Extratropics (75S-20S), Tropics (20S-20N), and Northern Extratropics (20N-75N)). Again, STAR is the satellite dataset that is most closely aligned to radiosonde datasets over 5 year time periods, with RSS in almost as good agreement with the radiosondes in the tropics.

Discussion

Table 3.

RMS Differences between 5-year trends.

Radiosonde Dataset	RSS	UAH	STAR
Globe 75S-75N			
HadAT	0.219	0.236	**0.198**
RAOBCORE	0.157	0.157	**0.150**
RICH	0.185	0.198	**0.163**
IUK	0.191	0.204	**0.181**
RATPAC	0.236	0.230	**0.221**
RATPAC_RH	0.238	0.243	**0.215**
South 75S-20S			
HadAT	0.334	0.322	**0.288**
RAOBCORE	0.250	0.234	**0.219**
RICH	0.316	0.305	**0.279**
IUK	0.256	0.252	**0.236**
RATPAC	0.258	**0.222**	0.251
RATPAC_RH	0.276	**0.218**	0.286
Tropics 20S-20N			
HadAT	0.283	0.297	**0.280**
RAOBCORE	**0.234**	0.252	0.235
RICH	0.237	0.257	**0.228**
IUK	**0.248**	0.259	0.255
RATPAC	0.270	0.294	**0.267**
RATPAC_RH	**0.288**	0.386	0.315
North 20N-75N			
HadAT	0.205	0.228	**0.189**
RAOBCORE	0.133	**0.132**	0.146
RICH	0.149	0.174	**0.139**
IUK	0.167	0.189	**0.159**
RATPAC	0.262	0.258	**0.242**
RATPAC_RH	0.271	0.272	**0.241**

In all cases we find that there are several time periods during which there is substantial disagreement between 5-year trends in radiosonde datasets and 5-year trends in the satellite datasets. When data from different satellite – radiosonde pairs is examined using both time series comparisons, and using 5-year trends, the results indicate that all satellite/sonde differences share common features, and that in most cases, the differences between radiosondes and satellites is much larger than between different satellite datasets, or between different radiosonde datasets. Given the current state of

knowledge, it is difficult or impossible to figure out whether this commonality is due to shared problems in the satellite datasets, or to shared problems with the radiosonde datasets, or a combination of both. The problem is that both types of datasets are likely to retain common biases within their respective types. There are several potential causes of common biases, including shared source data within each dataset type, or common errors or problems in the homogenization methods.

The approach of using short term trends for the analysis highlights the fact that the radiosonde and satellite datasets are quite different on short (i.e. 5-year) time scales. Note that sometimes these differences cancel over longer time periods, leading to (perhaps false) conclusions about the agreement between satellite and radiosonde datasets on multi-decadel time scales.

Since differences between all satellite datasets and the radiosonde datasets are larger than the differences between the satellite datasets, we conclude that trying to determine which satellite dataset is "better" based on comparisons with radiosonde datasets is an activity with limited value. We have shown that the assertions made by Randall and Herman (2008) and Christy et al (2008, 2010) that UAH is in better agreement with radiosondes than RSS and STAR is not valid outside the limited range of time that is the focus of these works.

Work Planned for the Next 6 months

- Extending this analysis to the lower stratosphere channel (TLS)
- Continue investigating sampling differences between RSS and STAR datasets.
- Investigate the impact of varying the swath width of the data included on the RSS data. Currently, RSS uses the central 5 fields of view, while STAR uses the central 7 fields of view.

References

Christy, John R., and Norris, William B. 2009. Discontinuity Issues with Radiosonde and Satellite Temperatures in the Australian Region 1979–2006. *Journal of Atmospheric and Oceanic Technology* 26, no. 3: 508-522.

Christy, John R., William B. Norris, Roy W. Spencer, and Justin J. Hnilo. 2007. Tropospheric temperature change since 1979 from tropical radiosonde and satellite measurements. *Journal of Geophysical Research* 112. doi:doi:10.1029/2005JD006881. file://X:\DOCUMENT\Journal_Papers\Christy_JGR_2007.pdf.

Christy,, John R., Benjamin M. Herman, R. Peilke, Sr., P. Klotzbach, Richard T. McNider, Justin J. Hnilo, Roy W. Spencer, T. Chase, and David H. Douglass. 2010. What do Observational Datasets Say about Modeled Tropospheric Temperature Trends sinde 1979? *Remote Sensing* 2: 2148-2196. doi:doi: 10.3390/rs2092148.

Free, M., Dian J. Seidel, James K. Angell, John R. Lanzante, Imke Durre, and Thomas C. Peterson. 2005. Radiosonde Atmospheric Temperature Products for Assessing Climate (RATPAC): A new dataset of large-area anomaly time series. *Journal of Geophysical Research* 110. doi:doi:10.1029/2005JD006169. file://X:\DOCUMENT\Journal_Papers\Free_et_al_2005_JGR_RATPAC.pdf.

Haimberger, Leopold. 2007. Homogenization of radiosonde temperature time series using innovation statistics. *Journal of Climate* 20, no. 7: 1377-1403.

Haimberger, Leopold, Christina Tavolato, and S Sperka. 2008. Towards the elimination of warm bias in historic radiosonde records -- some new results from a comprehensive intercomparison of upper air data. *Journal of Climate* 21: 4587-4606.

Lanzante, John R., S. Klein, and Dian J. Seidel. 2003. Temporal homogenization of monthly radiosonde temperature data. Part I: Methodology. *Journal of Climate* 16: 224-240.

Mears, Carl A., and Frank J. Wentz. 2009. 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.

Randall, Robb M, and Benjamin M. Herman. 2008. Using limited time period trends as a means to determine attribution of discrepancies in microwave sounding unit-derived tropospheric temperature time series. *Journal of Geophysical Research* 113. doi:doi:10.1029/2007JD008864. file://X:\DOCUMENT\Journal_Papers\Randall_and_Herman_2008.pdf.

Randel, William J., and Fei Wu. 2006. Biases in stratospheric and tropospheric temperature trends derived from historical radiosonde data. *Journal of Climate* 19, no. 10: 2094-2104.

Sherwood, Steven C. 2007. Simultaneous detection of climate change and observing biases in a network with incomplete sampling. *Journal of Climate* 20, no. 15: 4047-4062.

Sherwood, Steven C., Cathryn L. Meyer, Robert J. Allen, and Holly A. Titcher. 2008. Robust tropospheric warming revealed by iteratively homogenized radiosonde data. *Journal of Climate* 21, no. 20 (October 1): 5336-5352.

Thorne, Peter W., David E. Parker, Simon F.B. Tett, Phillip D. Jones, Mark McCarthy, Holly Coleman, and Philip Brohan. 2005. Revisiting radiosonde upper-air temperatures from 1958 to 2002. *Journal of Geophysical Research* 110. doi:doi:10.1029/2004JD005753. X:\DOCUMENT\Journal_Papers\thorne_JGR_2005.pdf.