

Reconciling tropospheric temperature trends  
from the microwave sounding unit

Stephen Po-Chedley

A dissertation submitted in partial fulfillment  
of the requirements for the degree of

Master of Science

University of Washington

2012

Program Authorized to Offer Degree: Department of Atmospheric Sciences



University of Washington

**Abstract**

Reconciling tropospheric temperature trends  
from the microwave sounding unit

Stephen Po-Chedley

Chair of the Supervisory Committee:  
Professor Qiang Fu  
Department of Atmospheric Sciences

The University of Alabama at Huntsville (UAH), Remote Sensing Systems (RSS), and the National Oceanic and Atmospheric Administration (NOAA) have constructed long-term temperature records for deep atmospheric layers using satellite microwave sounding unit (MSU) and advanced microwave sounding unit (AMSU) observations. However, these groups disagree on the magnitude of global temperature trends since 1979, including the trend for the mid-tropospheric layer (TMT). This study evaluates the selection of the MSU TMT warm target factor for the NOAA-9 satellite using five homogenized radiosonde products as references. The analysis reveals that the UAH TMT product has a positive bias of  $0.051 \pm 0.031$  in the warm target factor that artificially reduces the global TMT trend by an estimated  $0.042 \text{ K decade}^{-1}$  for 1979 - 2009. Accounting for this bias, we estimate that the global UAH TMT trend should increase from  $0.038 \text{ K decade}^{-1}$  to  $0.080 \text{ K decade}^{-1}$ , effectively eliminating the trend difference between UAH and RSS and decreasing the trend difference between UAH and NOAA by 47%. This warm target factor bias directly affects the UAH lower tropospheric (TLT) product and tropospheric temperature trends derived from a combination of TMT and lower stratospheric (TLS) channels.



## TABLE OF CONTENTS

	Page
List of Figures . . . . .	iii
List of Tables . . . . .	vii
Chapter 1: Introduction . . . . .	1
1.1 History of the microwave sounding unit discrepancies . . . . .	1
1.2 The expectation of surface warming and temperature amplification . . . . .	11
Chapter 2: MSU and radiosonde datasets . . . . .	15
2.1 Radiosondes . . . . .	19
2.2 MSU/AMSU observations . . . . .	22
2.3 Dataset intercomparison . . . . .	24
Chapter 3: Treatment of the NOAA-9 Satellite . . . . .	34
3.1 The NOAA-9 warm target factor . . . . .	34
3.2 Determining the bias in the NOAA-9 target factor . . . . .	38
Chapter 4: Conclusion . . . . .	45
Appendix A: UAH response to this work . . . . .	57
A.1 Our Response to Recent Criticism of the UAH Satellite Temperatures . . . . .	57
A.2 The Lower Tropospheric Temperatures (TLT) . . . . .	59
A.3 The Mid-Tropospheric Temperature (TMT) . . . . .	60
A.4 The Bottom Line . . . . .	64
A.5 References . . . . .	65
Appendix B: Addressing criticisms of the UAH team . . . . .	66
B.1 Into context . . . . .	66
B.2 Summary . . . . .	72

Appendix C: MSU/AMSU diurnal adjustment . . . . . 77

## LIST OF FIGURES

Figure Number	Page	
1.1	Weighting functions for the various channels of the microwave sounding unit. The T24 weighting function is for the tropics. Weighting functions are provided by Remote Sensing Systems. . . . .	3
1.2	Zonal mean atmospheric temperature change from 1890 to 1999 ( $^{\circ}\text{C}$ per century) as simulated by the PCM model from (a) solar forcing, (b) volcanoes, (c) well-mixed greenhouse gases, (d) tropospheric and stratospheric ozone changes, (e) direct sulphate aerosol forcing and (f) the sum of all forcings. Plot is from 1,000 hPa to 10 hPa (shown on left scale) and from 0 km to 30 km (shown on right). Based on Santer et al. (2003a) (IPCC, 2007). . . . .	11
1.3	Theoretical temperature change throughout the troposphere for a one degree temperature rise at the surface for different thermodynamic conditions. If surface warming occurs in a dry adiabatic environment, there would be no amplification aloft relative to the surface. On the other hand, if warming were moist adiabatic, there would be pronounced warming in the upper troposphere. The moist adiabatic profile assumes 100% relative humidity at the surface. . . . .	13
1.4	Surface warming compared to tropospheric warming in the full troposphere in atmosphere-ocean coupled historical runs from Community Model Intercomparison Project 3 and observations for 1979 - 2000 and $20^{\circ}\text{S}$ - $20^{\circ}\text{N}$ . There are two sets of observations, because different surface datasets show varying degree of warming in the tropics. In this figure we used GISS as the lower bound for tropical surface warming and HadCRUT3v as the upper bound. Radiosonde datasets are denoted by colored open circles and MSU/AMSU datasets are denoted by colored open squares. The model ensemble means are all other symbols. Figure adapted from Santer et al. (2005). . . . .	14
2.1	Periods in which each satellite is incorporated in the homogenized TMT product for each team. . . . .	15
2.2	Red dots indicate locations of the radiosonde stations in the RICH and RAOBCORE products that were utilized in this work. . . . .	17
2.3	T24 time series for the three MSU/AMSU datasets averaged over the globe from 1979 - 2009. The reference period in which anomalies are calculated is 1995 - 2005. Time series are smoothed to reduce high frequency variations. . . . .	24

2.4	As in Figure 2.3, but for the tropics (30°S - 30°N).	25
2.5	T24 time series for the five radiosonde datasets averaged over the globe from 1979 - 2009. The IUK analysis only lasts until 2005 and the RATPAC-B homogenization effectively lasts until 1997 (afterwards radiosondes may contain biases). The reference period in which anomalies are calculated is 1995 - 2005. Time series are smoothed to reduce high frequency variations.	25
2.6	As in Figure 2.5, but for the tropics (30°S - 30°N).	26
2.7	T24 difference time series for the three MSU/AMSU datasets averaged over the globe from 1979 - 2009. Time series are smoothed to reduce high frequency variations.	26
2.8	As in Figure 2.7, but for the tropics (30°S - 30°N).	27
2.9	T24 difference time series for the three MSU/AMSU datasets (collocated with radiosondes) relative to the mean of the radiosonde datasets averaged over the globe from 1979 - 2009. Time series are smoothed to reduce high frequency variations.	27
2.10	As in Figure 2.9, but for the tropics (30°S - 30°N).	28
2.11	Global tropospheric trends for different deep layers. Synthetic satellite channels have been computed for various radiosonde products (solid squares). MSU/AMSU trends and statistical errors (95% confidence interval) have also been computed (open black markers). Further, the MSU/AMSU trends were calculated at grid points collocated with radiosonde observations for comparison. The collocated MSU/AMSU trends are denoted by open colored markers in which the color represents the radiosonde dataset that the MSU/AMSU dataset was collocated with and the marker type (diamond, circle, and X) denote the MSU/AMSU dataset. Trends are for 1979 - 2005 since the IUK dataset is only available through 2005.	30
2.12	As in Figure 2.11, but for the tropical region bound by 30° South - 30° North.	31
2.13	Global T24 trend map for a) NOAA, b) RSS, and c) UAH from 1979 - 2009.	32
2.14	Global T24 trend difference maps for 1979 - 2009.	33
3.1	a) TMT warm target factors used for different MSU teams. b) Satellites used in the RSS TMT merge (Mears and Wentz, 2009a). Note that the satellites used are different for the various MSU teams.	36
3.2	Scatter plot of MSU - MSU TMT difference series versus the warm target temperature over the NOAA-9 time period. m represents the slope for each of the relationships. The variance explained from the leftmost subplot to the rightmost subplot is 0.69, 0.81, and 0.25, respectively. The error is the 95% confidence interval in the linear fit.	37

3.3	Scatter plot of the mean of the five collocated <i>UAH – RADIOSONDE</i> difference series versus the warm target temperature. $\Delta\alpha = -slope$ because the $\alpha \cdot T_{TARGET}$ term is subtracted from the measured brightness temperature to obtain the calibrated UAH brightness temperature ( $T_{MSU}$ ). The error is the 95% confidence interval in the linear fit. . . . .	40
3.4	Estimate of the spatial impact of the UAH NOAA-9 warm target bias on UAH TMT trends (K decade <sup>-1</sup> ). . . . .	42
A.1	Global MSU-MSU TMT difference time series. . . . .	64
B.1	UAH team replicating our procedure using US Viz Sondes (Dr. John Christy, personal communication, 2012). The x-axis represents the target temperature anomaly and the y-axis is UAH-radiosonde. . . . .	69
B.2	Effects of sample size on statistics in this analysis. We computed the $r^2$ value, p-value, and $\Delta\alpha$ value for the regression of UAH - RICH Radiosondes versus $T_{Target}$ for different numbers stations (collocated with UAH data over the NOAA-9 time period). For this calculation we randomly sub-sample a certain number of stations (x-axis) and create an average, collocated UAH-Radiosonde difference time series, which we then regress against the warm target temperature. We redo this calculation 1,000 times for each number of stations sampled and then present the mean p-value, $r^2$ value, and $\Delta\alpha$ value. The shaded region around the $\Delta\alpha$ value is the 95% confidence interval from the sub-sampling statistics only; it does not contain the error in the regression itself. The results become significant when about 35 stations are included in the global average; below this number, the signal to noise ratio is too low. . . . .	70
C.1	Local equatorial crossing time (LECT) for the satellites utilized in the MSU/AMSU datasets. Note that some of the satellites, such as the satellite from 1995 - 2005, can drift by more than six hours. Since the satellites are quasi-sun-synchronous, the satellites pass the equator 12 hours apart on the ascending and descending node. So an LECT of “4” means the satellite crosses the equator at 4 AM and 4 PM local time. . . . .	78
C.2	Tropical diurnal correction used by RSS for land and ocean based on the National Center for Atmospheric Research Community Climate System Model version 3. NOAA utilizes the same correction, but scales it to reduce error residuals for overlapping satellites. . . . .	79
C.3	Estimated tropical (30°S - 30°N) diurnal corrections for each satellite. We’ve offset the corrections such that the mean difference between pairs of satellites approaches zero. . . . .	80
C.4	RSS - UAH TLT difference series in the tropics (30°S - 30°N) along with the scaled mean of the diurnal corrections estimated from Figure C.3. . . . .	81

C.5	Estimated sensitivity of the tropical diurnal cycle corrections to the phase and amplitude of the applied diurnal cycle. . . . .	82
-----	--	----

## LIST OF TABLES

Table Number	Page	
1.1	Current least squares linear trend values (1979 - 2011) for NOAA, RSS, and UAH for various channels in units of $\text{K decade}^{-1}$ . . . . .	6
2.1	Tropical and global trends for the datasets used in this work. Trends are calculated over 1979 - 2005 to accommodate the IUK dataset. Trend estimates are the least-squares linear fits in units of $\text{K decade}^{-1}$ . The confidence intervals are the 95% confidence intervals for the linear regression accounting for autocorrelation. NOAA does not provide a TLT product. . . . .	29
3.1	Correlation coefficients for MSU (column) - REFERENCE (row) versus the global mean warm target temperature for NOAA-9 during January 1985 to February 1987. Values denoted by an asterisk are significant with 95% confidence. The “Radiosonde Mean” is the correlation coefficient of the mean of the five UAH - REFERENCE time series versus the global warm target temperature for NOAA-9. . . . .	39
3.2	$\Delta\alpha_9$ values and 95% confidence intervals derived from our least-squares linear fit. These values are the magnitude of the slope of the linear relationship between UAH (column) - REFERENCE (row) versus the global mean warm target temperature over the NOAA-9 operational period. This value should be subtracted from $\alpha_9$ to correct for the non-optimal selection of warm target factor. The “Radiosonde Mean” is the same as that from Table 1 and Figure 2. . . . .	41
A.1	MSU - Radiosonde error characteristics presented by the UAH team. . . . .	63
B.1	TMT error characteristics for the different MSU groups compared to the radiosonde references used in this work. For this comparison we used detrended, collocated time series from 1979 - 2009 and averaged the results for the five radiosonde datasets. . . . .	68

## ACKNOWLEDGMENTS

I would like to thank those that made this work possible through their unending support of me, personally and professionally. This includes my family, friends, and colleagues.

This work was supported by the National Science Foundation Graduate Research Fellowship (DGE-0718124), NOAA Grant NA08OAR4310725, and NESDIS-NESDISPO-2009-2001589 (SDS-09-15).

## Chapter 1

## INTRODUCTION

**1.1 History of the microwave sounding unit discrepancies**

A great deal of attention has been focused on satellite-derived temperature trends in the troposphere. This attention was largely a result of work by Spencer and Christy (1990). In their groundbreaking effort, the team found no warming trend throughout the 1980s in passive microwave observations, even though large temperature trends at the surface existed. In the late 1990s and early 2000s, radiosonde and satellite-derived measurements of temperature changes in the lower troposphere showed little warming or even cooling (Hurrell et al., 2000), leading some scientists to question the validity of the global surface station network that measured relatively large and positive temperature trends at the Earth’s surface (NRC, 2000). A report by the Panel on Reconciling Temperature Observations released a National Academies Report in 2000 that concluded that the surface warming “during the last 20 years is undoubtedly real and is substantially greater than the average rate of warming during the twentieth century. The disparity between surface and upper air trends in no way invalidates the conclusion that surface temperature has been rising” (NRC, 2000). The panel went on to explain that algorithm changes in the satellite-derived measurements helped reconcile some of the surface versus upper-air differences and that natural variability and ozone depletion might lead to differential trends between the surface and troposphere. In this work, we will focus on the remaining discrepancies related to the satellite observations.

In the early 2000s, efforts were being made to better understand biases in satellite observations (e.g. Christy et al., 1998, 2000). Until 2003, the only climate quality satellite dataset of tropospheric temperatures was a product of the University of Alabama Huntsville (UAH) Team, which used observations from a number of polar orbiting satellites that carried the microwave sounding unit (MSU) (Spencer and Christy, 1990, 1992a,b; Christy et al., 1998, 2000, 2003). With the launch of NOAA-15 in 1998, the MSU was succeeded by the ad-

vanced microwave sounding unit (AMSU). The MSU and AMSU are instruments flown on National Oceanic and Atmospheric Administration (NOAA) polar orbiting satellites as well as the National Aeronautics and Space Administration’s (NASA) AQUA satellite and the European Organisation for the Exploitation of Meteorological Satellite’s (EUMETSAT) MetOp-A satellite. These instruments are radiometers originally designed to aid in weather prediction via passive microwave observations of temperature, but have been utilized for climate applications in part because of their global coverage and long record (over a number of satellites). The passive microwave observations from these satellites measure the temperature of deep atmospheric layers. The mid-tropospheric channel (referred to as TMT or T2)<sup>1</sup> of the MSU detects non-negligible emissions from the stratosphere, so the UAH team linearly combined different view angles from TMT to create a lower tropospheric channel free of this contamination (referred to as TLT) (Spencer and Christy, 1992b). An alternative measure of tropospheric temperature is derived using measurements from MSU channel 2 (TMT) and channel 4 (lower stratosphere - TLS). This channel, referred to as T24, has more weight in the mid- and upper-troposphere and will be utilized throughout this work (Fu et al., 2004). The weighting functions for the bulk atmospheric layers used in this study are given in Figure 1.1.

---

<sup>1</sup>In Christy et al. (2003), the authors redefine T2 observations as TMT, because they begin to incorporate AMSU data into the analysis and the near-equivalent AMSU channel is channel 5. Throughout this work, we broadly define MSU channel 2 or AMSU channel 5 observations as TMT, referring to these channels. We similarly utilize TMT when referring to the mid-tropospheric product (combining MSU and AMSU observations), which is technically a composite of the near nadir Earth views of MSU channel 2 and AMSU channel 5. We also refer to TLT and T24 as “channels,” although this is not strictly true; T24 is a linear combination of MSU channels 2 and 4 and TLT is a linear combination of different TMT view angles.

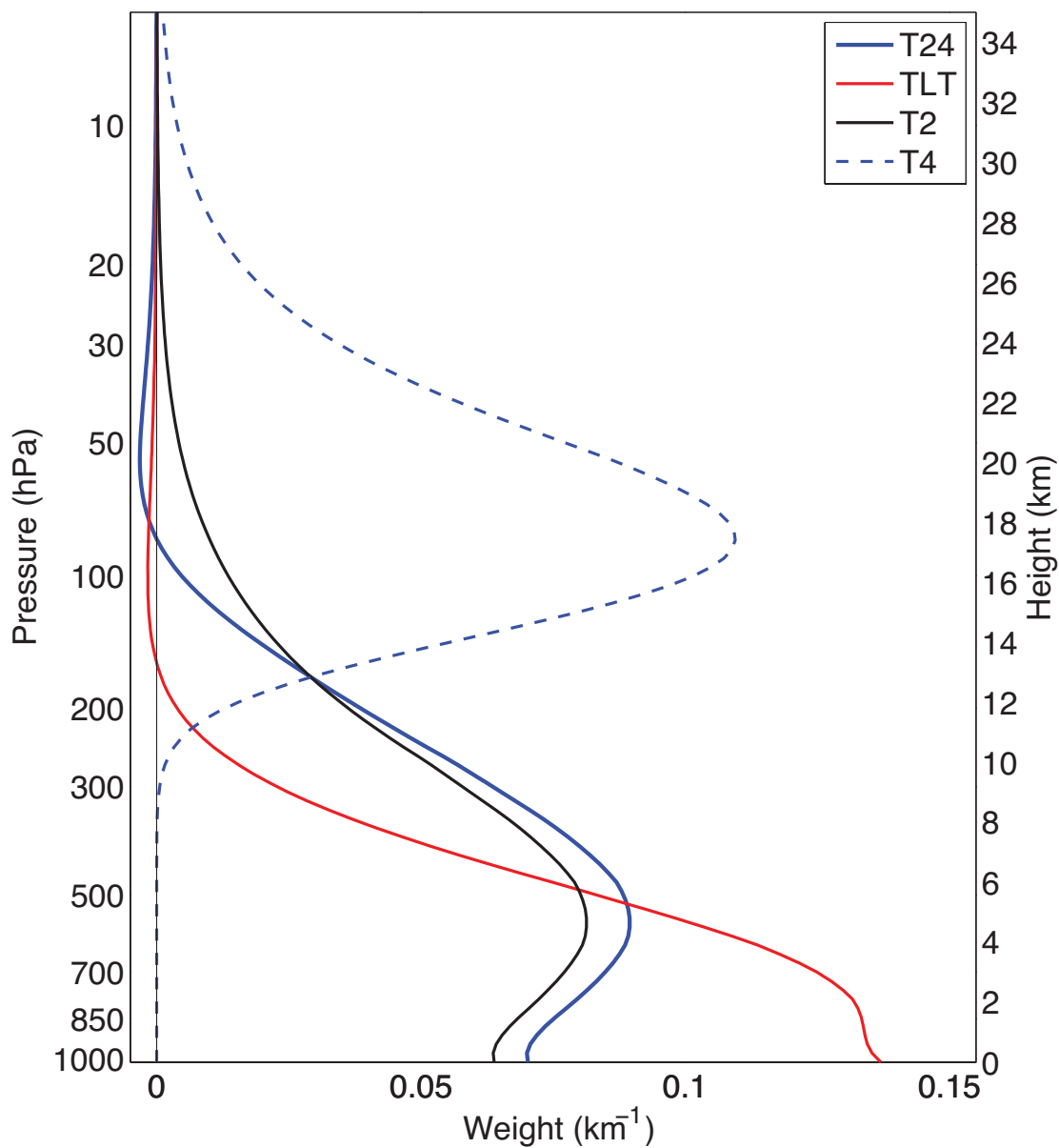


Figure 1.1: Weighting functions for the various channels of the microwave sounding unit. The T24 weighting function is for the tropics. Weighting functions are provided by Remote Sensing Systems.

Important biases were discovered in the TLT product, which led to large positive adjustments to the UAH temperature trends including biases due to the decay of satellite orbits (TLT increased 0.07 - 0.10 K decade<sup>-1</sup>) (Wentz and Schabel, 1998; Christy et al., 1998) and a bias in the treatment of a diurnal cycle drift correction for the NOAA-11 satellite (TLT increased 0.035 K decade<sup>-1</sup>) (Fu and Johanson, 2005; Mears and Wentz, 2005; Christy and Spencer, 2005a). In the former correction, satellites lost altitude due to increases in friction associated with solar events, which introduced a spurious decline in TLT temperatures. The loss of altitude led to an increase in the temperatures measured in the limb views of the satellite, without a compensating increase in the near-nadir views of the satellite. TLT is a linear combination of the view angles such that:

$$TLT = 4T_N - 3T_L \quad (1.1)$$

where  $T_N$  is the near-nadir view angles and  $T_L$  are the limb views of the sounder. As a result of the altitude changes (which caused an increase in temperatures in the limb view with altitude loss), the TLT suffered from large spurious cooling (Wentz and Schabel, 1998). In 2003, Remote Sensing Systems (RSS) released an alternative MSU TMT temperature dataset that reported considerably more warming ( $\sim 0.09$  K decade<sup>-1</sup>) compared to the UAH dataset (Mears et al., 2003; Mears and Wentz, 2009a; Mears et al., 2011), but did not produce a complementary TLT product at this time.

In 2004, an alternative tropospheric measurement was developed utilizing measurements from TMT and the lower stratospheric channel (TLS or T4) of the MSU, referred to as T24 (Fu et al., 2004). This full tropospheric temperature measurement removed stratospheric contamination from TMT, demonstrated that UAH TLT likely had undiscovered biases because of the inconsistency between UAH T24 and UAH TLT (TLT trends were negative, even though T24 trends were positive), and that the RSS T24 measurement was in excellent agreement with global circulation models. This served as an important contribution at a time when upper air datasets showed cooling or only slight warming in the troposphere.

Later, in 2005, RSS released a TLT product that showed warming consistent with the T24 measurement. RSS TLT also contained considerably more warming relative to UAH TLT and highlighted an error with the UAH TLT product, which was corrected the same

year (Fu and Johanson, 2005; Mears and Wentz, 2005; Christy and Spencer, 2005b; Mears and Wentz, 2009b). The 2005 UAH correction was based on the diurnal drift of the satellite. As the polar orbiting satellites age, they suffer from east-west orbital drift, which changes the local sampling time. As a result, the long-term temperature measurements are confounded by changes in the sampling of the diurnal cycle. RSS applies a diurnal cycle correction based on the diurnal cycle of the Community Climate System Model 3.0 (Mears et al., 2003) whereas UAH produces a TMT diurnal correction based on differences in the eastward and westward view angle of the satellite; the cross-swath difference in sampling is about 80 minutes at the equator (Christy et al., 2000). For the TLT channel, UAH “uses a regression-derived diurnal correction based on 1 year of data from co-orbiting advanced MSU (AMSU) satellites that measure a total of six nominal local satellite overpass times. Three assumed diurnal functions are fitted by regression to the grid point, monthly averaged brightness temperatures for AMSU channels 3 through 10. The three functions include a 24-hour trace of the solar flux, the time integrated solar flux, and a linear term representing infrared cooling” (Randall and Herman, 2008). In the 2005 correction, UAH had applied a diurnal drift correction of the wrong sign to the NOAA-11 satellite for TLT, which artificially reduced the UAH TLT trend.

After these corrections, most of the global trend inconsistencies between satellite and surface measurements had been reconciled and more attention was focused on the ratio of troposphere to surface warming in models versus observations, particularly in the tropics (e.g. Karl et al., 2006; IPCC, 2007; Douglass et al., 2008; Santer et al., 2008; Christy et al., 2010; Fu et al., 2011; Thorne et al., 2011b; Klotzbach et al., 2009, 2010). In Table 1.1, we present the current linear temperature trends for NOAA, RSS, and UAH updated through 2011. We see that the tropospheric trends (TLT, T24, and TMT with some stratospheric contamination) are all positive, which is consistent with surface warming. But global TMT and T24 trends for UAH are very different compared to NOAA and RSS; this work helps reconcile these longstanding differences. Furthermore, UAH has reduced tropical trends for all of the tropospheric trend values relative to RSS and NOAA. While we do not reconcile this important difference here, we propose lines of research that may help address this problem in Appendix C.

Table 1.1: Current least squares linear trend values (1979 - 2011) for NOAA, RSS, and UAH for various channels in units of K decade<sup>-1</sup>.

Channel	Region	NOAA	RSS	UAH
TLT	Global	N/A	0.139	0.137
	Tropical	N/A	0.125	0.072
T24	Global	0.197	0.141	0.113
	Tropical	0.177	0.138	0.075
TMT	Global	0.130	0.083	0.048
	Tropical	0.131	0.101	0.040
TLS	Global	-0.322	-0.302	-0.382
	Tropical	-0.324	-0.264	-0.313

In 2006, the CCSP Report concluded that:

“Previously reported discrepancies between the amount of warming near the surface and higher in the atmosphere have been used to challenge the reliability of climate models and the reality of human induced global warming. Specifically, surface data showed substantial global-average warming, while early versions of satellite and radiosonde data showed little or no warming above the surface. This significant discrepancy no longer exists because errors in the satellite and radiosonde data have been identified and corrected. New data sets have also been developed that do not show such discrepancies.”

The report goes on to explain that:

“In the tropics, the agreement between models and observations depends on the time scale considered. For month-to-month and year-to-year variations, models and observations both show amplification (i.e., the month-to-month and year-to-year variations are larger aloft than at the surface). This is a consequence of relatively simple physics, the effects of the release of latent heat as air rises

and condenses in clouds. The magnitude of this amplification is very similar in models and observations. On decadal and longer time scales, however, while almost all model simulations show greater warming aloft (reflecting the same physical processes that operate on the monthly and annual time scales), most observations show greater warming at the surface.

These results could arise either because “real world” amplification effects on short and long time scales are controlled by different physical mechanisms, and models fail to capture such behavior; or because non-climatic influences remaining in some or all of the observed tropospheric data sets lead to biased long-term trends; or a combination of these factors. The new evidence in this Report favors the second explanation.”

The procedure for merging individual satellites (currently 15 are utilized) into a continuous, climate-quality lower and mid-tropospheric time series is quite complicated and must account for the aforementioned non-climatic biases such as the diurnal cycle drift effect (e.g. Trenberth and Hurrell, 1997; Spencer and Christy, 1992a,b; Christy et al., 1998, 2000; Fu and Johanson, 2005; Mears and Wentz, 2005), biases due to the decay of satellite orbits (e.g. Wentz and Schabel, 1998), and the influence of the instrument body temperature on the measured radiance (e.g. Christy et al., 1998, 2000; Prabhakara et al., 2000; Mears et al., 2003; Zou et al., 2006). Other biases such as frequency shifts in the radiometer passband (e.g. Zou and Wang, 2011; Lu et al., 2011) and the transition from MSU to AMSU data may also be important (e.g. Mears and Wentz, 2009a).

There are compelling reasons to believe that some of the remaining discrepancies are related to the non-linear calibration of the satellite. Christy et al. (2000) and Mo (1994) showed that the non-linear calibration coefficients used in the temperature retrievals can vary from the pre-launch calibration. RSS and UAH use a linear correction factor based on that target calibration temperature of the satellite to account for perturbations from the pre-launch calibration coefficients. This calibration procedure is used to address “the instrument body effect,” because the changes in the non-linear calibration equation after launch produce Earth brightness temperature residuals related to the instrument body

temperature if left uncorrected. In 2006, NOAA Center for Satellite Applications and Research (STAR) developed yet another independent MSU/AMSU temperature dataset for TMT and TLS (but not TLT) (Zou et al., 2006, 2009; Zou and Wang, 2011). The NOAA STAR dataset differs from UAH and RSS because it determines satellite post-launch calibration coefficients from simultaneous nadir overpasses (SNO) of co-orbiting satellites and subsequently updates the level one data using the derived coefficients. Calibration using this method reduces the differences between five-day (pentad) grid point averages of overlapping satellites by an order of magnitude. Further, NOAA utilizes the RSS diurnal cycle correction, but applies a scaling factor to correct for biases in the climate model's diurnal cycle (Zou and Wang, 2009). The NOAA STAR TMT dataset has larger global trends than both RSS and UAH.

Differences in the removal of the instrument body temperature effect for NOAA-9 has been highlighted as one of the most important differences between UAH and RSS, accounting for  $\geq 50\%$  of the TMT trend difference (Mears et al., 2003; Karl et al., 2006). The impact of the different methods to remove the instrument body effect has been considered as a part of the structural uncertainty that results from the application of a reasonable set of processing choices (Thorne et al., 2005a; Karl et al., 2006). Christy and Norris (2004) compared the UAH TLT record with radiosondes before and after the NOAA-9 time period and found no significant differences, leading them to conclude that there was no bias in their treatment of NOAA-9 for TLT or TMT. This study assumed that radiosondes could be utilized as global references over a decade and then assume that the results they found for TLT (UAH had greater consistency with radiosondes) apply to TMT because of the similarity in the merging procedure for both products. We come to a different conclusion with a physically based test of the calibration in question (for the TMT product directly). Using radiosonde measurements as a reference, we show that the treatment of the instrument body temperature effect (i.e. warm target calibration) by UAH has a significant bias for the NOAA-9 satellite (Po-Chedley and Fu, 2012). The datasets used in this work are briefly discussed in Chapter 2 and our methods are developed in Chapter 3. For completeness, we have also included criticisms of our work from the UAH team (Appendix A) and some information that addresses their concerns (Appendix B). Removing this bias helps to

reconcile trend differences between UAH and RSS/NOAA.

Another important difference between RSS and UAH occurs at the introduction of NOAA-12 data around 1992. During this time, the RSS TLT tropical temperature has a step-like increase relative to UAH (Christy et al., 2007). Christy et al. (2007) suggests that the step like change is a result of a spurious drift in the latter half of the NOAA-11 time series, which leads to a sudden increase in brightness temperature when NOAA-11 is combined with NOAA-12 in late 1991. The same study demonstrates that this jump is evident for RSS when using radiosondes as references, but there is no significant jump for the UAH dataset. Randall and Herman (2008) compare short (five and ten-year) trend differences between MSU/AMSU observations and radiosondes to suggest that the RSS TLT diurnal correction for NOAA-11 is too large. Other comparisons with radiosondes and surface datasets come to similar conclusions about a residual bias in the RSS TLT dataset during this time period, some of which indicate that the RSS diurnal drift correction is too large (e.g. Christy and Norris, 2006; Christy et al., 2010, 2011; Bengtsson and Hodges, 2011). It has been pointed out that these analyses have focused on the time period after 1992 when RSS differs from radiosondes, but ignored other time periods when other substantial MSU/AMSU-radiosonde disagreements were present (Mears and Wentz, 2012). Further, since both radiosonde and MSU/AMSU observations may contain time varying biases, it is not possible to draw conclusions about the validity of a dataset based on simple comparisons, especially when the internal uncertainty of both datasets is not accounted for (Mears et al., 2011; Mears and Wentz, 2012). Unfortunately, substantial long-term errors may remain in homogenized radiosonde data (Lanzante et al., 2003a; Titchner et al., 2009; Mears and Wentz, 2012), especially in the tropics (Randel and Wu, 2006; Sherwood et al., 2005, 2008).

We briefly discuss the role of the diurnal drift correction in reconciling MSU/AMSU trends and actions that may be taken to help determine biases in the implementation in the diurnal drift correction. This information is included in Appendix C.

In summary, there exists three independent and up-to-date MSU/AMSU datasets (UAH, RSS, and NOAA). RSS and UAH utilize the Christy et al. (2000) correction for changes in the satellite calibration coefficient whereas NOAA uses the SNO method (and then the

Christy correction). Because of slight differences in these calibration procedures, each group obtains different calibration coefficients, especially for the NOAA-9 satellite. NOAA and RSS use a similar diurnal correction based on a climate model, whereas UAH uses cross-scan differences to account for the diurnal cycle in TMT and information from co-orbiting AMSU satellites for TLT. As a result of differences in the merging procedure for the various MSU teams, global TMT trends from 1979 - 2009 differ by more than a factor of three. The global TMT trends for 1979 - 2009 are  $0.127 \text{ K decade}^{-1}$ ,  $0.080 \text{ K decade}^{-1}$ , and  $0.038 \text{ K decade}^{-1}$  for NOAA, RSS, and UAH, respectively. Understanding the discrepancies in TMT trends among various teams is critically important to reliably derive tropospheric temperature trends based on satellite-borne MSU/AMSU observations (e.g. Karl et al., 2006). Although the TMT product suffers from contamination from the stratosphere (Fu et al., 2004), it is utilized for deriving upper (T24) and lower (TLT) tropospheric warming.

## 1.2 The expectation of surface warming and temperature amplification

In some of the earliest attempts to model the Earth's climate system, Manabe and Wetherald (1967) demonstrated that increasing carbon dioxide has the effect of cooling the stratosphere and warming the surface and troposphere with enhanced warming in the tropical upper troposphere (Manabe and Wetherald, 1975; Thorne et al., 2011a). This has become a robust feature of global circulation models (GCMs) (IPCC, 2007). Figure 1.2 demonstrates the impact of various forcings on the observed change over the 20th century.

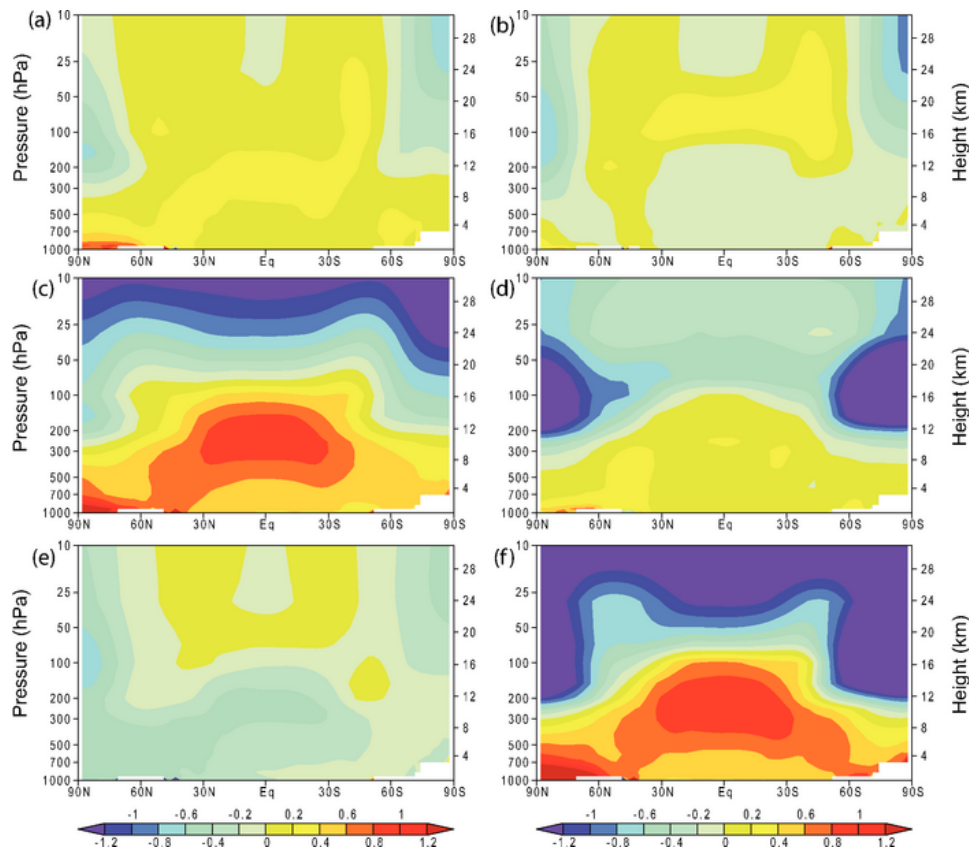


Figure 1.2: Zonal mean atmospheric temperature change from 1890 to 1999 ( $^{\circ}\text{C}$  per century) as simulated by the PCM model from (a) solar forcing, (b) volcanoes, (c) well-mixed greenhouse gases, (d) tropospheric and stratospheric ozone changes, (e) direct sulphate aerosol forcing and (f) the sum of all forcings. Plot is from 1,000 hPa to 10 hPa (shown on left scale) and from 0 km to 30 km (shown on right). Based on Santer et al. (2003a) (IPCC, 2007).

The tropical troposphere maintains a horizontally homogeneous temperature field, as a result of the small Coriolis force (e.g. Bretherton and Sobel, 2003). Heating in the free troposphere is quickly spread throughout the tropics via internal gravity waves, which ensure that the deep tropics maintain weak temperature gradients (Bretherton and Sobel, 2003). The temperature in the tropical troposphere roughly follows a moist adiabatic profile that is tied to the mean tropical sea surface temperature with some fixed relative humidity (SST) (Stone and Carlson, 1979). More precisely, convective adjustment only acts in regions of high precipitation, which means that the tropical tropospheric temperatures will be locked to the SST in convective regions via a moist adiabatic relationship on timescales greater than  $\sim 10$  days (but not necessarily the tropical mean SST) (Sobel et al., 2002). Under anthropogenic climate change simulations in the tropics, increased warming occurs aloft because the warming follows the moist adiabatic lapse rate (MALR). In Figure 1.3, we show the warming from MALR theory (Stone and Carlson, 1979). As latent heat is released, there is pronounced warming in the tropical upper troposphere (near  $\sim 200$ hPa). Models and observations robustly show this relationship on interannual timescales, but not all observations show amplified warming relative to the surface on decadal timescales (Santer et al., 2005). We demonstrate the decadal scaling relationship for models and observations in Figure 1.4. In Figure 1.4, we utilize a synthetic satellite channel, T24, as a measure of the full tropospheric warming in observations and models. If warming were perfectly moist adiabatic and locked to the surface temperature rise, we would expect T24 to warm  $\sim 1.6$  K for every 1 K surface temperature rise.

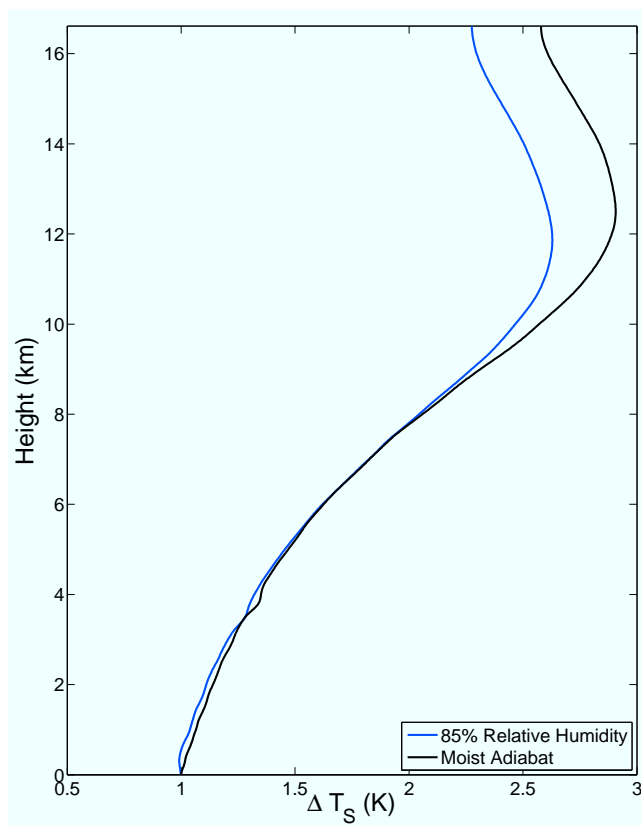


Figure 1.3: Theoretical temperature change throughout the troposphere for a one degree temperature rise at the surface for different thermodynamic conditions. If surface warming occurs in a dry adiabatic environment, there would be no amplification aloft relative to the surface. On the other hand, if warming were moist adiabatic, there would be pronounced warming in the upper troposphere. The moist adiabatic profile assumes 100% relative humidity at the surface.

From theory and GCM results, we expect the troposphere to warm under anthropogenic climate change, with enhanced warming in the tropical troposphere. At some level in the tropical troposphere, the MALR no longer holds and so the amplified warming from the moist adiabatic relationship is not necessarily expected to the tropical tropopause. Recent studies have shown that GCMs may over-predict the warming in the tropical upper troposphere (Bengtsson and Hodges, 2011; Fu et al., 2011).

Over the last decade, not all observational studies have demonstrated tropical tropospheric warming and amplification on decadal time scales, even though amplification has

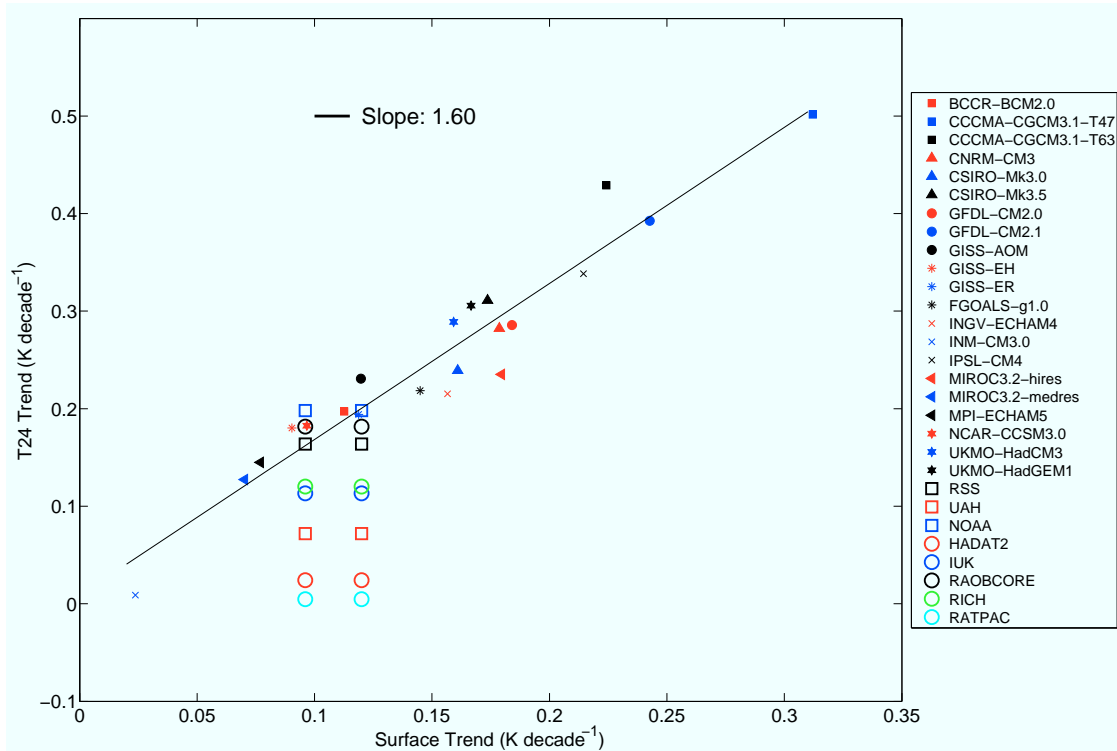


Figure 1.4: Surface warming compared to tropospheric warming in the full troposphere in atmosphere-ocean coupled historical runs from Community Model Intercomparison Project 3 and observations for 1979 - 2000 and 20°S - 20°N. There are two sets of observations, because different surface datasets show varying degree of warming in the tropics. In this figure we used GISS as the lower bound for tropical surface warming and HadCRUT3v as the upper bound. Radiosonde datasets are denoted by colored open circles and MSU/AMSU datasets are denoted by colored open squares. The model ensemble means are all other symbols. Figure adapted from Santer et al. (2005).

been considered a “fingerprint” of anthropogenic global warming (e.g. Santer et al., 2003b; Allen and Sherwood, 2008). A number of these studies considered the role of climate variability and generally concluded that there are either remaining residual errors in the satellite records that effect long-term trends or that there are different physical processes that effect tropospheric temperature on interannual and decadal timescales (Gaffen et al., 2000; Santer et al., 2000; Hegerl and Wallace, 2002; Santer et al., 2003b, 2005). For these reasons, we are trying to diagnose residual errors in the satellite records.

## Chapter 2

## MSU AND RADIOSONDE DATASETS

We focus the bulk of our analysis on monthly TMT data during the NOAA-9 time period (January 1985 - February 1987). In our comparison, we utilize UAH T2 (i.e. TMT) V5.3, RSS TMT V3.3, and NOAA TMT V2.0 (Christy et al., 2003; Mears and Wentz, 2009a; Zou et al., 2006). For each gridded dataset we calculate anomalies based on a common reference period (1995 - 2005) to facilitate comparison. The satellite record is composed of a number of individual satellites that have been homogenized into one continuous time series. Figure 2.1 shows the periods in which each team utilizes each of the individual satellites for the TMT record.

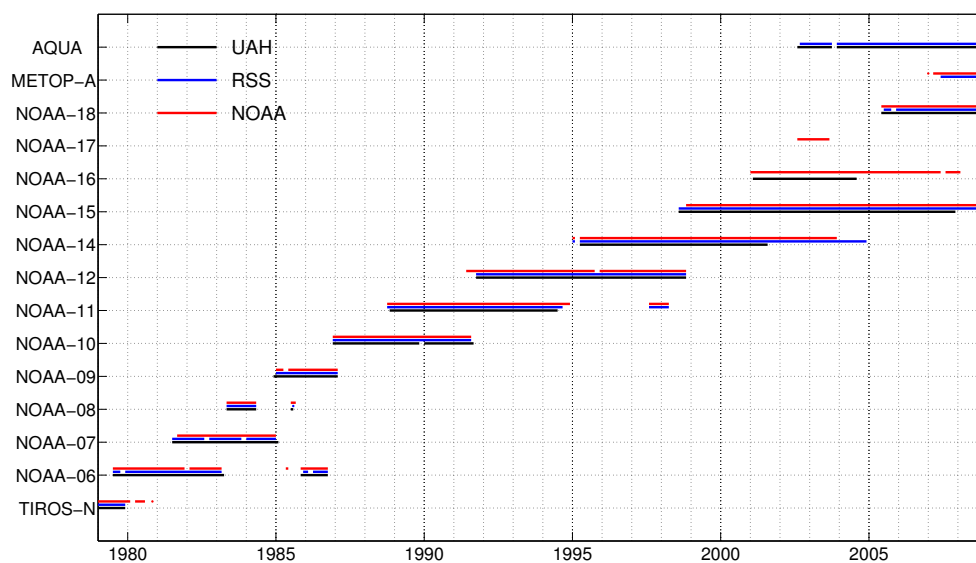


Figure 2.1: Periods in which each satellite is incorporated in the homogenized TMT product for each team.

Radiosondes will be used as an independent reference for the MSU/AMSU-derived datasets. We use station data from five global, homogenized, monthly-mean radiosonde datasets, including RICH, RAOBCORE V1.4, HadAT2, RATPAC-B, and IUK A and B (Dr. L. Haimberger, personal communication, 2011; Haimberger, 2007; Haimberger et al., 2008; Thorne et al., 2005b; McCarthy et al., 2008; Free et al., 2005; Sherwood, 2007). We used these five datasets to demonstrate that our results are robust regardless of the radiosonde dataset used, even though each dataset was formed with different homogenization techniques and different sets of radiosonde observations. A brief summary of the radiosonde datasets are provided in this chapter.

For the radiosonde data, we use the same reference period (1995 - 2005) to derive anomalies and apply a weighting function from RSS to monthly mean data to produce synthetic MSU brightness temperatures. The synthetic MSU brightness temperature is given by Equation 2.1:

$$T_B = \int_{p_S}^0 w(p) \cdot T(p) dp + \epsilon_S \cdot w_S \cdot T_S \quad (2.1)$$

where

$$w_S = 1 - \int_{p_S}^0 w(p) \cdot dp, \quad (2.2)$$

$T_B$  represents the synthetic satellite channel brightness temperature,  $\epsilon_S$  is the surface emissivity (we used 0.9 for land and 0.5 for ocean),  $p$  is pressure,  $T$  is the temperature as a function of height, and  $w$  is the weight at each level. The “S” subscript denotes the surface level. The radiosonde temperature profile was interpolated to the finer scale weighting function height levels.

For each radiosonde station within each product we required that at least 90% of the time series (1979 - 2009) be available with enough data to fill in at least 85% of the weighting function in order to produce a time series for that station. As a result of these constraints, only a subset of each radiosonde dataset was used. In this study, we used 384 of 2,881 radiosonde stations for RICH and RAOBCORE V1.4, 219 of 676 stations for HadAT2, 45 of 85 stations for RATPAC-B, and 451 of 527 stations for IUK A and B. If enough data existed for both the 00Z and 12Z time series, the anomalies for each were averaged together to form a single time series for each station. For each radiosonde dataset, MSU data was

temporally and spatially collocated with radiosonde station data to form corresponding MSU time series. From Figure 2.2, we see that radiosonde sampling is not globally uniform and can be quite poor in some latitude bands.

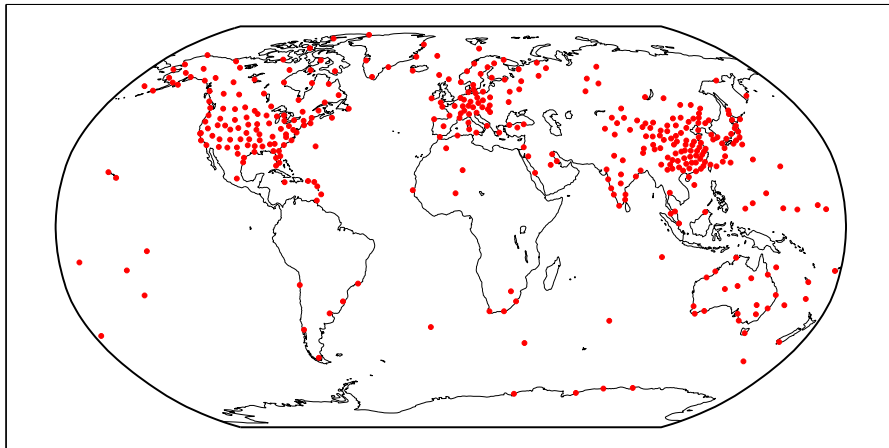


Figure 2.2: Red dots indicate locations of the radiosonde stations in the RICH and RAOB-CORE products that were utilized in this work.

Microwave radiation contributions from the surface can be important for microwave sounding measurements. Since radiosonde datasets do not provide a surface skin temperature measurement, we tried several methods to estimate surface temperature (which will serve as our skin temperature). Some studies (e.g Randall and Herman, 2008) utilize the 1000 hPa measurement in place of the skin temperature. Others utilize measurements from global surface temperature datasets, such as the HadCRUT dataset.

Since not all datasets provided us with a 1000 hPa measurement and because we are utilizing individual station data and not gridded products for all comparisons, we utilized a different technique. To compute synthetic satellite channels, we extrapolated the surface temperature assuming a linear lapse rate using available temperature measurements in the

lower troposphere. Further, if the Earth’s surface at a given station is above mean sea level, we added additional weight to the surface temperature measurement in our brightness temperature calculation (Equation 2.2). This would have the most pronounced effect on the TLT channel, but the global trend difference for TLT between 1) our lapse rate assumption for the surface temperature and 2) a surface temperature measurement utilizing surface datasets [GHCN (Jones and Moberg, 2003), HadCRUT3v (Brohan et al., 2006), and GISS (Hansen et al., 2010)] for the surface temperature is less than  $0.02 \text{ K decade}^{-1}$ . For all of the radiosonde datasets, we assumed that the anomaly at the top of the atmosphere was zero.

Channel T24, the full tropospheric channel, was designed to remove stratospheric contamination from the TMT channel. It is a linear combination of MSU channel 2 (TMT) and channel 4 (TLS) and is given by Fu et al. (2004):

$$T24 = a_{TMT} \cdot TMT + a_{TLS} \cdot TLS \quad (2.3)$$

which effectively subtracts off the stratospheric influence in TMT using the TLS channel. In the tropics  $a_{TMT} = 1.1$  and  $a_{TLS} = -0.1$ . Since radiosondes have large biases in the stratosphere (e.g Randel and Wu, 2006), we calculate T24 from radiosondes following Johanson and Fu (2006):

$$T_B = \frac{\int_{p_S}^{p_T} w(p) \cdot T(p) dp + \epsilon_S \cdot w_S \cdot T_S}{\int_{p_S}^{p_T} w(p) + w_S} \quad (2.4)$$

where the tropopause height (tropopause is denoted with the “T” subscript) is derived from the annual average NCEP/NCAR reanalysis tropopause height at each latitude (Kalnay et al., 1996). In this case, the weighting function is for the TMT channel. For the MSU/AMSU measurements the  $a_{TMT}$  and  $a_{TLS}$  coefficients are a function of latitude, accounting for tropopause height (Celeste Johanson, personal communication, 2010; Fu and Johanson, 2004; Johanson and Fu, 2006). In the global average  $a_{TMT} = 1.15$  and  $a_{TLS} = -0.15$ .

We had considered using SSM/I water vapor data as a proxy for tropospheric temperature following the work of Wentz and Schabel (2000), but decided that water vapor was not a robust enough reference for our purposes. Wentz and Schabel (2000) find that water vapor

scales as 6.7% per K in the deep tropics (20°S - 20°N) and found that water vapor column (WVC) anomaly correlated very well with the UAH TLT product (correlation coefficient exceeded 0.9). In trying to repeat their finding, we found TLT versus WVC had a much reduced correlation coefficient ( $\sim 0.6$ ), unless we smoothed the data as had been done by Wentz and Schabel (2000). Typical global radiosonde/MSU correlations (unsmoothed) have correlation coefficients exceeding 0.9. Furthermore, the MSU/AMSU minus WVC-derived temperature (using the 6.7%  $\text{K}^{-1}$  scaling in the deep tropics) had a standard deviation roughly three times larger than that of MSU/AMSU minus radiosondes. Since monthly anomalies are important to our analysis, we elected not to use WVC, although it might be a useful metric on interannual or decadal timescales.

## **2.1 Radiosondes**

### *2.1.1 HadAT2*

The HadAT2 dataset was produced by Thorne et al. (2005b) and has been maintained as an up-to-date homogenized, global radiosonde dataset. HadAT2 contains 676 stations with 9 mandatory reporting levels from 850 hPa to 30 hPa from a number of raw radiosonde datasets. The product utilizes “neighbor buddy checks” to find and correct inhomogeneities at individual stations.

It was noted that stations in the tropics utilize nearly all tropical stations for homogenization since the tropical atmosphere is spatially homogenous. Importantly, if systematic regional biases exist, they may not be detected and may remain in the regional time series. This potential bias is especially important in the tropics, which has fewer stations relative to the northern mid-latitudes.

### *2.1.2 IUK A and B*

IUK A and B is a dataset that was produced by Sherwood et al. (2008), following up on previous work that recommended the use of a technique called iterative universal kriging (Sherwood, 2007). The dataset differs from others in that it employs wind shear data, which has few time dependent inhomogeneities, to help determine step-wise changes in tempera-

ture. The iterative procedure also uses differences between day and night observations to detect change points. The dataset consists of Group A (460 stations) and Group B (67 stations), with Group B being a set of stations that are more difficult to homogenize because of a lack of 00 UTC and 12 UTC data. In the tropics, about one-quarter of the stations are in the B Group compared to 10% in the northern hemisphere. The product contains 10 pressure levels from 850 hPa to 30 hPa and the record spans through 2005.

Sherwood et al. (2008) suggest that “artifacts remain in the troposphere in some of the stations from 5°S - 20°N, because trends there are too low compared to those at other latitudes.” For this work we utilize the data from both groups A and B.

### *2.1.3 RICH and RAOBCORE V1.4*

Haimberger (2007) have utilized the differences between the background forecasts of ECMWF Re-Analysis minus the original radiosonde observations (the difference is referred to as an “innovation”) to detect and adjust radiosonde inhomogeneities. This analysis method is known as RAOBCORE. Since the ECMWF analysis utilizes MSU/AMSU measurements, it is not completely independent of the satellite data utilized in this work. Unlike the IUK dataset, day minus night differences are not utilized to detect change points, but the analysis is successful in eliminating biases related to the daytime heating of radiosondes. A new version, RAOBCORE V1.5, was released since this work began, but we were not able to consider the new data.

Haimberger et al. (2008) and Haimberger et al. (2011) also utilize neighbor station comparisons to make adjustments independently of the ECMWF reanalysis innovations, which can be effected by the radiosonde biases themselves. This analysis is known as RICH.

Both RICH and RAOBCORE provide 16 levels from 1000 hPa to 10 hPa, though many stations lack data throughout the stratosphere and are not continuous throughout the period considered here (1979 - 2009).

#### 2.1.4 *RATPAC-B*

Free et al. (2005) utilized the LKS dataset (Lanzante et al., 2003a,b) and append it using uncorrected IGRA radiosonde network data. The LKS dataset utilized metadata records, statistical analysis, and expert review by a panel of scientists to homogenize a network of 87 radiosondes through 1997. This dataset is essentially uncorrected after 1997, but the alternative dataset, RATPAC-A, is meant for large-scale averages. Since we are collocating MSU and radiosonde data in this analysis, we utilize RATPAC-B, which has data for individual stations. Because we focus on the NOAA-9 period for much of this work (1985 - 1987), this makes RATPAC-B a useful comparison dataset.

The RATPAC-B dataset includes a surface temperature measurement and 12 pressure levels from 850 hPa to 30 hPa.

## 2.2 MSU/AMSU observations

### 2.2.1 University of Alabama, Huntsville (UAH)

The University of Alabama, Huntsville (UAH) dataset was the original climate record from the microwave sounding unit and has undergone a number of version changes since its establishment in 1990 (Spencer and Christy, 1990, 1992a,b; Christy et al., 1995, 1998, 2000, 2003). UAH produces homogenized satellite products for the following channels: TLT (originally T2LT), TMT (MSU channel 2), and TLS (MSU channel 4). We also calculate a T24 product from channels TMT and TLS.

One of the main differences between UAH and NOAA/RSS is that UAH utilizes what it calls a “backbone” merging procedure. UAH selects a set of core reference satellites that are adjusted relative to one another iteratively. For example, in the 1990s, UAH calculates the NOAA-11 biases relative to NOAA-10 using the NOAA-10/NOAA-11 overlap and then subsequently calculates the NOAA-12 biases relative to NOAA-11 (Christy et al., 2007). The UAH procedure tends to ignore short overlapping periods and, when possible, utilizes satellite overlaps of about two years to calculate intersatellite biases. UAH also utilizes co-orbiting satellites (for TLT) and various view angles (for TMT) to sample and remove the diurnal cycle drift of the satellites over time.

For this work we utilized UAH Version 5.3. After we began this work, UAH released Version 5.4, but the release only modified the reference period for the anomaly calculations, which would not effect any of the results in this work since we have calculated all anomalies using the 1995 - 2005 reference period.

UAH is planning on releasing a new dataset this year that utilizes a different technique for the non-linear radiometer calibration and the diurnal drift correction (Dr. John Christy, personal communication, 2012).

### 2.2.2 Remote Sensing Systems (RSS)

Since 2003, RSS has been used in tandem with the UAH dataset for climate applications (Mears et al., 2003; Mears and Wentz, 2005, 2009a,b; Mears et al., 2011). The differences between RSS and UAH have typically been considered structural uncertainties (e.g Karl

et al., 2006). In other words, both UAH and RSS utilize reasonable processing choices and it is impossible to determine which is more accurate.

RSS, unlike UAH, utilizes all available satellite overlaps when doing inter-satellite calibrations. This difference has been highlighted as especially important during time periods when satellite overlaps are short in the 1980s. Another important difference is that RSS utilizes a climate model for its diurnal cycle corrections.

RSS, like UAH, has homogenized satellite products for TLT, TMT, and TLS. In this work, we utilize the latest RSS dataset, RSS Version 3.3. We also calculate a T24 product using channels TMT and TLS.

### *2.2.3 National Oceanic and Atmospheric Administration (NOAA)*

The NOAA MSU/AMSU dataset is relatively new compared to RSS and UAH and was released in 2006 (Zou et al., 2006, 2009; Zou and Wang, 2011). The NOAA dataset is similar to RSS in its merging procedure, but NOAA adjusts the level 1 radiance data starting with NOAA-10 using simultaneous nadir overpasses of co-orbiting satellites (Zou et al., 2009). NOAA also utilizes the RSS diurnal cycle correction, but scales it by about 90% to minimize the residual errors for overlapping satellites.

NOAA does not produce a TLT product, but has channels TMT and TLS available, from which we can compute a T24 tropospheric product. We are utilizing NOAA Version 2.0 for this analysis.

## 2.3 Dataset intercomparison

### 2.3.1 T24 time series comparisons

In this section, we present some full tropospheric (T24) time series for the MSU/AMSU and radiosonde datasets (see Figure 1.1 for the T24 weighting function). We utilize T24 because NOAA does not yet produce a TLT product, but T24 largely eliminates stratospheric contamination, which makes it a worthwhile measure of tropospheric temperature. T24 is also advantageous because its weighting function is elevated relative to TLT, so the influence of the surface and diurnal drift effects should be reduced.

In Figures 2.3 and 2.4, we present time series for the T24 temperature evolution in the global mean and tropics, respectively.

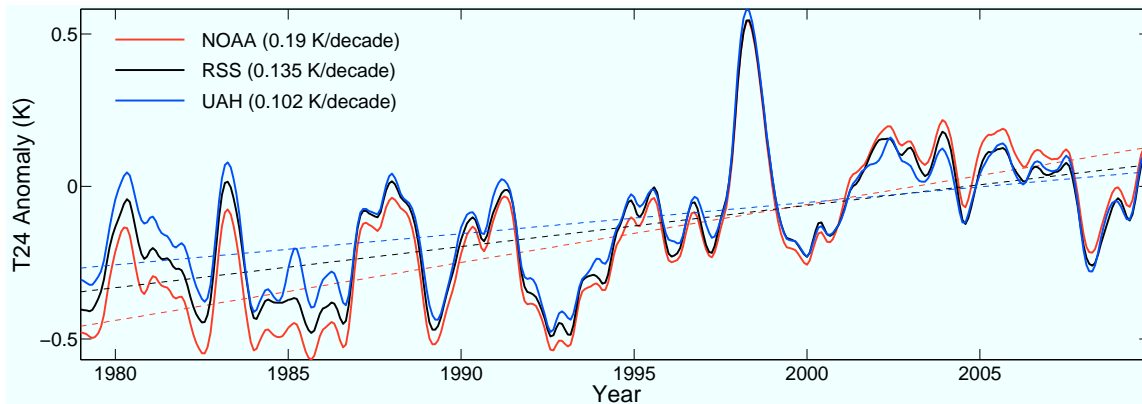


Figure 2.3: T24 time series for the three MSU/AMSU datasets averaged over the globe from 1979 - 2009. The reference period in which anomalies are calculated is 1995 - 2005. Time series are smoothed to reduce high frequency variations.

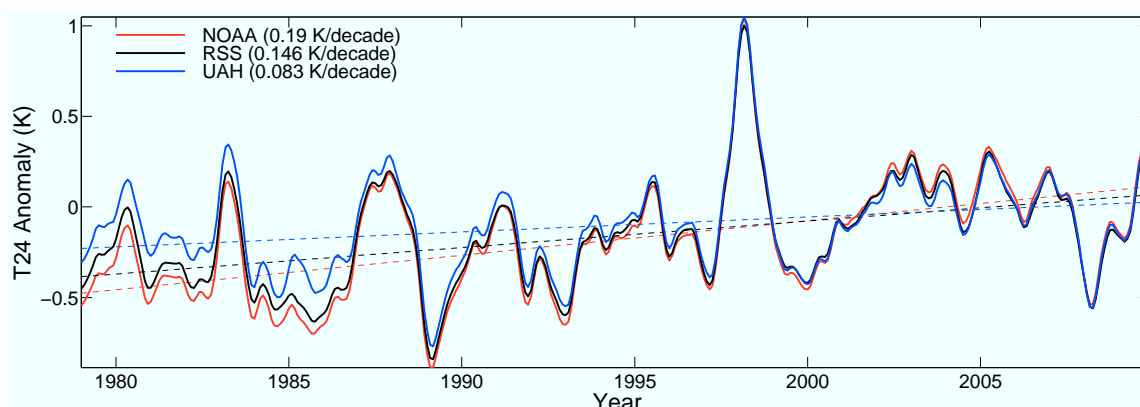


Figure 2.4: As in Figure 2.3, but for the tropics ( $30^{\circ}\text{S} - 30^{\circ}\text{N}$ ).

In this work radiosondes will act as independent measures of tropospheric temperature. We have calculated synthetic satellite brightness temperatures for the various radiosonde products. Their time series for channel T24 are given in Figures 2.5 and 2.6 for the global mean and tropics, respectively.

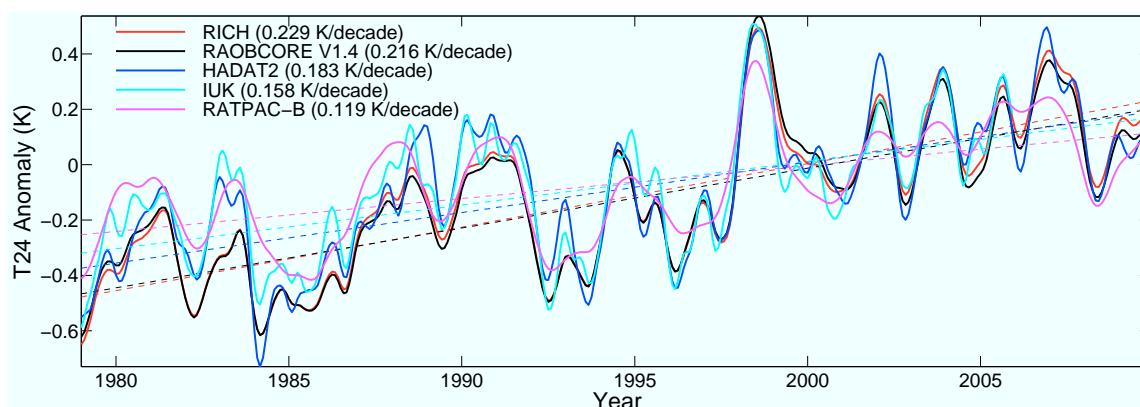


Figure 2.5: T24 time series for the five radiosonde datasets averaged over the globe from 1979 - 2009. The IUK analysis only lasts until 2005 and the RATPAC-B homogenization effectively lasts until 1997 (afterwards radiosondes may contain biases). The reference period in which anomalies are calculated is 1995 - 2005. Time series are smoothed to reduce high frequency variations.

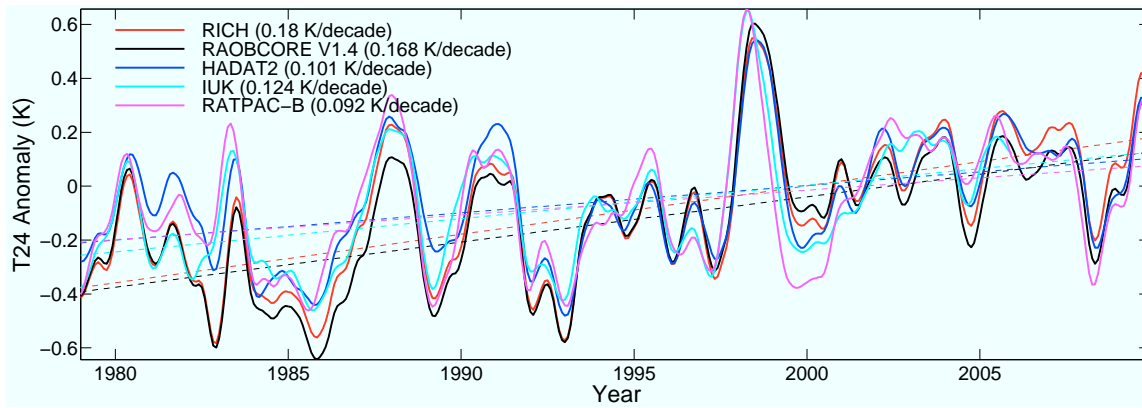


Figure 2.6: As in Figure 2.5, but for the tropics ( $30^{\circ}\text{S} - 30^{\circ}\text{N}$ ).

For this analysis, we are attempting to identify and reconcile differences in the MSU/AMSU products. To assist in identifying differences, we have taken the “difference series” for the various MSU/AMSU products in Figure 2.7 and 2.8. This allows us to see the relative differences between the datasets. For example, it is evident in Figure 2.7 that NOAA warms relative to RSS and UAH over the entire time series. We also see that, for example, UAH is different from RSS and NOAA over the 1985 - 1987 time period.

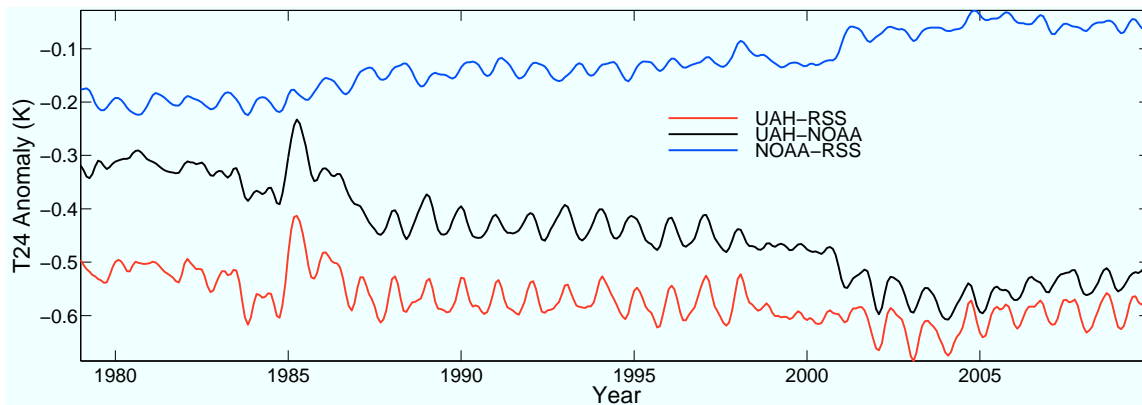


Figure 2.7: T24 difference time series for the three MSU/AMSU datasets averaged over the globe from 1979 - 2009. Time series are smoothed to reduce high frequency variations.

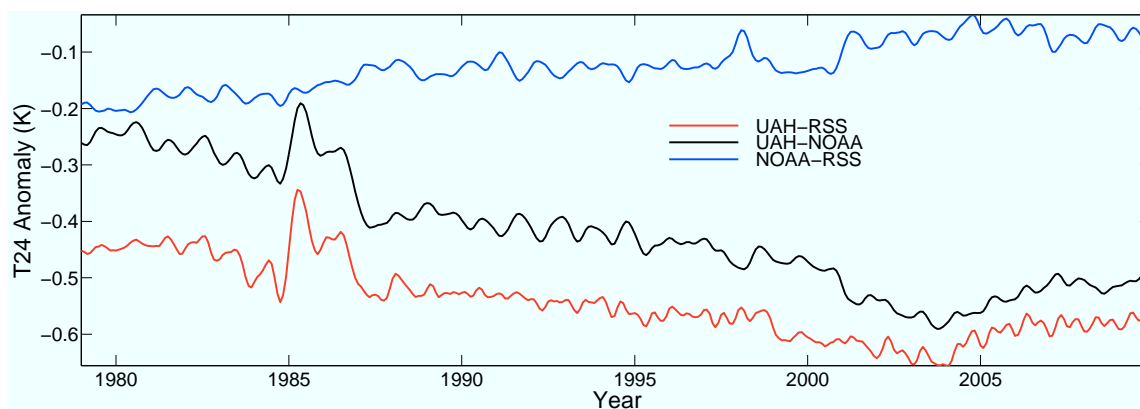


Figure 2.8: As in Figure 2.7, but for the tropics ( $30^{\circ}\text{S} - 30^{\circ}\text{N}$ ).

In this analysis, we will be utilizing radiosondes as references. To demonstrate the differences between MSU/AMSU observations and radiosonde observations, we took the MSU/AMSU minus radiosonde difference series in Figures 2.9 and 2.10.

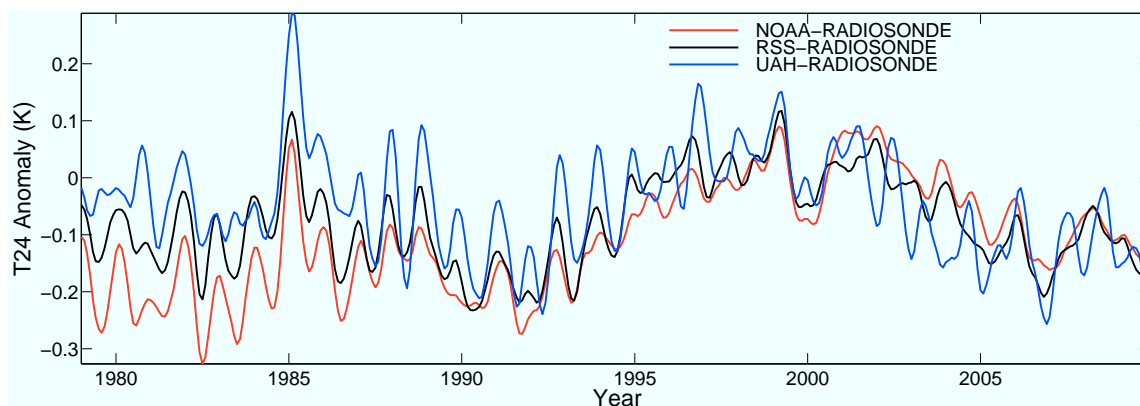


Figure 2.9: T24 difference time series for the three MSU/AMSU datasets (collocated with radiosondes) relative to the mean of the radiosonde datasets averaged over the globe from 1979 - 2009. Time series are smoothed to reduce high frequency variations.

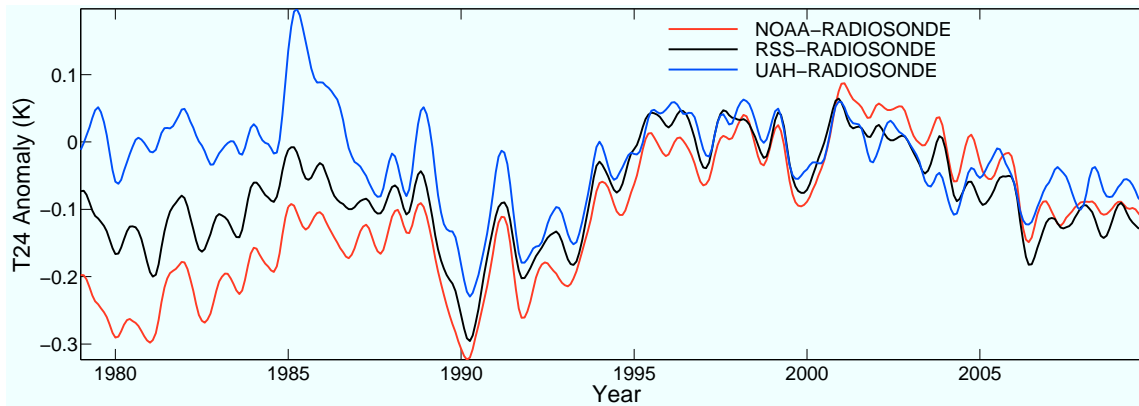


Figure 2.10: As in Figure 2.9, but for the tropics ( $30^{\circ}\text{S} - 30^{\circ}\text{N}$ ).

### 2.3.2 Current tropospheric temperature trend characteristics

Here we compare various radiosonde and MSU/AMSU trend estimates for the troposphere. To facilitate comparison between satellite and radiosonde datasets, we computed synthetic satellite channels for the radiosondes. The weighting functions for the bulk atmospheric layers used in this study are given in Figure 1.1.

In Table 2.1 we present the least squares linear trends from 1979 - 2005 for the datasets used in this work. To illustrate the effect of sampling at different radiosonde locations, we present Figures 2.11 and 2.12. Figure 2.11 illustrates the global mean trend estimates for the various upper air datasets and for different measures of tropospheric temperature. Since tropospheric amplification is expected to be prominent in the tropics, tropical trend estimates are provided in Figure 2.12. Note that it is often concluded that substantial residual errors likely remain in tropical radiosonde datasets where a sparse radiosonde network exists and radiation effects are significant for the often daytime only observations (Randel and Wu, 2006; Sherwood et al., 2005, 2008; Titchner et al., 2009; Mears and Wentz, 2012). In each comparison, the spatial sampling of each upper-air dataset can make an important difference in the trend estimate, though the effect is smaller in the tropics. In general, NOAA STAR has the greatest trends, UAH has the smallest trends, and RSS is approximately in between the two.

Table 2.1: Tropical and global trends for the datasets used in this work. Trends are calculated over 1979 - 2005 to accommodate the IUK dataset. Trend estimates are the least-squares linear fits in units of K decade<sup>-1</sup>. The confidence intervals are the 95% confidence intervals for the linear regression accounting for autocorrelation. NOAA does not provide a TLT product.

Group	Global Trends			Tropical Trends (30°S - 30°N)		
	TLT	T24	TMT	TLT	T24	TMT
NOAA	X	0.22 ± 0.08	0.15 ± 0.07	X	0.23 ± 0.14	0.17 ± 0.12
RSS	0.16 ± 0.06	0.16 ± 0.07	0.10 ± 0.07	0.17 ± 0.12	0.19 ± 0.14	0.13 ± 0.12
UAH	0.14 ± 0.07	0.12 ± 0.07	0.05 ± 0.07	0.10 ± 0.12	0.11 ± 0.14	0.05 ± 0.12
RICH	0.25 ± 0.07	0.23 ± 0.06	0.16 ± 0.06	0.18 ± 0.08	0.17 ± 0.07	0.13 ± 0.06
RAOBCORE	0.24 ± 0.07	0.22 ± 0.07	0.16 ± 0.06	0.16 ± 0.08	0.18 ± 0.07	0.14 ± 0.07
HadAT2	0.23 ± 0.07	0.19 ± 0.07	0.12 ± 0.06	0.13 ± 0.07	0.10 ± 0.07	0.06 ± 0.07
IUK	0.18 ± 0.06	0.16 ± 0.06	0.10 ± 0.06	0.14 ± 0.07	0.12 ± 0.08	0.08 ± 0.07
RATPAC-B	0.16 ± 0.05	0.13 ± 0.06	0.06 ± 0.06	0.11 ± 0.07	0.10 ± 0.09	0.03 ± 0.08

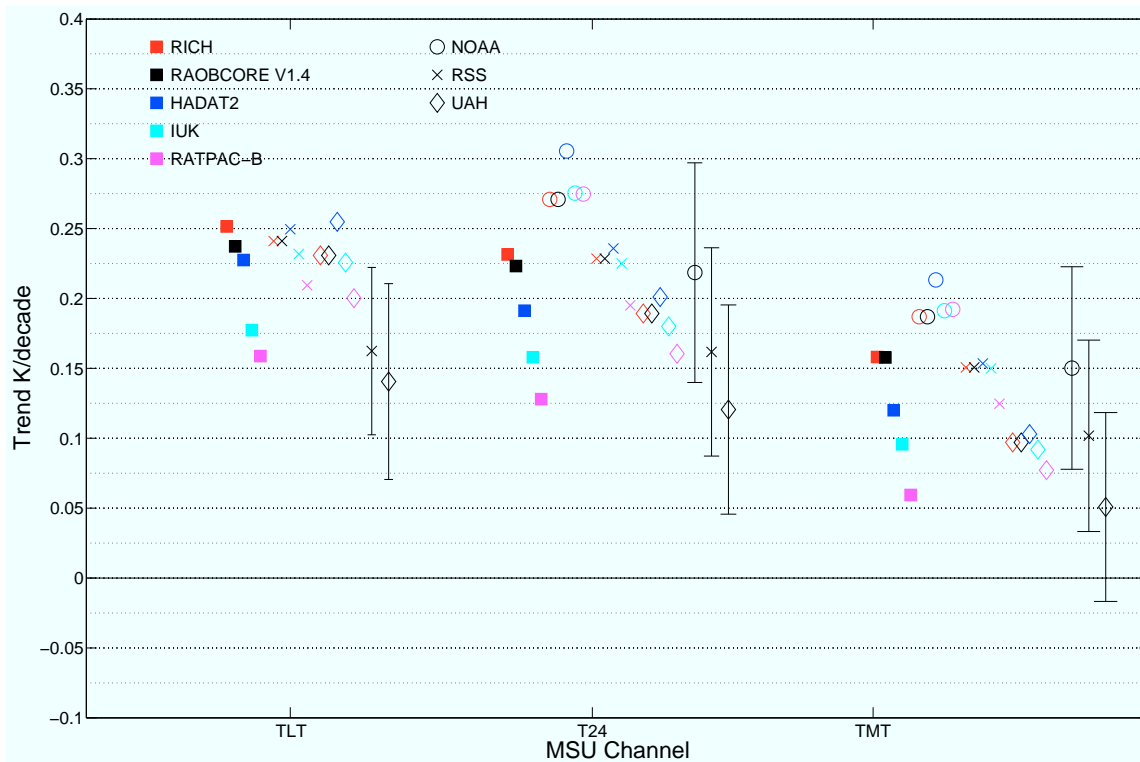


Figure 2.11: Global tropospheric trends for different deep layers. Synthetic satellite channels have been computed for various radiosonde products (solid squares). MSU/AMSU trends and statistical errors (95% confidence interval) have also been computed (open black markers). Further, the MSU/AMSU trends were calculated at grid points collocated with radiosonde observations for comparison. The collocated MSU/AMSU trends are denoted by open colored markers in which the color represents the radiosonde dataset that the MSU/AMSU dataset was collocated with and the marker type (diamond, circle, and X) denote the MSU/AMSU dataset. Trends are for 1979 - 2005 since the IUK dataset is only available through 2005.

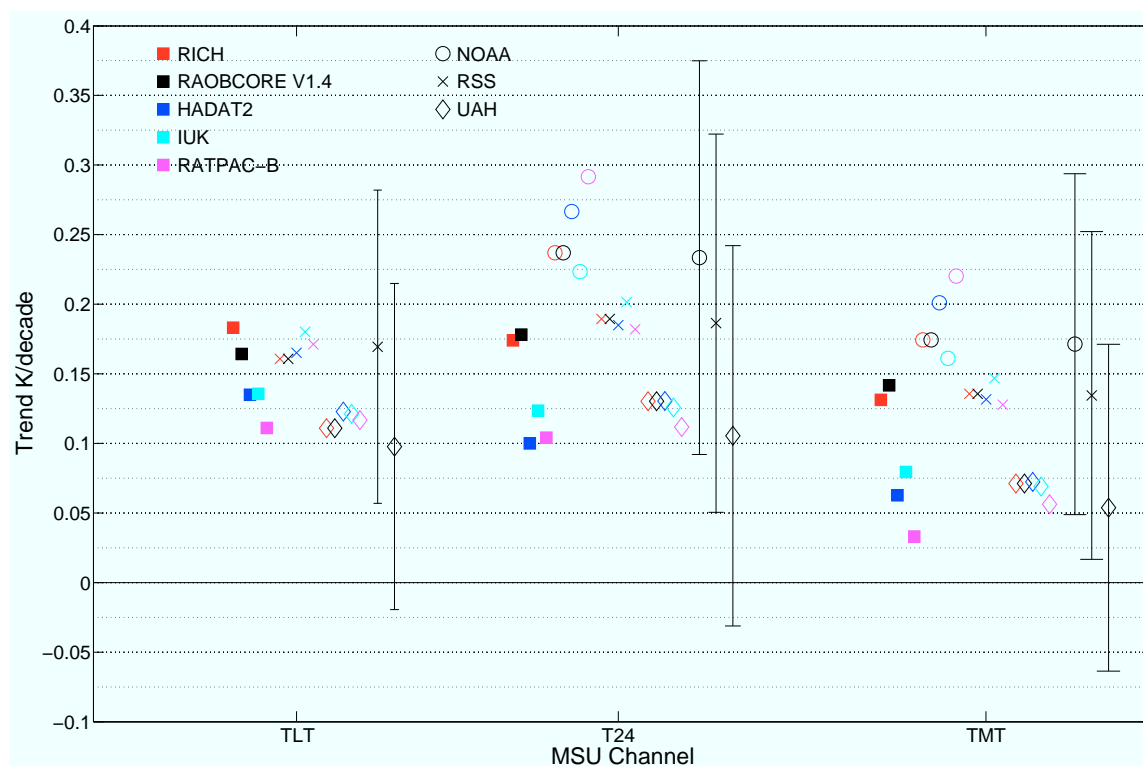


Figure 2.12: As in Figure 2.11, but for the tropical region bound by  $30^\circ$  South -  $30^\circ$  North.

In order to understand the spatial trend differences in trends derived by different MSU/AMSU teams, we present the global T24 trend maps in Figure 2.13 and the differences between the groups in Figure 2.14.

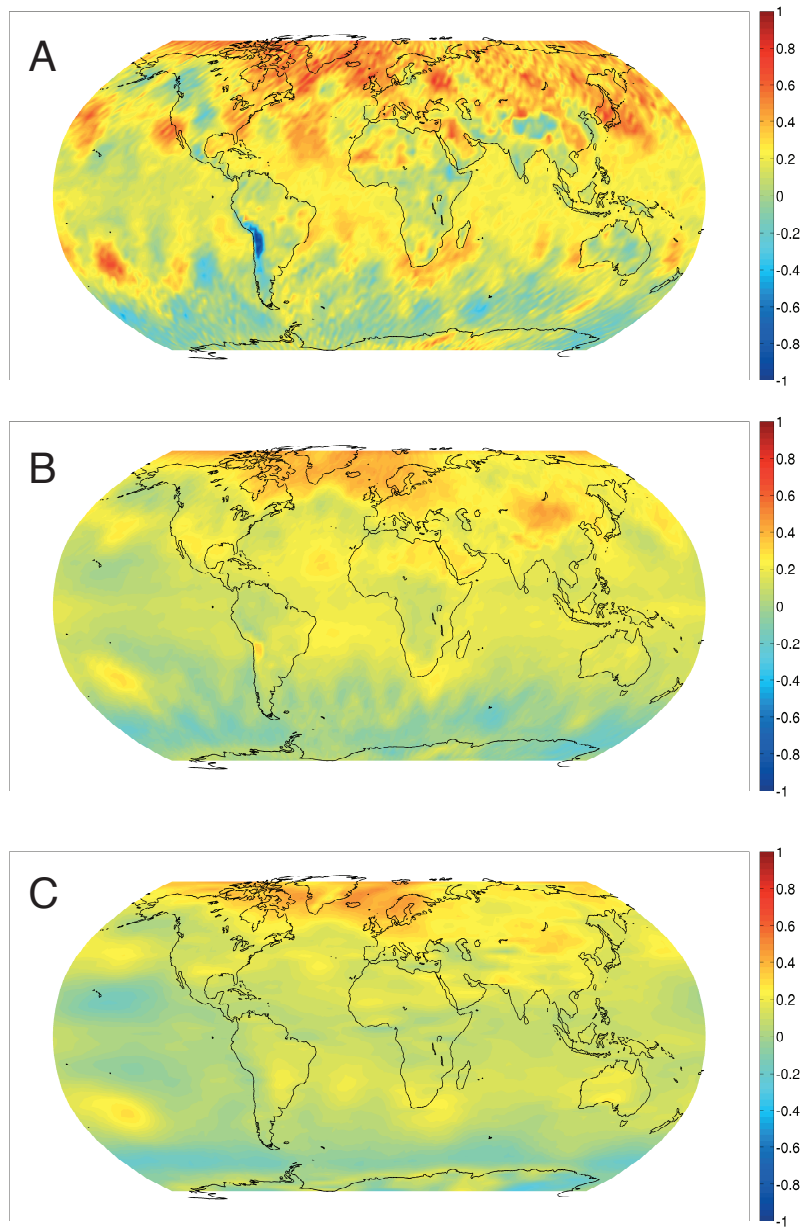


Figure 2.13: Global T24 trend map for a) NOAA, b) RSS, and c) UAH from 1979 - 2009.

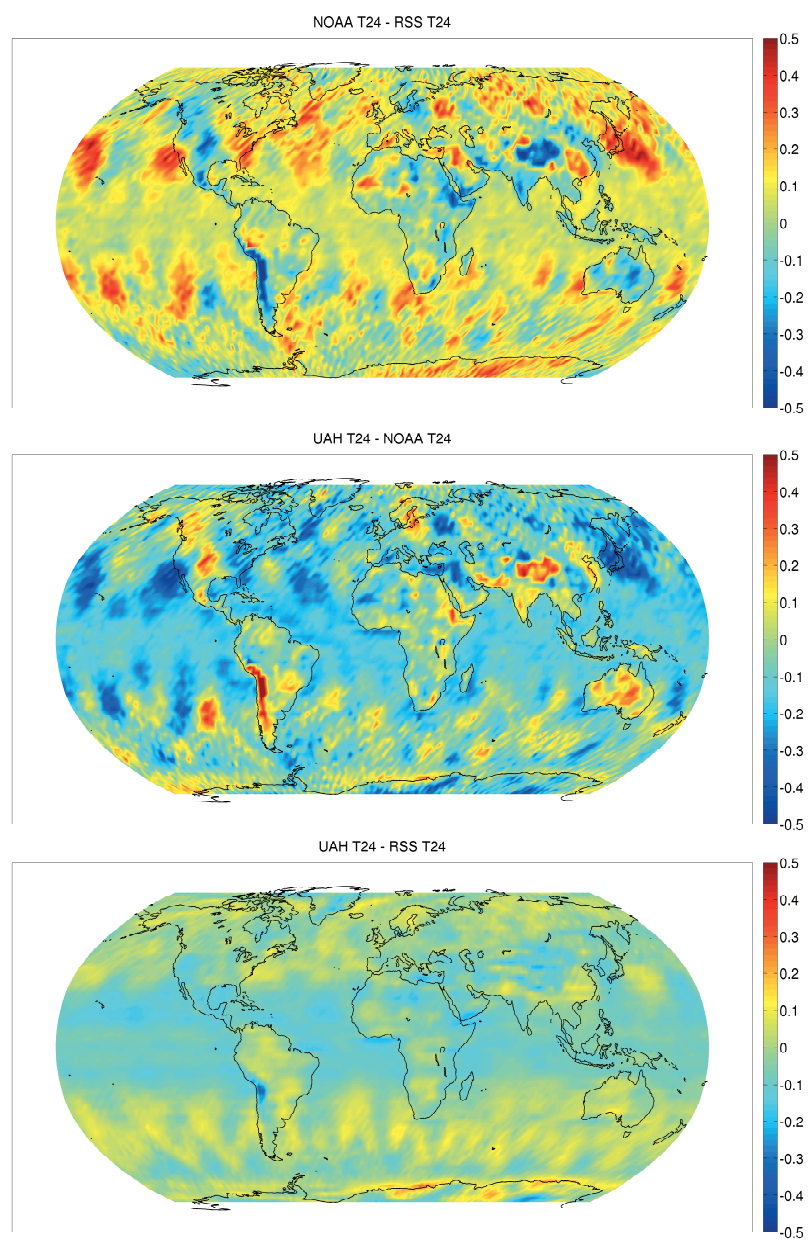


Figure 2.14: Global T24 trend difference maps for 1979 - 2009.

## Chapter 3

## TREATMENT OF THE NOAA-9 SATELLITE

**3.1 The NOAA-9 warm target factor**

UAH, RSS, and NOAA remove the effect of the instrument body temperature on the measured Earth radiance using an equation developed by Christy et al. (2000), which can be written as

$$T_{MEAS,i} = T_o + A_i + \alpha_i T_{TARGET,i} + \Delta T_{MSU,i} \quad (3.1)$$

where  $T_{MEAS,i}$  is the measured brightness temperature for the  $i$ th satellite,  $T_o$  is the actual Earth brightness temperature,  $A_i$  is a constant offset,  $T_{TARGET,i}$  is the temperature anomaly of the warm calibration target on the satellite as measured by platinum resistance thermometers,  $\alpha_i$  is the warm target factor, and  $\Delta T_{MSU,i}$  represents unresolved residual errors (Mears et al., 2003). Note that  $A_i$  is on the order of 0.01 - 1 K (Mears et al., 2003) and  $\Delta T_{MSU,i}$  is on the order of 0.1 K (Zou et al., 2009), but these values vary in time, space, and by satellite.

The correction of the instrument body temperature effect is to account for post-launch changes in the non-linear operational calibration (e.g. Zou et al., 2006). Equation 1 is applied to co-orbiting satellite pairs in order to determine the warm target factor for the  $i$ th and  $j$ th satellite by minimizing square of the measured brightness temperature differences for collocated observations averaged over 5 days (pentads):

$$T_{diff}(t_n) = T_{MEAS,i}(t_n) - T_{MEAS,j}(t_n) \quad (3.2)$$

where  $T_{diff}$  is the difference in the brightness temperature measured by the two satellites (Christy et al., 2000; Mears et al., 2003).

Figure 3.1 shows the warm target factors used by different MSU teams for each satellite and the time periods for which each satellite is incorporated into the TMT product. NOAA

and RSS use oceanic pentads to determine the warm target factor in order to minimize the potential influence of the diurnal cycle drift effect (Mears et al., 2003; Zou et al., 2009) whereas UAH uses both land and ocean pentads to determine warm target factor. There is typically good agreement for the selection of the warm target factor between UAH and RSS. An important exception is for NOAA-9, when the RSS and UAH target factor differs by more than a factor of two: 0.040 for RSS versus 0.099 for UAH (Dr. J. Christy, personal communication, 2011; Mears and Wentz, 2009a). This difference arises because the UAH team does not consider short satellite overlaps, which reduces the number of pentad pairs used to determine the warm target factors (Mears et al., 2003). Target factors for the NOAA team starting with the NOAA-10 satellite are different in part because the team first addresses non-linear calibration issues using simultaneous nadir overpasses of co-orbiting satellites and then apply the instrument body effect corrections (Zou et al., 2006, 2009; Zou and Wang, 2010). The NOAA team warm target factor for the NOAA-9 satellite is 0.025 (Drs. C. Z. Zou and W. H. Wang, personal communication, 2011). The large difference in the NOAA-9 target factor between UAH and RSS (NOAA) implies that one or more teams is over or underestimating the influence of the warm target temperature on the measured Earth radiance. As such, artificial residuals will remain in the temperature time series.

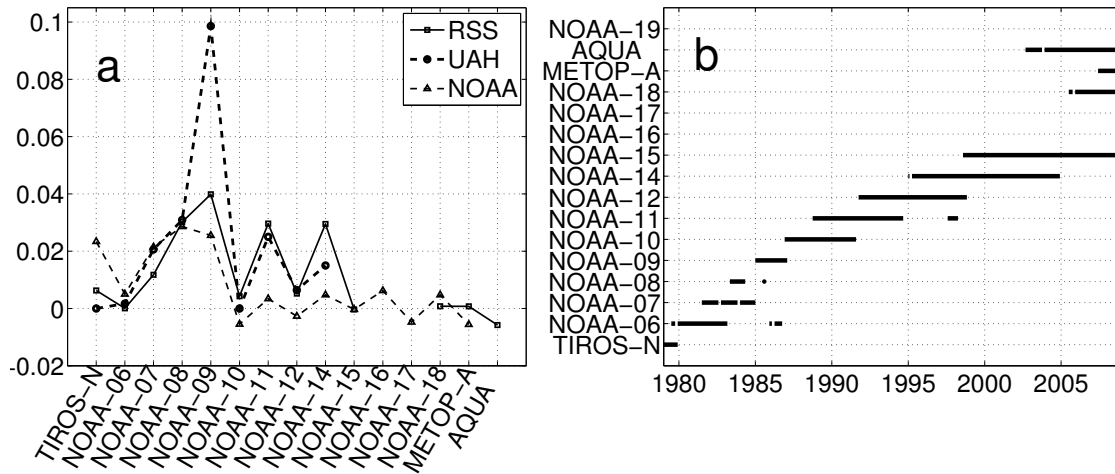


Figure 3.1: a) TMT warm target factors used for different MSU teams. b) Satellites used in the RSS TMT merge (Mears and Wentz, 2009a). Note that the satellites used are different for the various MSU teams.

We find that the difference between any two team's TMT anomaly series is significantly correlated (95% confidence) with the global mean NOAA-9 warm target temperature from January 1985 - February 1987 (26 months). We demonstrate the relationship between the TMT differences and the target temperature in Figure 3.2. For example, the correlation coefficient ( $r$ ) for  $UAH - NOAA$  and  $UAH - RSS$  versus  $T_{TARGET}$  is -0.90 and -0.83, respectively. This implies that the warm target calibration does explain some of the differences between the MSU datasets. As a result of the warm target temperature drift during NOAA-9's operational life, these differences will also affect the merged TMT trends. In this study, we utilize radiosondes as references to find biases in the warm target factor,  $\alpha_i$ .

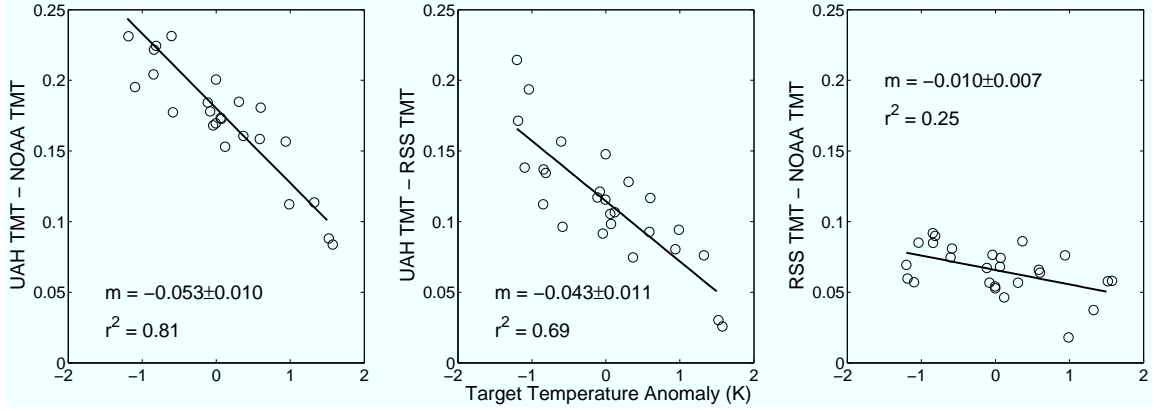


Figure 3.2: Scatter plot of MSU - MSU TMT difference series versus the warm target temperature over the NOAA-9 time period.  $m$  represents the slope for each of the relationships. The variance explained from the leftmost subplot to the rightmost subplot is 0.69, 0.81, and 0.25, respectively. The error is the 95% confidence interval in the linear fit.

If we define the reported brightness temperature as the measured brightness temperature minus the constant offset and the influence of the instrument body effect, i.e.,  $T_{MSU,i} = T_{MEAS,i} - A_i - \alpha_i T_{TARGET,i}$ , then Eq. 3.1 becomes:

$$T_{MSU,i} = T_o + \Delta T_{MSU,i}, \quad (3.3)$$

which is simply the actual Earth brightness temperature plus the unresolved errors in the MSU/AMSU measurement.

We can similarly define the radiosonde temperature measurement as the sum of the actual Earth brightness temperature and some unresolved error:

$$T_R = T_o + \Delta T_R \quad (3.4)$$

where  $T_R$  is the radiosonde measurement,  $T_o$  represents the actual signal, and  $\Delta T_R$  represents the measurement error of the radiosonde. By taking the difference of Eq. 3.3 and Eq. 3.4, we have

$$T_{MSU} - T_R = (T_o + \Delta T_{MSU}) - (T_o + \Delta T_R) \quad (3.5)$$

where the subscripts MSU and R refer to MSU/AMSU and radiosonde measurements, respectively. This leaves

$$T_{MSU} - T_R = \Delta T_{MSU} - \Delta T_R. \quad (3.6)$$

Note that the MSU error may be written as  $\Delta T_{MSU} = -\Delta\alpha_i T_{TARGET,i} + \epsilon_i$  where  $\epsilon_i$  represents any unresolved errors unrelated to the warm target calibration and  $\Delta\alpha_i$  represents a bias in the warm target factor. The radiosonde measurement error (i.e.  $\Delta T_R$ ) is unrelated to the satellite warm target temperature. We will examine the correlation between  $T_{MSU} - T_R$  (i.e.  $\Delta T_{MSU} - \Delta T_R$ ) and the warm target temperature,  $T_{TARGET}$ . We expect no (a non-zero) correlation if the instrument body effect is (is not) removed from the MSU/AMSU dataset. Some of the problems that bias radiosondes, such as solar heating effects and instrument changes (e.g. Gaffen, 1994; Sherwood et al., 2005), have no physical relationship with the warm target temperature and thus have little effect on our procedure. Note that by assuming that  $\Delta T_R$  cannot explain the variance in  $T_{TARGET}$  we are not assuming that  $\Delta T_R \sim 0$ . For our application radiosondes act as independent observations to remove the real signal of the Earth radiance.

## 3.2 Determining the bias in the NOAA-9 target factor

### 3.2.1 Quantifying the bias

Since the NOAA-9 target factors are very different for each group, we expect that some fraction of the TMT anomaly series will be related to the warm target temperature for one or more MSU teams. We can rewrite Eq. 3.1 as:

$$T_{MEAS,i} = T_o + A_i + (\alpha_{i,o} + \Delta\alpha_i)T_{TARGET,i} + \epsilon_i \quad (3.7)$$

where  $\alpha_{i,o}$  is the optimized warm target factor and  $\Delta\alpha_i$  is a measure of the over or underestimate of the warm target factor. If  $\Delta\alpha_i \neq 0$ , residuals related to the warm target factor will remain in the TMT time series.

We examine difference time series for each MSU team's TMT product with a reference time series derived from radiosondes (see Eq. 3.6). If this difference time series over the NOAA-9 period (January 1985 - February 1987) is correlated with the global average warm

target temperature, we can conclude that the warm target calibration for NOAA-9 is artificially affecting TMT. The slope of the temperature residuals versus the warm target temperature is a measure of the warm target factor bias,  $\Delta\alpha_i$ .

In Table 3.1, we present the correlation coefficients for  $MSU - REFERENCE$  versus  $T_{TARGET,9}$  over the NOAA-9 period. We see that the warm target factor is significantly related to the UAH TMT product, regardless of the reference dataset. RSS and NOAA are not significantly related to the warm target temperature for any of the radiosonde reference datasets.

Table 3.1: Correlation coefficients for MSU (column) - REFERENCE (row) versus the global mean warm target temperature for NOAA-9 during January 1985 to February 1987. Values denoted by an asterisk are significant with 95% confidence. The ‘‘Radiosonde Mean’’ is the correlation coefficient of the mean of the five UAH - REFERENCE time series versus the global warm target temperature for NOAA-9.

$\downarrow REFERENCE \downarrow$	<i>NOAA</i>	<i>RSS</i>	<i>UAH</i>
HadAT2	-0.044	-0.087	-0.404*
IUK	0.146	0.105	-0.443*
RICH	-0.143	-0.245	-0.536*
RAOBCORE	-0.106	-0.203	-0.510*
RATPAC	-0.235	-0.357	-0.758*
Radiosonde Mean	-0.077	-0.166	-0.573*

In Figure 3.3 we show the relationship for UAH-RADIOSONDE versus  $T_{TARGET,9}$  over the NOAA-9 period. In this case, we averaged the five collocated (UAH-RADIOSONDE) time series and regressed this average time series against  $T_{TARGET,9}$ . This method has the advantage of reducing random noise related to the different radiosonde datasets. This radiosonde mean estimate for  $\Delta\alpha_9$  for UAH is  $0.051 \pm 0.031$  (95% confidence). We estimate the  $\Delta\alpha_9$  value for UAH, because it is significantly correlated with every radiosonde reference. The sign difference between the slope in Figure 3.3 and  $\Delta\alpha_9$  results because the instrument body effect correction is subtracted from the measured brightness temperature (the reported

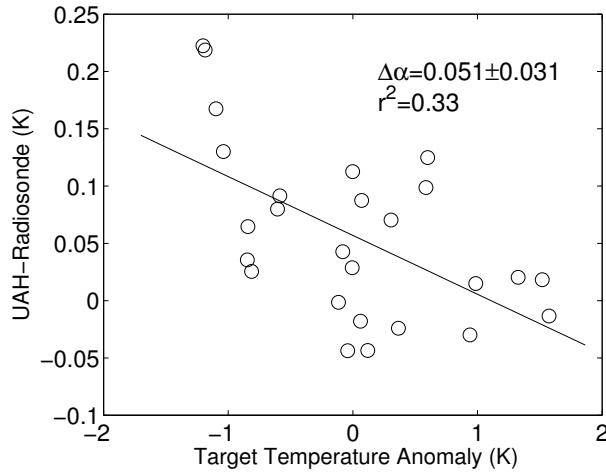


Figure 3.3: Scatter plot of the mean of the five collocated *UAH* – *RADIOSONDE* difference series versus the warm target temperature.  $\Delta\alpha = -slope$  because the  $\alpha \cdot T_{TARGET}$  term is subtracted from the measured brightness temperature to obtain the calibrated UAH brightness temperature ( $T_{MSU}$ ). The error is the 95% confidence interval in the linear fit.

brightness temperature will depend on  $T_{TARGET}$  through  $-\Delta\alpha \cdot T_{TARGET}$ ).

In Table 3.2, we provide the  $\Delta\alpha_9$  value for UAH compared to each reference dataset, which is the magnitude of the slope of the relationship between UAH minus REFERENCE and  $T_{TARGET,9}$ . This value should be subtracted from  $\alpha_9$  to correct the UAH TMT time series. Using our radiosonde mean estimate, we find that UAH overestimated the warm target factor for NOAA-9 by  $0.051 \pm 0.031$ . In calculating the confidence interval, we used the fit error from Figure 3.3. This value compares favorably to the warm target factor difference for UAH minus RSS (NOAA). The warm target factor difference between UAH and RSS (UAH and NOAA) is  $\Delta\alpha_9 = 0.059$  ( $\Delta\alpha_9 = 0.073$ ).

Table 3.2:  $\Delta\alpha_9$  values and 95% confidence intervals derived from our least-squares linear fit. These values are the magnitude of the slope of the linear relationship between UAH (column) - REFERENCE (row) versus the global mean warm target temperature over the NOAA-9 operational period. This value should be subtracted from  $\alpha_9$  to correct for the non-optimal selection of warm target factor. The “Radiosonde Mean” is the same as that from Table 1 and Figure 2.

$\downarrow$ REFERENCE $\downarrow$	UAH $\Delta\alpha_9$
HadAT2	$0.041 \pm 0.039$
IUK	$0.038 \pm 0.032$
RICH	$0.056 \pm 0.037$
RAOBCORE	$0.051 \pm 0.036$
RATPAC	$0.071 \pm 0.026$
Radiosonde Mean	$0.051 \pm 0.031$

### 3.2.2 The effect of the NOAA-9 warm target bias on the UAH TMT trend

In order to estimate the effect of the NOAA-9 warm target bias on the global mean TMT trend, we began by determining the trend in the global mean warm target temperature during NOAA-9’s operational period (January 1985 through February 1987) using a least-squares linear fit. For the global average warm target temperature, we obtain a trend of  $1.15 \text{ K year}^{-1}$ . Its impact on the global UAH TMT product over NOAA-9 is then  $-1.15 \text{ K year}^{-1} \cdot \Delta\alpha_9 \cdot t$ , where  $t$  is time in years starting from January 1985. To estimate the impact of this bias on the long-term trend, we set a synthetic time series as zero from January 1979 through December 1984,  $-1.15 \text{ K year}^{-1} \cdot \Delta\alpha_9 \cdot t$  during January 1985 through February 1987, and  $-1.15 \text{ K year}^{-1} \cdot \Delta\alpha_9 \cdot \frac{26}{12}$  for March 1987 through December 2009. We then take the least-squares linear trend of this synthetic time series to estimate the impact on the long-term trend. Using this approach, we estimate the effect of the NOAA-9 target factor on the long term UAH TMT trend as  $-0.042 \text{ K decade}^{-1}$  (1979 - 2009). This estimate indicates that the global UAH TMT trends are artificially reduced by this bias.

In Figure 3.4, we use the same procedure to estimate the spatial trend effect at each grid

point because the  $T_{TARGET,9}$  drift is not spatially homogeneous. Spatial differences in Figure 3.4 are due to differences in the target temperature drift over the NOAA-9 operational period, since we used a spatially constant  $\Delta\alpha_9$  value (MSU teams use a constant value for  $\alpha_9$ ). Fundamentally, the target temperature is related to the solar zenith angle, which influences the heating and cooling of the instrument (Zou and Wang, 2011). The instrument body effect does not explain differences in land versus ocean trends for the different MSU/AMSU teams, because the effect is related to the non-linear radiometer calibration and does not depend on the surface type (Mears et al., 2003). The pronounced striping in Figure 3.4 may be a result of non-uniform sampling of the seasonal cycle over NOAA-9’s life. A similar observation was noted by Mears et al. (2003) when looking at the differences between ascending and descending nodes of the satellite orbit.

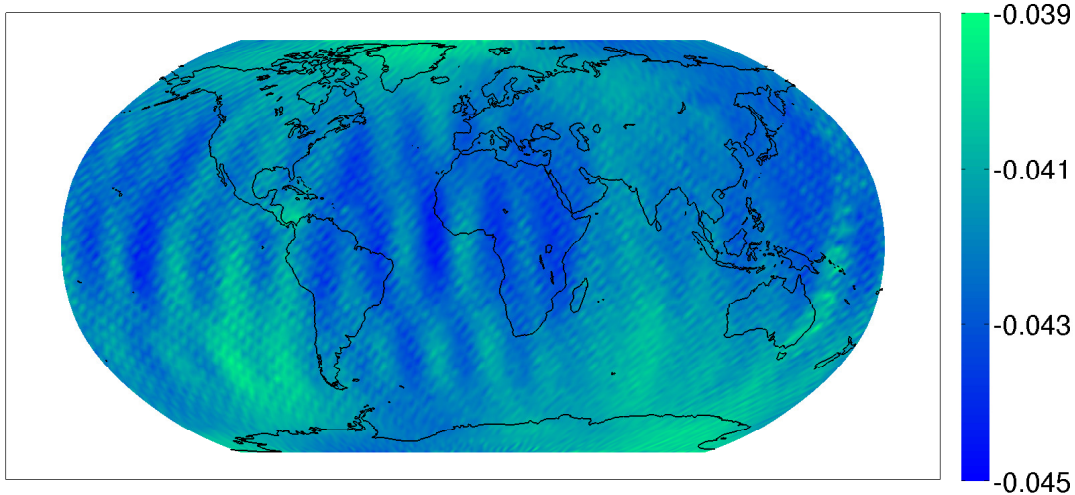


Figure 3.4: Estimate of the spatial impact of the UAH NOAA-9 warm target bias on UAH TMT trends ( $\text{K decade}^{-1}$ ).

Our global mean trend sensitivity to the NOAA-9 target factor is similar to that estimated by Mears et al. (2003). When the RSS team used the UAH NOAA-9 TMT target factor, they found a trend decrease of  $0.073 \text{ K decade}^{-1}$ . On the other hand, when the UAH

team used the RSS target factor they found a trend increase of  $0.05 \text{ K decade}^{-1}$  (Mears et al., 2003). Comparisons of the present result with those in Mears et al. (2003) differ in part because of version changes, including target factor changes for both groups. We also expect the trend effect of the NOAA-9 target factor to be smaller now since the time series is longer. We find a UAH trend increase of  $0.064 \text{ K decade}^{-1}$  when the trend spans 1979 - 2002, consistent with the findings in Mears et al. (2003). While our estimate of the UAH TMT trend sensitivity to the warm target factor is similar to estimates by the UAH and RSS teams, our study indicates that the UAH team will need to incorporate an optimal NOAA-9 target factor into their merging procedure for an accurate trend estimate.

After the estimate in this work was published (Po-Chedley and Fu, 2012), the UAH team performed a sensitivity analysis and determined that the effect of the NOAA-9 target factor is only  $0.022 \text{ K decade}^{-1}$  (see Appendix A). We obtain a similar estimate of the trend effect when we estimate the drift in NOAA-9 between the NOAA-6/NOAA-9 overlap and the NOAA-9/NOAA-10 overlap. This indicates that our estimate for the bias in the UAH trend may be too large.

Importantly, the NOAA-9 target factor cannot simply be adjusted in an ad hoc manner, since the target factor is selected to minimize the errors between overlapping satellites. It has been pointed out that by decreasing the target factor for NOAA-9, the inter-satellite differences increase for the satellites that UAH considers (Appendix A). Mears et al. (2003) found that utilizing the UAH warm target factor increases the inter-satellite difference trends for satellites not considered by UAH. The fact that the influence of the NOAA-9 is evident in the final merged TMT product indicates that the UAH merging algorithm is not robust in determining the calibration coefficients. This may occur because of errors in the diurnal drift calibration (Dr. Carl Mears, personal communication, 2012) or a result of the satellite pairs used in the bias correction procedure.

Current global TMT trends for NOAA, RSS, and UAH are  $0.127 \text{ K decade}^{-1}$ ,  $0.080 \text{ K decade}^{-1}$ , and  $0.038 \text{ K decade}^{-1}$  (1979 - 2009). By correcting the bias in the UAH warm target factor, the global UAH TMT trend becomes  $0.080 \text{ K decade}^{-1}$ , which effectively eliminates the global UAH and RSS trend difference and reduces the global UAH and NOAA trend difference by 47%.

A key consideration for tropospheric trend interpretation is the contamination of the TMT product by the stratosphere (e.g. Fu and Johanson, 2005; Johanson and Fu, 2006). Using a combination of TMT and TLS (referred to as T24) (Fu et al., 2004; Johanson and Fu, 2006), the UAH T24 trend will increase if this bias is taken into account. We carried out an identical analysis for TLT, but there was no statistically significant relationship between UAH (RSS) minus REFERENCE and the global warm target temperature during the NOAA-9 period for four of five radiosonde references. This result is unsurprising because the TLT product is a linear combination of signals from different view angles, which amplifies noise by a factor of three compared to TMT (Hurrell and Trenberth, 1998; Christy et al., 2000; Mears and Wentz, 2009b). Importantly, UAH uses the same target factors for the TLT product (Christy et al., 2000) (RSS uses  $\alpha_9 = 0.049$ ) and there is a significant relationship for UAH minus RSS TLT versus  $T_{TARGET,9}$  ( $r=0.64$ ,  $\Delta\alpha_9 = 0.054$ , 95% confidence). The NOAA-9 warm target bias thus has a similar effect on the UAH TLT product.

There is likely a residual bias related to the NOAA and/or RSS NOAA-9 warm target factor, even though the present study indicates that RSS and NOAA have no significant biases. Note that TMT residuals for RSS minus NOAA are significantly related to  $T_{TARGET,9}$  ( $r=-0.501$ ). This indicates that some of the differences between RSS and NOAA during this time are related to the instrument body temperature effect, but because neither dataset is significantly related to the warm target temperature when using radiosondes as references, the bias is too small to be quantified by radiosonde observations.

## Chapter 4

**CONCLUSION**

Using radiosondes as references, we were able to attribute a bias in the NOAA-9 warm target factor to the UAH team and quantify its magnitude. The bias was statistically significant and compared well with the UAH minus RSS (NOAA) warm target value differences. We estimate that the UAH NOAA-9 warm target bias is  $0.051 \pm 0.031$ . Accounting for this problem, the global UAH TMT trend increases by an estimated  $0.042 \text{ K decade}^{-1}$  (1979 - 2009), which reconciles a majority of the current global trend difference between UAH and RSS. Since 1) our estimate of the trend impact is simple, 2) the bias in the NOAA-9 satellite might affect other satellites, and 3) UAH finds that the trend increases by just  $0.022 \text{ K decade}^{-1}$  when utilizing our target factor, an alternative merging procedure may be required to accurately determine the impact of this bias on the UAH TMT trend. Regardless, the UAH trend should increase when this problem is taken into account. Since the UAH TMT warm target factors are also used for the TLT product (Christy et al., 2000) and the T24 product utilizes TMT, T24 and TLT trends are also affected by this bias.

There is likely a residual bias related to the NOAA and/or RSS NOAA-9 warm target factor, even though the present study finds that RSS and NOAA have no significant biases. Note that TMT residuals for  $RSS - NOAA$  are significantly related to  $T_{TARGET,9}$  ( $r = -0.501$ ). This indicates that some of the differences between RSS and NOAA during this time are related to the instrument body temperature effect, but because neither dataset is significantly related to the warm target temperature when using radiosondes as references, the bias is too small to be quantified by radiosonde observations.

Creating climate-quality satellite temperature datasets is a challenging process that requires constant attention as new biases are discovered (e.g. Wentz and Schabel, 1998; Christy et al., 2000; Fu and Johanson, 2005; Mears and Wentz, 2005). Independent measurements of atmospheric temperature such as those from radiosondes will continue to be an important

tool in evaluating satellite temperature products over limited time periods.

## BIBLIOGRAPHY

- Allen, R. J. and S. C. Sherwood, 2008: Warming maximum in the tropical upper troposphere deduced from thermal winds. *Nature Geosci*, **1 (6)**, 399–403, URL <http://dx.doi.org/10.1038/ngeo208>.
- Anthes, R. A., et al., 2008: The cosmic/formosat-3 mission: Early results. *Bulletin of the American Meteorological Society*, **89 (3)**, 313–333, URL <http://dx.doi.org/10.1175/BAMS-89-3-313>.
- Bengtsson, L. and K. Hodges, 2011: On the evaluation of temperature trends in the tropical troposphere. *Climate Dynamics*, **36**, 419–430, URL <http://dx.doi.org/10.1007/s00382-009-0680-y>, 10.1007/s00382-009-0680-y.
- Bretherton, C. S. and A. H. Sobel, 2003: The gill model and the weak temperature gradient approximation. *Journal of the Atmospheric Sciences*, **60 (2)**, 451–460, URL [http://dx.doi.org/10.1175/1520-0469\(2003\)060<0451:TGMATW>2.0.CO;2](http://dx.doi.org/10.1175/1520-0469(2003)060<0451:TGMATW>2.0.CO;2).
- Brohan, P., J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones, 2006: Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850. *J. Geophys. Res.*, **111 (D12)**, URL <http://dx.doi.org/10.1029/2005JD006548>.
- Christy, J. R. and W. B. Norris, 2004: What may we conclude about global tropospheric temperature trends? *Geophys. Res. Lett.*, **31 (6)**, URL <http://dx.doi.org/10.1029/2003GL019361>.
- Christy, J. R. and W. B. Norris, 2006: Satellite and *viz*-radiosonde intercomparisons for diagnosis of nonclimatic influences. *Journal of Atmospheric and Oceanic Technology*, **23 (9)**, 1181–1194, URL <http://dx.doi.org/10.1175/JTECH1937.1>.
- Christy, J. R., W. B. Norris, R. W. Spencer, and J. J. Hnilo, 2007: Tropospheric temperature

- change since 1979 from tropical radiosonde and satellite measurements. *J. Geophys. Res.*, **112 (D6)**, URL <http://dx.doi.org/10.1029/2005JD006881>.
- Christy, J. R. and R. W. Spencer, 2005a: Correcting temperature data sets. *Science*, **310 (5750)**, 972–973, doi:10.1126/science.310.5750.972, URL <http://www.sciencemag.org/content/310/5750/972.short>, <http://www.sciencemag.org/content/310/5750/972.full.pdf>.
- Christy, J. R. and R. W. Spencer, 2005b: Correcting temperature data sets. *Science*, **310 (5750)**, 972.
- 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 (9)**, 1153–1170.
- Christy, J. R., R. W. Spencer, and E. S. Lobl, 1998: Analysis of the merging procedure for the msu daily temperature time series. *Journal of Climate*, **11 (8)**, 2016 – 2041.
- Christy, J. R., R. W. Spencer, and R. T. McNider, 1995: Reducing noise in the msu daily lower-tropospheric global temperature dataset. *Journal of Climate*, **8 (4)**, 888–896, URL [http://dx.doi.org/10.1175/1520-0442\(1995\)008<0888:RNITMD>2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(1995)008<0888:RNITMD>2.0.CO;2).
- Christy, J. R., R. W. Spencer, and W. B. Norris, 2011: The role of remote sensing in monitoring global bulk tropospheric temperatures. *Int. J. Remote Sens.*, **32 (3)**, 671–685, doi:10.1080/01431161.2010.517803, URL <http://dx.doi.org/10.1080/01431161.2010.517803>.
- 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 (5)**, 613–629.
- Christy, J. R., et al., 2010: What do observational datasets say about modeled tropospheric temperature trends since 1979? *Remote Sensing*, **2 (9)**, 2148–2169, doi:10.3390/rs2092148.

- Dai, A. and K. E. Trenberth, 2004: The diurnal cycle and its depiction in the community climate system model. *Journal of Climate*, **17** (5), 930–951, URL [http://dx.doi.org/10.1175/1520-0442\(2004\)017<0930:TDCAID>2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(2004)017<0930:TDCAID>2.0.CO;2).
- Douglass, D. H., J. R. Christy, B. D. Pearson, and S. F. Singer, 2008: A comparison of tropical temperature trends with model predictions. *International Journal of Climatology*, **28** (13), 1693–1701, doi:10.1002/joc.1651.
- Free, M., D. J. Seidel, J. K. Angell, J. Lanzante, I. Durre, and T. C. Peterson, 2005: Radiosonde atmospheric temperature products for assessing climate (ratpac): A new data set of large-area anomaly time series. *J. Geophys. Res.*, **110** (D22).
- Free, M., et al., 2002: Creating climate reference datasets: Cards workshop on adjusting radiosonde temperature data for climate monitoring. *Bulletin of the American Meteorological Society*, **83** (6), 891–899, URL [http://dx.doi.org/10.1175/1520-0477\(2002\)083<0891:CCRDCW>2.3.CO;2](http://dx.doi.org/10.1175/1520-0477(2002)083<0891:CCRDCW>2.3.CO;2).
- Fu, Q. and C. M. Johanson, 2004: Stratospheric influences on msu-derived tropospheric temperature trends: A direct error analysis. *Journal of Climate*, **17** (24), 4636–4640, URL <http://dx.doi.org/10.1175/JCLI-3267.1>.
- Fu, Q. and C. M. Johanson, 2005: Satellite-derived vertical dependence of tropical tropospheric temperature trends. *Geophys. Res. Lett.*, **32** (10).
- Fu, Q., C. M. Johanson, S. G. Warren, and D. J. Seidel, 2004: Contribution of stratospheric cooling to satellite-inferred tropospheric temperature trends. *Nature*, **429** (6987), 55–58.
- Fu, Q., S. Manabe, and C. M. Johanson, 2011: On the warming in the tropical upper troposphere: Models versus observations. *Geophys. Res. Lett.*, **38** (15).
- Gaffen, D. J., 1994: Temporal inhomogeneities in radiosonde temperature records. *J. Geophys. Res.*, **99** (D2), 3667–3676.
- Gaffen, D. J., B. D. Santer, J. S. Boyle, J. R. Christy, N. E. Graham, and R. J. Ross, 2000: Multidecadal changes in the vertical temperature structure of the tropical

- troposphere. *Science*, **287** (5456), 1242–1245, doi:10.1126/science.287.5456.1242, URL <http://www.sciencemag.org/content/287/5456/1242.abstract>, <http://www.sciencemag.org/content/287/5456/1242.full.pdf>.
- Haimberger, L., 2007: Homogenization of radiosonde temperature time series using innovation statistics. *Journal of Climate*, **20** (7), 1377–1403.
- Haimberger, L., C. Tavalato, and S. Sperka, 2008: Toward elimination of the warm bias in historic radiosonde temperature records—some new results from a comprehensive inter-comparison of upper-air data. *Journal of Climate*, **21** (18), 4587–4606.
- Haimberger, L., C. Tavalato, and S. Sperka, 2011: Homogenization of the global radiosonde dataset through combined comparison with reanalysis background series and neighboring stations. *Journal of Climate*.
- Hansen, J., R. Ruedy, M. Sato, and K. Lo, 2010: Global surface temperature change. *Rev. Geophys.*, **48** (4), URL <http://dx.doi.org/10.1029/2010RG000345>.
- Hegerl, G. C. and J. M. Wallace, 2002: Influence of patterns of climate variability on the difference between satellite and surface temperature trends. *Journal of Climate*, **15** (17), 2412–2428, URL [http://dx.doi.org/10.1175/1520-0442\(2002\)015<2412:IOP0CV>2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(2002)015<2412:IOP0CV>2.0.CO;2).
- Ho, S.-p., et al., 2009: Estimating the uncertainty of using gps radio occultation data for climate monitoring: Intercomparison of champ refractivity climate records from 2002 to 2006 from different data centers. *J. Geophys. Res.*, **114** (D23), URL <http://dx.doi.org/10.1029/2009JD011969>.
- Hurrell, J. W. and K. E. Trenberth, 1998: Difficulties in obtaining reliable temperature trends: Reconciling the surface and satellite microwave sounding unit records. *Journal of Climate*, **11** (5), 945–967.
- Hurrell, J. W., K. E. Trenberth, S. J. Brown, and J. R. Christy, 2000: Comparison of tropospheric temperatures from radiosondes and satellites: 1979–

98. *Bulletin of the American Meteorological Society*, **81 (9)**, 2165–2177, URL [http://dx.doi.org/10.1175/1520-0477\(2000\)081<2165:COTTFR>2.3.CO;2](http://dx.doi.org/10.1175/1520-0477(2000)081<2165:COTTFR>2.3.CO;2).
- IPCC, 2007: *Climate change 2007 : the physical science basis : contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge; New York.
- Johanson, C. M. and Q. Fu, 2006: Robustness of tropospheric temperature trends from msu channels 2 and 4. *Journal of Climate*, **19 (17)**, 4234–4242.
- Jones, P. D. and A. Moberg, 2003: Hemispheric and large-scale surface air temperature variations: An extensive revision and an update to 2001. *Journal of Climate*, **16 (2)**, 206–223, URL [http://dx.doi.org/10.1175/1520-0442\(2003\)016<0206:HALSSA>2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(2003)016<0206:HALSSA>2.0.CO;2).
- Kalnay, E., et al., 1996: The ncep/ncar 40-year reanalysis project. *Bulletin of the American Meteorological Society*, **77 (3)**, 437–471, URL [http://dx.doi.org/10.1175/1520-0477\(1996\)077<0437:TNYRP>2.0.CO;2](http://dx.doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2).
- Karl, T. R., S. J. Hassol, C. D. Miller, and W. L. Murray, (Eds.) , 2006: *Temperature Trends in the Lower Atmosphere: Steps for Understanding and Reconciling Differences*. Synthesis and Assessment Product 1.1, Climate Change Science Program and the Subcommittee on Global Change Research, Washington, DC.
- Klotzbach, P. J., S. Pielke, Roger A., J. Pielke, Roger A., J. R. Christy, and R. T. McNider, 2009: An alternative explanation for differential temperature trends at the surface and in the lower troposphere. *J. Geophys. Res.*, **114 (D21)**, URL <http://dx.doi.org/10.1029/2009JD011841>.
- Klotzbach, P. J., S. Pielke, Roger A., J. Pielke, Roger A., J. R. Christy, and R. T. McNider, 2010: Correction to &#8220;an alternative explanation for differential temperature trends at the surface and in the lower troposphere&#8221;. *J. Geophys. Res.*, **115 (D1)**, URL <http://dx.doi.org/10.1029/2009JD013655>.
- Lanzante, J. R., S. A. Klein, and D. J. Seidel, 2003a: Temporal homogenization of monthly

- radiosonde temperature data. part i: Methodology. *Journal of Climate*, **16** (2), 224–240, URL [http://dx.doi.org/10.1175/1520-0442\(2003\)016<0224:THOMRT>2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(2003)016<0224:THOMRT>2.0.CO;2).
- Lanzante, J. R., S. A. Klein, and D. J. Seidel, 2003b: Temporal homogenization of monthly radiosonde temperature data. part ii: Trends, sensitivities, and msu comparison. *Journal of Climate*, **16** (2), 241–262, URL [http://dx.doi.org/10.1175/1520-0442\(2003\)016<0241:THOMRT>2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(2003)016<0241:THOMRT>2.0.CO;2).
- Lu, Q., W. Bell, P. Bauer, N. Bormann, and C. Peubey, 2011: Characterising the fy-3a microwave temperature sounder using the ecmwf model. *Journal of Atmospheric and Oceanic Technology*.
- Manabe, S. and R. T. Wetherald, 1967: Thermal equilibrium of the atmosphere with a given distribution of relative humidity. *Journal of the Atmospheric Sciences*, **24** (3), 241–259, URL [http://dx.doi.org/10.1175/1520-0469\(1967\)024<0241:TEOTAW>2.0.CO;2](http://dx.doi.org/10.1175/1520-0469(1967)024<0241:TEOTAW>2.0.CO;2).
- Manabe, S. and R. T. Wetherald, 1975: The effects of doubling the co2 concentration on the climate of a general circulation model. *Journal of the Atmospheric Sciences*, **32** (1), 3–15, URL [http://dx.doi.org/10.1175/1520-0469\(1975\)032<0003:TEODTC>2.0.CO;2](http://dx.doi.org/10.1175/1520-0469(1975)032<0003:TEODTC>2.0.CO;2).
- McCarthy, M. P., H. A. Titchner, P. W. Thorne, S. F. B. Tett, L. Haimberger, and D. E. Parker, 2008: Assessing bias and uncertainty in the hadat-adjusted radiosonde climate record. *Journal of Climate*, **21** (4), 817–832.
- 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** (22), 3650–3664.
- Mears, C. A. and F. J. Wentz, 2005: The effect of diurnal correction on satellite-derived lower tropospheric temperature. *Science*, **309** (5740), 1548–1551, doi:10.1126/science.1114772.
- 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** (6), 1040 – 1056.

- Mears, C. A. and F. J. Wentz, 2009b: Construction of the rss v3.2 lower-tropospheric temperature dataset from the msu and amsu microwave sounders. *Journal of Atmospheric and Oceanic Technology*, **26 (8)**, 1493–1509.
- Mears, C. A. and F. J. Wentz, 2012: Assessing the value of satellite-radiosonde comparisons in ascertaining errors in climate data records. *J. Geophys. Res.*
- Mears, C. A., F. J. Wentz, P. Thorne, and D. Bernie, 2011: Assessing uncertainty in estimates of atmospheric temperature changes from msu and amsu using a monte-carlo estimation technique. *J. Geophys. Res.*, **116 (D8)**.
- Mo, T., 1994: A study of the microwave sounding unit on the noaa-12 satellite. *Geoscience and Remote Sensing Symposium, 1994. IGARSS '94. Surface and Atmospheric Remote Sensing: Technologies, Data Analysis and Interpretation., International*, Vol. 3, 1535–1537 vol.3, doi:10.1109/IGARSS.1994.399490.
- NRC, 2000: *Reconciling Observations of Global Temperature Change*. National Academy Press, Washington, D.C.
- Po-Chedley, S. and Q. Fu, 2012: A bias in the midtropospheric channel warm target factor on the noaa-9 microwave sounding unit. *Journal of Atmospheric and Oceanic Technology*, **29 (5)**, 646–652, URL <http://dx.doi.org/10.1175/JTECH-D-11-00147.1>.
- Prabhakara, C., J. Iacovazzi, R., J.-M. Yoo, and G. Dalu, 2000: Global warming: Evidence from satellite observations. *Geophys. Res. Lett.*, **27 (21)**, 3517–3520.
- Randall, R. M. and B. 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. *J. Geophys. Res.*, **113 (D5)**, URL <http://dx.doi.org/10.1029/2007JD008864>.
- Randel, W. J. and F. Wu, 2006: Biases in stratospheric and tropospheric temperature trends derived from historical radiosonde data. *Journal of Climate*, **19 (10)**, 2094–2104, URL <http://dx.doi.org/10.1175/JCLI3717.1>.

- Santer, B. D., et al., 2000: Interpreting differential temperature trends at the surface and in the lower troposphere. *Science*, **287** (5456), 1227–1232, doi:10.1126/science.287.5456.1227, URL <http://www.sciencemag.org/content/287/5456/1227.abstract>, <http://www.sciencemag.org/content/287/5456/1227.full.pdf>.
- Santer, B. D., et al., 2003a: Contributions of anthropogenic and natural forcing to recent tropopause height changes. *Science*, **301** (5632), 479–483, doi:10.1126/science.1084123, URL <http://www.sciencemag.org/content/301/5632/479.abstract>, <http://www.sciencemag.org/content/301/5632/479.full.pdf>.
- Santer, B. D., et al., 2003b: Influence of satellite data uncertainties on the detection of externally forced climate change. *Science*, **300** (5623), 1280–1284, doi:10.1126/science.1082393, URL <http://www.sciencemag.org/content/300/5623/1280.abstract>, <http://www.sciencemag.org/content/300/5623/1280.full.pdf>.
- Santer, B. D., et al., 2005: Amplification of surface temperature trends and variability in the tropical atmosphere. *Science*, **309** (5740), 1551–1556, doi:10.1126/science.1114867.
- Santer, B. D., et al., 2008: Consistency of modelled and observed temperature trends in the tropical troposphere. *International Journal of Climatology*, **28** (13), 1703–1722, doi:10.1002/joc.1756.
- Sherwood, S. C., 2007: Simultaneous detection of climate change and observing biases in a network with incomplete sampling. *Journal of Climate*, **20** (15), 4047–4062.
- Sherwood, S. C., J. R. Lanzante, and C. L. Meyer, 2005: Radiosonde daytime biases and late-20th century warming. *Science*, **309** (5740), 1556–1559, doi:10.1126/science.1115640.
- Sherwood, S. C., C. L. Meyer, R. J. Allen, and H. A. Titchner, 2008: Robust tropospheric warming revealed by iteratively homogenized radiosonde data. *Journal of Climate*, **21** (20), 5336–5352, URL <http://dx.doi.org/10.1175/2008JCLI2320.1>.

- Sobel, A. H., I. M. Held, and C. S. Bretherton, 2002: The enso signal in tropical tropospheric temperature. *Journal of Climate*, **15** (18), 2702–2706, URL [http://dx.doi.org/10.1175/1520-0442\(2002\)015<2702:TESITT>2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(2002)015<2702:TESITT>2.0.CO;2).
- Spencer, R. W. and J. R. Christy, 1990: Precise monitoring of global temperature trends from satellites. *Science*, **247** (4950), 1558–1562, doi:10.1126/science.247.4950.1558.
- Spencer, R. W. and J. R. Christy, 1992a: Precision and radiosonde validation of satellite gridpoint temperature anomalies. part i: Msu channel 2. *Journal of Climate*, **5** (8), 847–857.
- Spencer, R. W. and J. R. Christy, 1992b: Precision and radiosonde validation of satellite gridpoint temperature anomalies. part ii: A tropospheric retrieval and trends during 1979–90. *Journal of Climate*, **5** (8), 858–866.
- Stone, P. H. and J. H. Carlson, 1979: Atmospheric lapse rate regimes and their parameterization. *Journal of the Atmospheric Sciences*, **36** (3), 415–423, URL [http://dx.doi.org/10.1175/1520-0469\(1979\)036<0415:ALRRAT>2.0.CO;2](http://dx.doi.org/10.1175/1520-0469(1979)036<0415:ALRRAT>2.0.CO;2).
- Thorne, P. W., J. R. Lanzante, T. C. Peterson, D. J. Seidel, and K. P. Shine, 2011a: Tropospheric temperature trends: history of an ongoing controversy. *Wiley Interdisciplinary Reviews: Climate Change*, **2** (1), 66–88, doi:10.1002/wcc.80, URL <http://dx.doi.org/10.1002/wcc.80>.
- Thorne, P. W., D. E. Parker, J. R. Christy, and C. A. Mears, 2005a: Uncertainties in climate trends: Lessons from upper-air temperature records. *Bulletin of the American Meteorological Society*, **86** (10), 1437–1442.
- Thorne, P. W., D. E. Parker, S. F. B. Tett, P. D. Jones, M. McCarthy, H. Coleman, and P. Brohan, 2005b: Revisiting radiosonde upper air temperatures from 1958 to 2002. *J. Geophys. Res.*, **110** (D18).
- Thorne, P. W., et al., 2007: Tropical vertical temperature trends: A real discrepancy? *Geophys. Res. Lett.*, **34** (16).

- Thorne, P. W., et al., 2011b: A quantification of uncertainties in historical tropical tropospheric temperature trends from radiosondes. *J. Geophys. Res.*, **116** (D12), URL <http://dx.doi.org/10.1029/2010JD015487>.
- Titchner, H. A., P. W. Thorne, M. P. McCarthy, S. F. B. Tett, L. Haimberger, and D. E. Parker, 2009: Critically reassessing tropospheric temperature trends from radiosondes using realistic validation experiments. *Journal of Climate*, **22** (3), 465–485, URL <http://dx.doi.org/10.1175/2008JCLI2419.1>.
- Trenberth, K. E. and J. W. Hurrell, 1997: How accurate are satellite ‘thermometers’? *Nature*, **389** (6649), 342–342.
- Wentz, F. J. and M. Schabel, 1998: Effects of orbital decay on satellite-derived lower-tropospheric temperature trends. *Nature*, **394** (6694), 661–664.
- Wentz, F. J. and M. Schabel, 2000: Precise climate monitoring using complementary satellite data sets. *Nature*, **403** (6768), 414–416, URL <http://dx.doi.org/10.1038/35000184>.
- 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** (7), 1661–1681.
- Zou, C.-Z., M. D. Goldberg, Z. Cheng, N. C. Grody, J. T. Sullivan, C. Cao, and D. Tarpley, 2006: Recalibration of microwave sounding unit for climate studies using simultaneous nadir overpasses. *J. Geophys. Res.*, **111** (D19).
- Zou, C.-Z. and W. Wang, 2009: Diurnal drift correction in the nesdis/star msu/amsu atmospheric temperature climate data record. *Proc. SPIE*, **7456** (745616).
- Zou, C.-Z. and W. Wang, 2010: Stability of the msu-derived atmospheric temperature trend. *Journal of Atmospheric and Oceanic Technology*, **27** (11), 1960–1971.
- Zou, C.-Z. and W. Wang, 2011: Inter-satellite calibration of amsu-a observations for weather and climate applications. *J. Geophys. Res.*

## Appendix A

## UAH RESPONSE TO THIS WORK

***A.1 Our Response to Recent Criticism of the UAH Satellite Temperatures***

May 9th, 2012

by John R. Christy and Roy W. Spencer

University of Alabama in Huntsville

A new paper by Stephen Po-Chedley and Quang Fu (2012) (hereafter PCF) was sent to us at the end of April 2012 in page-proof form as an article to appear soon in the Journal of Atmospheric and Oceanic Technology. The topic of the paper is an analysis of a single satellite's impact on the rarely-used, multi-satellite deep-layer global temperature of the mid-troposphere or TMT. Some of you have been waiting for our response, but this was delayed by the fact that one of us (J. Christy) was out of the country when the UW press release was issued and just returned on Tuesday the 8th.

There are numerous incorrect and misleading assumptions in this paper. Neither one of us was aware of the paper until it was sent to us by Po-Chedley two weeks ago, so the paper was written and reviewed in complete absence of the authors of the dataset itself. In some cases this might be a normal activity, but in a situation where complicated algorithms are involved, it is clear that PCF did not have a sufficient understanding of the construction methodology.

By way of summary, here are our main conclusions regarding the new PCF paper:

- 1) the authors' methodology is qualitative and irreproducible
- 2) the authors are uninformed on the complexity of the UAH satellite merging algorithm
- 3) the authors use the RSS (Remotes Sensing Systems) satellite dataset as verification for their proposed UAH NOAA-9 calibration target adjustment for TMT, but barely mention that their TLT (lower tropospheric) results are insignificant and that trends are essentially

identical between UAH and RSS without any adjustment in the NOAA-9 calibration coefficient

4) the authors neglected the main TMT differences among the datasets - and instead try to explain the UAH v. RSS trend difference by only two years of NOAA-9 data, while missing all of the publications which document other issues such as RSS problems with applying the diurnal correction.

The paper specifically claims to show that a calibration target coefficient of one satellite, NOAA-9, should be a value different than that calculated directly from empirical data in UAHs version of the dataset. With an adjustment to the time series guesstimated by PCF, this increases the UAH overall global trend by +0.042 C/decade. Their new UAH trend, being +0.042 warmer, then becomes the same as the TMT trend from RSS. This, they conclude, indicates a verification of their exercise.

More importantly, with regard to the most publicized UAH dataset, the temperature of the lower troposphere (TLT), there was no similar analysis done by PCF - an indication that their re-calculations would not support their desired outcome for this dataset, as we shall demonstrate below.

All of this will soon be moot, anyway. Since last year we have been working on v6.0 of the UAH datasets which should be ready with the tropospheric temperature datasets before summer is out. These will include (1) a new, more defensible objective empirical calculation to correct for the drift of the satellites through the diurnal cycle, and (2) a new hot calibration target effective emissivity adjustment which results in better agreement between simultaneously operating satellites at the calibration step, making the post-calibration hot-target adjustment PCF criticizes unnecessary. So, since our new v6.0 dataset is close to completion and submission for publication, we have chosen this venue to document PCFs misinformation in a rather informal, but reproducible, way rather than bother to submit a journal rebuttal addressing the older dataset. However, to show that version 5.4 of our datasets was credible, we discuss these issues below.

## ***A.2 The Lower Tropospheric Temperatures (TLT)***

We shall return to TMT below, but most of the research and popular use of the UAH datasets have focused on the lower tropospheric temperature, or TLT (surface to about 300 hPa, i.e. without stratospheric impact). Thus, we shall begin our discussion with TLT because it is rightly seen as a more useful variable because it documents the bulk heat content of the troposphere with very little influence from the stratosphere. And [this is important in the TMT discussion] the same hot-target coefficients for NOAA-9 were used in TLT as in TMT.

PCF focused on the deep layer TMT, i.e. temperature of the surface to about 75 hPa, which includes quite a bit of signal above 300 hPa. As such, TMT includes a good portion of the lower stratosphere - a key weakness when utilizing radiosondes which went through significant changes and adjustments during this time. [This was a period when many stations converted to the Vaisala 80 radiosonde which introduced temperature shifts throughout the atmosphere (Christy and Norris 2004).]

As indicated in their paper, it seems PCF's goal was to explain the differences in trend between RSS and UAH, but the history of this effort has always been to find error with UAH's products rather than in other products (as we shall see below). With us shut out of the peer-review cycle it is easy to assume an underlying bias of the authors.

Lord Kelvin told us that All science is numbers, so here are some numbers. First, let's look at the global trends of UAH and RSS for TLT (70S to 82.5N) for Jan 1979 to Apr 2012:

$$+0.137 \text{ C/decade UAH LT (70S-82.5N)} \quad +0.134 \text{ C/decade RSS LT (70S-82.5N)}$$

These trends are, for all practical purposes, identical. This, however, hides the fact that there are indeed differences between the two time series that, for one reason or another, are balanced out when calculating the linear trend over the entire 30+ year period. As several papers have documented (see Christy et al. 2011, or C11, for the list - by the way, C11 was not cited by PCF) the evidence indicates RSS contains a spurious warming in the 1990s then a spurious cooling from around 2002 onward (note that the RSS temperature anomaly for last month, April, 2012, was 0.08C cooler than our UAH anomaly).

This behavior arises, we believe, from an over-correction of the drift of the satellites by RSS (in the 1990s the satellites drifted to cooler times of day, so the correction must add warming, and in the 2000s the satellites drifted to warmer times of day so a correction is needed to cool things down.) These corrections are needed (except for the Aqua satellite operating since 2002, which has no diurnal drift and which we use as an anchor in the UAH dataset) but if not of the right magnitude they will easily affect the trend.

In a single paragraph, PCF admit that the UAH TLT time series has no significant hot-target relationship with radiosonde comparisons (which for TLT are more robust) over the NOAA-9 period. However, they then utilize circular reasoning to claim that since RSS and UAH have a bit of disagreement in that 2-year period, and RSS must be correct, that then means UAH has a problem. So, this type of logic, as stated by PCF, points to their bias - assume that RSS is correct which then implies UAH is the problem. This requires one to ignore the many publications that show the opposite.

Note too that in their press release, PCF claim that observations and models now are closer together for this key parameter (temperature of the bulk troposphere) if one artificially increases the trend in UAH data. This is a questionable claim as evidence shows TLT for CMIP3 and CMIP5 models averages about +0.26 C/decade (beginning in 1979) whereas UAH \*and\* RSS datasets are slightly below +0.14 C/decade, about a factor of 2 difference between models and observations. We shall let the reader decide if the PCF press-release claim is accurate.

The key point for the discussion here (and below) is that TLT uses the same hot-target coefficients as TMT, yet we see no problem related to it for the many evaluation studies we have published. Indeed this was the specific result found in Christy and Norris 2004 - again, work not cited by PCF.

### ***A.3 The Mid-Tropospheric Temperature (TMT)***

About 12 years ago we discovered that even though two different satellites were looking at the same globe at the same time, there were differences in their measurements beyond a simple bias (time-invariant offset). We learned that these were related to the variations in the temperature of the instrument itself. If the instrument warmed or cooled (differing solar

angles as it orbited or drifted), so did the calculated temperature. We used the thermistors embedded in the hot-target plate to track the instrument temperature, hence the metric is often called the hot target temperature coefficient.

To compensate for this error, we devised a method to calculate a coefficient that when multiplied by the hot target temperature would remove this variation for each satellite. Note that the coefficients were calculated from the satellite data, they were not estimated in an ad hoc fashion.

The calculation of this coefficient depends on a number of things, (a) the magnitude of the already-removed satellite drift correction (i.e. diurnal correction), (b) the way the inter-satellite differences are smoothed, and (c) the sequence in which the satellites are merged.

Since UAH and RSS perform these processes differently, the coefficients so calculated will be different. Again recall that the UAH (and RSS) coefficients are calculated from a system of equations, they are not invented. The coefficients are calculated to produce the largest decrease in inter-satellite error characteristics in each dataset.

To make a long story short, PCF focused on the 26-month period of NOAA-9 operation, basically 1985-86. They then used radiosondes over this period to estimate the hot-target coefficient as +0.048 rather than UAHs calculated value of +0.0986. [Note, the language in PCF is confusing, as we cannot tell if they conclude our coefficient is too high by 0.051 or should actually be 0.051. We shall assume they believe our coefficient is too high by 0.051 to give them the benefit of the doubt.]

Recall, radiosondes were having significant shifts with the levels monitored by TMT primarily with the switch to Vaisala 80 sondes, and so over small, 26-month periods, just about any result might be expected. [We reproduced PCF's Fig. 2 using only US VIZ sondes (which had no instrument changes in the 26-month period and span the globe from the western tropical Pacific to Alaska to the Caribbean Sea) and found an explained variance of less than 4% - an insignificant value.]

Another problematic aspect of PCF's methodology is that when looking at the merged time series, one does not see just NOAA-9s influence, but the impact of all of the other satellites which provided data during 1985-86, i.e. NOAA-6, -7 and -8 as well. So, it

is improper to assume one may pick out NOAA-9s impact individually from the merged satellite series.

That PCF had little understanding of the UAH algorithm is demonstrated by the following simple test. We substituted the PCF value of +0.048 directly into our code. The increase in trend over our v5.4 TMT dataset was only +0.022 C/decade for 1979-2009 (not 0.042), and +0.019 C/decade for 1979-2012.

To put it another way, PCF overestimated the impact of the NOAA-9 coefficient by a factor of about 2 when they artificially reconstructed our dataset using 0.048 as the NOAA-9 coefficient. In fact, if we use an implausible target coefficient of zero, we still cant return a trend difference greater than +0.037 C/decade. Thus PCF have incorrectly assumed something about the construction methodology of our time series that gave them a result which is demonstrated here to be faulty.

In addition, by changing the coefficient to +0.048 in an ad hoc fashion, they create greater errors in NOAA-9s comparisons to other satellites. Had they contacted us at any point about this, we would have helped them to understand the techniques. [There were 4 emails from Po-Chedley in Aug and Sep 2011, but this dealt with very basic facts about the dataset, not the construction methodology. Incidentally, these emails were exchanged well after C11 was published.]

PCF brought in a third dataset, STAR, but this one uses the same diurnal corrections and sequential merging methodology as RSS, so it is not a truly independent test. As shown in C11, STAR is clearly the outlier for overall trend values due to a different method of debiasing the various satellite data and a differing treatment of the fundamental brightness temperature calibration.

We have additional information regarding UAHs relatively low error statistics. Using radiosondes to evaluate microwave temperatures requires great care. In our tests, we concentrated on sondes which had documented characteristics and a high degree of consistency such as the US VIZ and Australian sondes. These comparisons have been published a number of times, but most recently updated in C11.

Here are the comparisons for the US VIZ radiosonde network (stretching from the western tropical Pacific to Alaska down across the conterminous US and to the Caribbean.)

Viz Sondes	Monthly St.	Annual St.	Monthly $r^2$	Annual $r^2$
TMT	Dev. Difference K	Dev. Difference K	Composite	Composite
UAH	0.088	0.037	0.90	0.96
RSSv3.2	0.104	0.065	0.89	0.91
NOAAv2.0	0.102	0.065	0.89	0.91

Table A.1: MSU - Radiosonde error characteristics presented by the UAH team.

As you can see, UAH MT provides the lowest error magnitudes and highest reproducibility of the three data sets. Similar results were found for the Australian comparisons.

For data through April 2012 we have the following global TMT trends: UAH +0.045, RSS +0.079 and STAR +0.124 C/decade. So, RSS, in the middle, is closer to UAH than STAR, yet PCF chose to examine UAH as the problem dataset. Had PCF wanted to pick some low-hanging fruit regarding the differences between UAH, RSS and STAR, they would have (a) looked at the diurnal differences between UAH and RSS (see publications) or (b) looked at a simple time series of differences between the three datasets (below). One thing that pops out is a spurious upward shift in STAR TMT relative to UAH and RSS of about +0.06 C on precisely 1 Jan 2001 - an obvious beginning-of-year glitch. Why not look there?

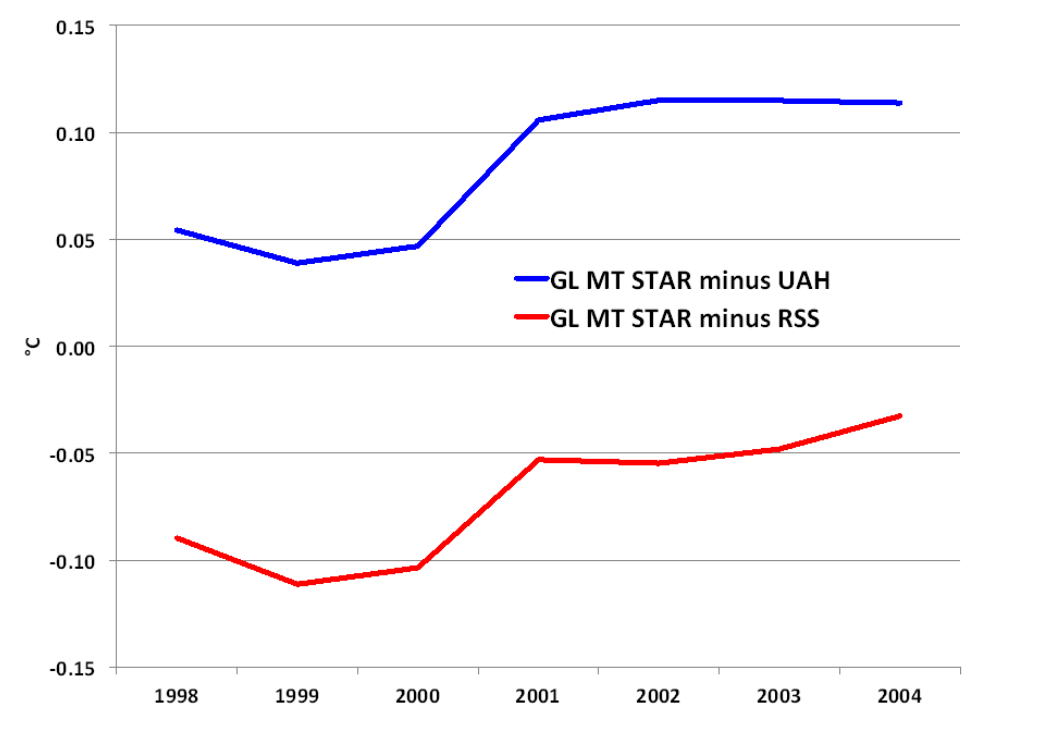


Figure A.1: Global MSU-MSU TMT difference time series.

#### A.4 *The Bottom Line*

In conclusion, we believe that the result in PCF was a rather uninformed attempt to find fault with the UAH global temperature dataset, using an ad hoc adjustment to a single, short-lived satellite while overlooking the greater problems which have been documented (published or as demonstrated in the figure above) regarding the other datasets.

And think about this. If PCF is correct that we should be using a revised NOAA-9 coefficient, and since we use the same coefficient in both TMT and TLT, then the near perfect agreement currently between RSS and UAH for TLT will disappear; our TLT trend will become warmer, and then RSS will have the lowest warming trend of all the satellite datasets. The authors of the new study cannot have it both ways, claiming their new

adjustment brings RSS and UAH closer together for TMT (a seldom used temperature index), but then driving the UAH and RSS trends for TLT farther apart, leaving RSS with essentially the same warming trend that UAH had before.

Since it is now within 3 months of the publication cutoff for research to be included in the IPCC AR5, one is tempted to conclude that PCF will be well-received by the Lead Authors (some of whom are closely associated with the RSS dataset) without critical evaluation such as briefly performed here. However, we cannot predict what the AR5 outcome will be or, for that matter, what waning influence the IPCC might still exert.

That PCF brushed aside the fact that the UAH and RSS trends for the LOWER troposphere are essentially identical (for which the UAH NOAA-9 coefficient is the same) seems to us to be a diversionary tactic we have seen before: create a strawman problem which will allow the next IPCC report to make a dismissive statement about the validity of an uncooperative dataset with a minimum of evidence. We hope that rationality instead prevails.

### **A.5 References**

Christy, J.R. and W. B. Norris, 2004: What may we conclude about global tropospheric temperature trends? *Geophys. Res. Lett.* 31, No. 6.

Christy, J.R., R.W. Spencer and W.B Norris (deceased), 2011: The role of remote sensing in monitoring global bulk tropospheric temperatures. *Int. J. Remote Sens.* 32, 671-685, DOI:10.1080/01431161.2010.517803.

Gohring, Nancy: FAQ: New research brings satellite measurements and global climate models closer. *UW Today.* 7 May, 2012 <<http://www.washington.edu/news/articles/satellite-temperature-measurements-faq>>.

Gohring, Nancy: New research brings satellite measurements and global climate models closer. *UW Today.* 7 May, 2012 <<http://www.washington.edu/news/articles/new-research-brings-satellite-measurements-and-global-climate-models-closer>>.

Po-Chedley, S. and Q. Fu, 2012: A bias in the midtropospheric channel warm target factor on the NOAA-9 Microwave Sounding Unit. *J. Atmos. Oceanic Tech.* DOI: 10.1175/JTECH-D-11-00147.1.

## Appendix B

### ADDRESSING CRITICISMS OF THE UAH TEAM

#### *B.1 Into context*

It has long been recognized that NOAA-9 represents a critical link in the MSU/AMSU TMT temperature merger (Christy et al., 1998; Mears et al., 2003; Karl et al., 2006). To quote the CCSP Report:

The difference between trends for T<sub>2</sub> [TMT] has received considerable attention. A close examination of the procedures suggests that about 50% of the discrepancy in trends is accounted for by a difference between the target factor for the NOAA-09 instrument deduced by the two groups. This difference mainly arises from the subsets of data used by the two groups when determining the satellite merging parameters (i.e., offsets and target factors). The UAH group emphasizes pairs of satellites that have long periods of overlap, and thus uses six pairs of satellites, while RSS uses all available (12) overlapping pairs of satellites.

Importantly, Dr. John Christy and Dr. Carl Mears were both authors of this report. During the development of this thesis, the global radiosonde datasets that we utilized consistently showed significant negative drifts in the UAH group during the NOAA-9 time period, utilizing T24 as a reference (which minimizes the influence of the stratosphere). T24 unfortunately is influenced by both channels TMT and TLS with about 90% of the information coming from the TMT channel. We wanted to isolate the problematic channel, TMT, which was known to be a cause of important differences between UAH and RSS/NOAA. We therefore focused our analysis on the TMT product (though we obtain similar results with T24, which effectively removes the influence of the stratosphere).

Figure 3.2 demonstrates that the differences over the NOAA-9 time period are related to differences in the instrument body calibration, which we would expect given that the

UAH target factor is about two times larger than that of RSS and NOAA (Figure 3.1). While Drs. Spencer and Christy are correct in pointing out that the instrument body effect and the diurnal drift correction are related, the relationship is not linear and given that the regressed differences between MSU groups match the differences in the target factor (i.e. the slopes in Figure 3.2 are approximately the same size as the differences in the actual target factors used), it is evident that the diurnal drift correction is not the main problem during the NOAA-9 time period, especially since the authors of the RSS and UAH dataset came to the same conclusion in the CCSP Report.

With a solid basis to suspect that a major difference (but not the only difference) between these groups is related to the instrument body effect, we decided to use radiosondes as an arbiter. We used five global (differences in target factors affect the instrument calibration nearly uniformly across the globe) radiosonde datasets that have undergone intensive and peer reviewed homogenization procedures. Furthermore, our method does not simply compare the evolution of radiosonde temperatures compared to that of MSU/AMSU temperatures. This is a method utilized in other studies (e.g. Christy and Norris, 2004), which come to different conclusions utilizing the TLT channel (and then assume that the conclusions apply to TMT). In our method, we are attempting to detect residuals related to the warm target temperature of the satellite, which, if present, suggest a mis-calibration of the satellite. We assumed that errors in the global mean radiosonde data are unrelated to the evolution of the warm target temperature. It is possible that this assumption is not true, but all five radiosonde datasets we utilized demonstrate a consistent story (UAH over-estimated its target factor) and the magnitude of the bias as measured by the radiosondes is consistent with the target factor differences between UAH and RSS/NOAA. We chose this method because it would help us identify the calibration problem at its root. We are also looking at the calibration bias in TMT directly. In Christy and Norris (2004), the time periods 1987 - 1990 and 1979 - 1982 are compared for UAH and RSS TLT versus radiosondes and it is determined that UAH **TLT** is more consistent (and therefore **TMT** for UAH likely has no calibration problem). This is important because radiosondes can contain undocumented and time varying biases, especially over long periods of time, so comparing radiosonde records on long timescales can be misleading. Further, TLT also has

a much larger diurnal drift correction, which makes it impossible to assume that both TMT and TLT are properly calibrated by analyzing TLT in isolation. Drs. Christy and Spencer repeated our analysis for US radiosondes which they assert are better references and get an insignificant result. This suggests that either 1) NOAA and RSS may have mis-calibrated their target factor and all five of the global formally homogenized datasets contain a spurious drift that happens to be related to UAH minus Radiosonde or 2) the US radiosondes that are used by Drs. Christy and Spencer lacks adequate sampling or have not properly been homogenized. We will demonstrate that the latter of these two possibilities is likely.

We measure similar MSU/AMSU versus radiosonde error characteristics compared to those reported by UAH (Appendix A). In Table B.1, we present the error characteristics measured in our global radiosonde datasets for each MSU/AMSU TMT product. The major difference is that our annual standard deviation of MSU minus radiosondes is similar for UAH and RSS/NOAA. UAH reports an annual standard deviation of differences that is about half as large for UAH compared to RSS and NOAA.

Table B.1: TMT error characteristics for the different MSU groups compared to the radiosonde references used in this work. For this comparison we used detrended, collocated time series from 1979 - 2009 and averaged the results for the five radiosonde datasets.

MSU - Sonde	Monthly St.	Annual St.	Monthly $r^2$	Annual $r^2$
TMT	Dev. Difference K	Dev. Difference K	Composite	Composite
UAH	0.086	0.063	0.89	0.87
RSSv3.2	0.090	0.068	0.88	0.85
NOAAv2.0	0.092	0.071	0.89	0.84

The UAH team attempted to replicate our analysis using “US Viz Sondes” and found that they could not detect a significant bias in their TMT product. Their results are presented in Figure B.1. The UAH team utilized these radiosondes because they believe they are consistent throughout the NOAA-9 time period. If they utilized the same set of radiosondes as in Christy and Norris (2006), this would be a set of 31 radiosondes.

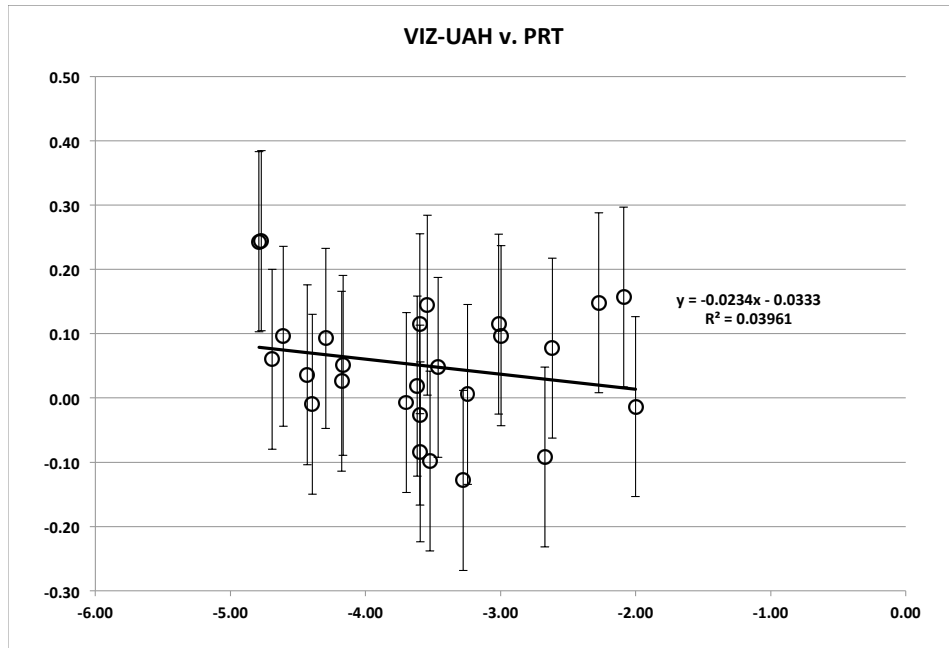


Figure B.1: UAH team replicating our procedure using US Viz Sondes (Dr. John Christy, personal communication, 2012). The x-axis represents the target temperature anomaly and the y-axis is UAH-radiosonde.

We found that our results are dependent on the sample size, which may be the reason for the discrepancy with UAH. In Figure B.2, we demonstrate the effect of sample size on our statistics. In this figure, we randomly sampled our 384 radiosonde stations from the RICH dataset, collocated those stations with the UAH product, and compared the mean UAH - RICH time series with the NOAA-9 target temperature time series over the NOAA-9 time period for different sample sizes. For each sample size, we repeated the calculation 1,000 times. The results are similar for RAOBCORE V1.4, RATPAC-B, HadAT2, and IUUK. This demonstrates that a large station sample size is needed to reduce the radiosonde noise

sufficiently enough to measure the bias in the NOAA-9 target factor, which might explain why the UAH comparison with US radiosondes yields different results. In other words, a large number of radiosondes are needed to detect this calibration error.

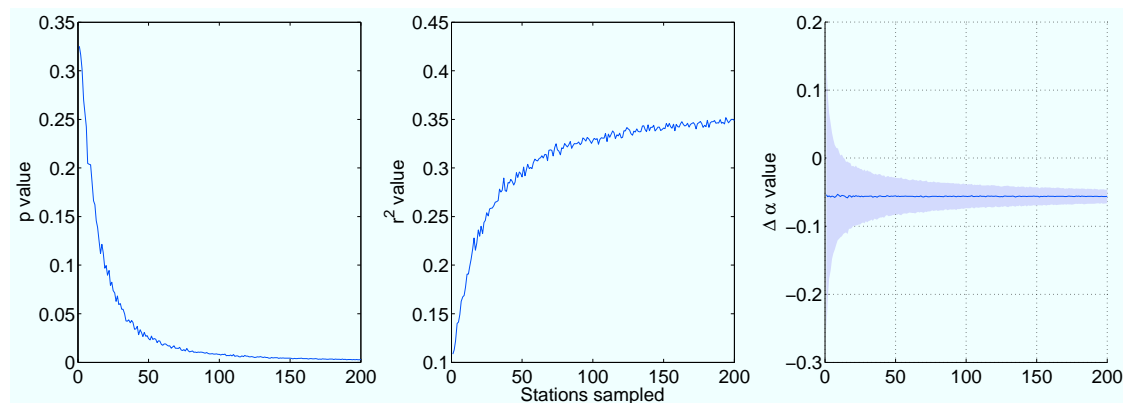


Figure B.2: Effects of sample size on statistics in this analysis. We computed the  $r^2$  value, p-value, and  $\Delta\alpha$  value for the regression of UAH - RICH Radiosondes versus  $T_{Target}$  for different numbers stations (collocated with UAH data over the NOAA-9 time period). For this calculation we randomly sub-sample a certain number of stations (x-axis) and create an average, collocated UAH-Radiosonde difference time series, which we then regress against the warm target temperature. We redo this calculation 1,000 times for each number of stations sampled and then present the mean p-value,  $r^2$  value, and  $\Delta\alpha$  value. The shaded region around the  $\Delta\alpha$  value is the 95% confidence interval from the sub-sampling statistics only; it does not contain the error in the regression itself. The results become significant when about 35 stations are included in the global average; below this number, the signal to noise ratio is too low.

An excellent point made by the UAH team is that our analysis is based on the UAH merged product, which includes the influence of NOAA-6, NOAA-7, NOAA-8, NOAA-9, and NOAA-10. Importantly, most of these satellites are incorporated into the time series for only a short period of time, but nonetheless may influence the result. It has been suggested (Dr. Carl Mears, personal communication, 2012) that we repeat the analysis using only NOAA-9 (corrected for the diurnal drift and the instrument body effect). This would be a worthwhile next step if UAH is willing to provide this data, but it is compelling that residuals related to the NOAA-9 target temperature are in the merged product (NOAA-9

has the most influence on the UAH TMT anomalies during this time period). If we are correct that NOAA-9 is not calibrated correctly, utilizing the NOAA-9 data in isolation could enhance our case.

One point made by UAH is that we have focused on UAH, when a number of other studies indicate that RSS and NOAA may have problems (e.g. Christy and Norris, 2004; Christy et al., 2010, 2011). Some of the claims made about RSS and NOAA having problems with their diurnal cycle adjustment should be explored further, but they were not the focus of this work and are irrelevant to the results presented here.

Another implied assumption is that we are attempting to find errors with the UAH dataset to match the UAH trend with the RSS trend. This is not true. The NOAA-9 TMT problem was well motivated in the literature and it is apparent in the latest MSU/AMSU datasets that the warm target factor for NOAA-9 helps explain differences between the datasets. We believe that the UAH TMT trend should increase when this bias is taken into account, but it is possible that the trend value could be below that of RSS (or it is possible that the value could exceed that of RSS); we do not believe that any of these possibilities discredits the bias presented here. Similarly, our trend estimate was a simple estimate and it was noted that “while our estimate of the UAH TMT trend sensitivity to the warm target factor is similar to estimates by the UAH and RSS teams, our study indicates that the UAH team will need to incorporate an optimal NOAA-9 target factor into their merging procedure for an accurate trend estimate” (Po-Chedley and Fu, 2012). We therefore are not concerned that their trend sensitivity does not match ours, although we note that our estimate compared well to previous estimates of the impact of this difference on the TMT trend.

Much of the criticism from UAH hinges on TLT, which was a secondary aspect of this work. The NOAA-9 TMT target factor difference was well established. We contend that this difference results because of a bias in the UAH calibration procedure. Since UAH uses the TMT target factor for TLT, this bias must effect the TLT product if the NOAA-9 target factor bias is real (even if the trend effect is different). This is true even if this would make the UAH TLT trend larger than that of RSS. One reason we may not have been able to measure the bias in the TLT product is because TLT time series has greater noise (Hurrell

and Trenberth, 1998; Mears and Wentz, 2009b; Christy et al., 2000). This is established by the UAH team: “Since these are linear operations, we will apply the  $T_W [\alpha]$  coefficients for  $T_2$  [TMT] to  $T_{2LT}$  [TLT] because  $T_{2LT}$  [TLT] has greater noise than  $T_2$  [TMT] due to the retrieval algorithm” (Christy et al., 2000). Just as in our work, Christy et al. (2000) found that there is too much noise in the TLT time series to obtain reliable warm target coefficients. Mears and Wentz (2009b) similarly discusses the complications of determining the warm target coefficients for TLT (though RSS does attempt to determine warm target coefficients for TLT).

Another criticism was that we utilized RSS as a reference for UAH. Our estimate of this bias comes solely from radiosondes. We compare the three satellite datasets, but our work is not justified by the closer agreement between the satellite groups. With that said, we did aim to reconcile the trend differences by looking for real biases (in any of the datasets) in keeping with the recommendations of the CCSP Report (Karl et al., 2006). Some of the MSU/AMSU differences posited in other works have been difficult to verify robustly, but this does not mean that we have not noticed differences during other periods.

Last, we did not intend for UAH to simply reduce its target factor. It is well understood that the satellites are merged in such a way to reduce the errors between overlapping satellites. Mears et al. (2003) noted that the difference trends between satellite pairs not considered in the UAH merging procedure are substantially enhanced with the large UAH target factor. Accounting for the NOAA-9 bias presented here might have a cascading effect that influences a number of satellite merging parameters. It is possible that this bias results because UAH does not use enough satellite pairs to constrain its merging parameters. Including more satellite pairs could further improve the UAH product. We have simply identified a problem that has important trend implications for the UAH team; only UAH can decide if and how to account for this bias.

## ***B.2 Summary***

- **UAH Claim:** “The authors methodology is qualitative and irreproducible.”

This is probably in reference to two claims made by UAH:

1) **UAH Claim:** “Recall, radiosondes were having significant shifts with the levels monitored by TMT primarily with the switch to Vaisala 80 sondes, and so over small, 26-month periods, just about any result might be expected. [We reproduced PCFs Fig. 2 using only US VIZ sondes (which had no instrument changes in the 26-month period and span the globe from the western tropical Pacific to Alaska to the Caribbean Sea) and found an explained variance of less than 4% - an insignificant value.]”

**Response:** UAH used a different radiosonde dataset and could not reproduce our measurement for the warm target bias. This does not mean our work is not reproducible. We show that a large sample of radiosondes is needed to detect the warm target bias, which we believe is a shortcoming of the UAH comparison to our work. Furthermore, when we use T24 to measure this bias, we obtain the same basic results, indicating that radiosonde discontinuities that largely effect the stratosphere did not influence our result.

2) **UAH Claim:** “We substituted the PCF value of +0.048 directly into our code. The increase in trend over our v5.4 TMT dataset was only +0.022 C/decade for 1979-2009 (not 0.042).”

**Response:** Our trend estimate, albeit simple, was in the range of previously published sensitivity experiments for the NOAA-9 target factor problem. Although our estimate may be too large (if the effective drift in the merged time series is between the NOAA-6/NOAA-9 overlap to the NOAA-9/NOAA-10 overlap), the full effect cannot be measured without UAH altering its merging procedure so that they minimize error residuals between overlapping satellites **and** remove the bias that we identified.

- **UAH Claim:** “The authors use the RSS (Remotes Sensing Systems) satellite dataset as verification for their proposed UAH NOAA-9 calibration target adjustment for TMT.”

The correction for the instrument body effect should not be different for UAH and RSS by a factor of two. This calibration is meant to resolve changes in the radiometer calibration. We use *radiosondes* from five, peer-reviewed datasets to 1) attribute the bias and 2) estimate the magnitude of this bias. Since we expect that an unbiased calibration coefficient should be the same for different MSU/AMSU teams, we compare the value to that of RSS and NOAA and find that a properly calibrated warm target factor is similar to their values (for which we detect no significant bias).

- **UAH Claim:** “[The authors] barely mention that their TLT (lower tropospheric) results are insignificant and that trends are essentially identical between UAH and RSS without any adjustment in the NOAA-9 calibration coefficient.”

We don’t pre-suppose that UAH and RSS trends should be the same upon accounting for this effect. It seems obvious that there are other important differences between the two time series that need to be resolved. So agreement between UAH and RSS TLT products does not mean that both groups are correct. The agreement could be a result of offsetting errors. The focus of this paper was on TMT, although we discuss reasons for an insignificant result in TLT. The title of the paper is “A Bias in the Midtropospheric Channel Warm Target Factor on the NOAA-9 Microwave Sounding Unit” and almost entirely discusses TMT. The paper has implications for TLT, because UAH uses the TMT warm target factor for TLT.

- **UAH Claim:** “The authors neglected the main TMT differences among the datasets and instead try to explain the UAH v. RSS trend difference by only two years of NOAA-9 data, while missing all of the publications which document other issues such as RSS problems with applying the diurnal correction.”

The NOAA-9 difference has been highlighted as an important difference in the literature by both RSS and UAH. We sought to help resolve this difference. Our results demonstrate that for all of the major radiosonde datasets, UAH contains a

significant bias for this specific problem (the calibration of the NOAA-9 satellite). This work helps resolve an important difference, but it does not apply to other time periods or calibration issues. After correcting for this bias, there are likely residual biases in the tropics due to differences in the diurnal drift calibration. Many of the papers that UAH alludes to regarding the diurnal drift are largely irrelevant to the NOAA-9 problem, though resolving other calibration problems should help reconcile differences between the two groups.

- **UAH Claim:** “With an adjustment to the time series guesstimated by PCF, this increases the UAH overall global trend by +0.042 C/decade. Their new UAH trend, being +0.042 warmer, then becomes the same as the TMT trend from RSS. This, they conclude, indicates a verification of their exercise.”

The good trend agreement between RSS and UAH after accounting for this bias does not verify our work. We do not make this claim. We expect that accounting for this error would increase agreement between datasets, since this has been highlighted (multiple times) as a structural uncertainty. By identifying a bias in the UAH merging procedure, we effectively remove this structural uncertainty; trend agreement between the groups should subsequently improve.

- **UAH Claim:** “More importantly, with regard to the most publicized UAH dataset, the temperature of the lower troposphere (TLT), there was no similar analysis done by PCF an indication that their re-calculations would not support their desired outcome for this dataset.”

We did not measure a statistically significant bias for UAH TLT. This could be because noise in the TLT product is 2 - 3 times higher than in TMT or it might be a result of a conflicting bias with the diurnal cycle correction, which is much larger for the TLT product compared to TMT. We know that this would have some effect on TLT, because UAH *a priori* apply the target factor calculated for TMT on TLT

(and not the other way around). So a bias in the target factor in TMT, must effect the TLT product. We expect that the TLT trend would also increase, possibly such that the UAH global trend would be larger than the RSS global trend.

- **UAH Claim:** “Since last year we have been working on v6.0 of the UAH datasets which should be ready with the tropospheric temperature datasets before summer is out. These will include (1) a new, more defensible objective empirical calculation to correct for the drift of the satellites through the diurnal cycle, and (2) a new hot calibration target effective emissivity adjustment which results in better agreement between simultaneously operating satellites at the calibration step, making the post-calibration hot- target adjustment PCF criticizes unnecessary. So, since our new v6.0 dataset is close to completion and submission for publication, we have chosen this venue to document PCFs misinformation in a rather informal, but reproducible, way rather than bother to submit a journal rebuttal addressing the older dataset. However, to show that version 5.4 of our datasets was credible, we discuss these issues below.”

We encourage UAH to publish their criticisms of our work in a peer-reviewed journal, especially if residuals related to the warm target temperature remain in the merged time series (in version 6.0). We maintain that this is a bias that artificially decreases the UAH trend. Further, since UAH applies the TMT warm target factor to TLT, this represents a bias in TLT as well.

## Appendix C

### MSU/AMSU DIURNAL ADJUSTMENT

As explained in Section 1.1, there have been a number of studies investigating slow evolving discrepancies in the tropics throughout the 1990s and many have suggested that these MSU/AMSU discrepancies are related to the RSS diurnal cycle correction. During this time, the east-west drift of the satellite leads to spurious cooling of the satellite, so MSU/AMSU teams add a correction to compensate for this effect. It has been suggested that RSS overcorrects this problem, which leads to spurious warming relative to UAH (Christy et al., 2007; Randall and Herman, 2008; Christy and Norris, 2006; Christy et al., 2010, 2011; Bengtsson and Hodges, 2011). Nearly all of these studies have utilized radiosondes, which often contain undocumented changes or incomplete metadata records, and even when changes are known they may not be fully removed (Randel and Wu, 2006). Therefore a better reference may be needed to help determine the source of this discrepancy (Mears and Wentz, 2012).

The diurnal correction is very important, especially over land (Mears et al., 2003). To demonstrate the importance of the diurnal drift correction, we show the drift of the MSU/AMSU satellites through the diurnal cycle in Figure C.1. Using the drift of the satellite and the diurnal cycle of the land and ocean, the appropriate correction can be applied to account for the satellite drift. In Figure C.2, we show the annual average diurnal cycle correction for the tropics used by RSS. This diurnal correction is based on five years of model output from the National Center for Atmospheric Research Community Climate System Model version 3 (Mears et al., 2003; Dai and Trenberth, 2004; Mears and Wentz, 2009b).

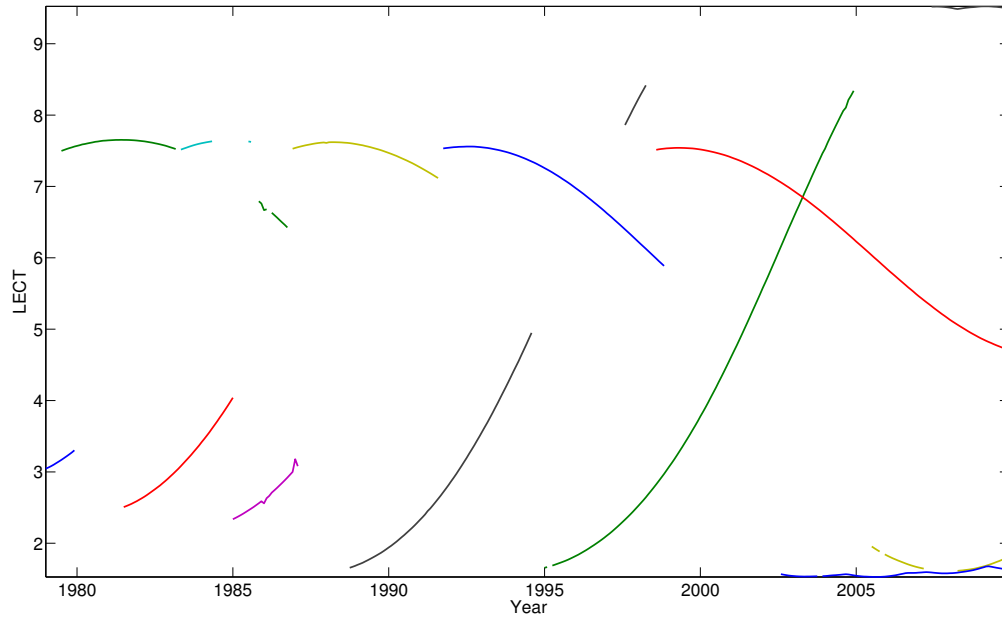


Figure C.1: Local equatorial crossing time (LECT) for the satellites utilized in the MSU/AMSU datasets. Note that some of the satellites, such as the satellite from 1995 - 2005, can drift by more than six hours. Since the satellites are quasi-sun-synchronous, the satellites pass the equator 12 hours apart on the ascending and descending node. So an LECT of “4” means the satellite crosses the equator at 4 AM and 4 PM local time.

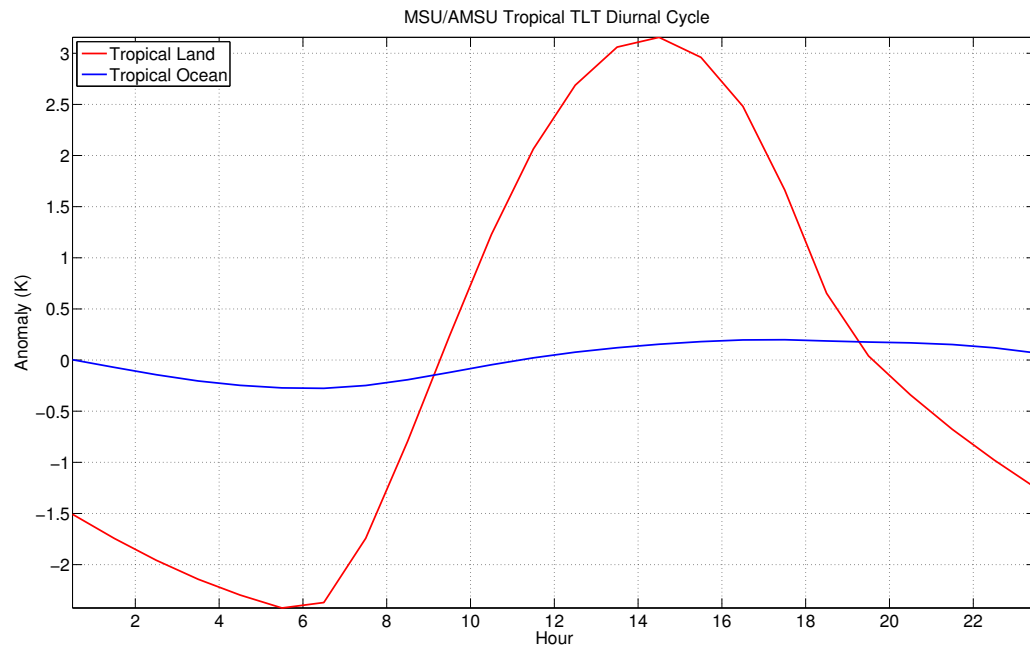


Figure C.2: Tropical diurnal correction used by RSS for land and ocean based on the National Center for Atmospheric Research Community Climate System Model version 3. NOAA utilizes the same correction, but scales it to reduce error residuals for overlapping satellites.

In Figure C.3, we approximate the tropical diurnal corrections for the various satellites utilized in the RSS TLT merging procedure. As suggested by various studies, the tropical diurnal cycle may explain some of the differences between RSS and UAH as seen in Figure C.4. We do not draw conclusions about the integrity of the diurnal cycle corrections utilized by UAH or RSS, but note that this correction may be important in reconciling tropical trends. Importantly, it would be ideal to use different methods to diagnose possible errors in the diurnal cycle calibration because radiosondes contain undocumented and potentially uncorrected time varying biases.

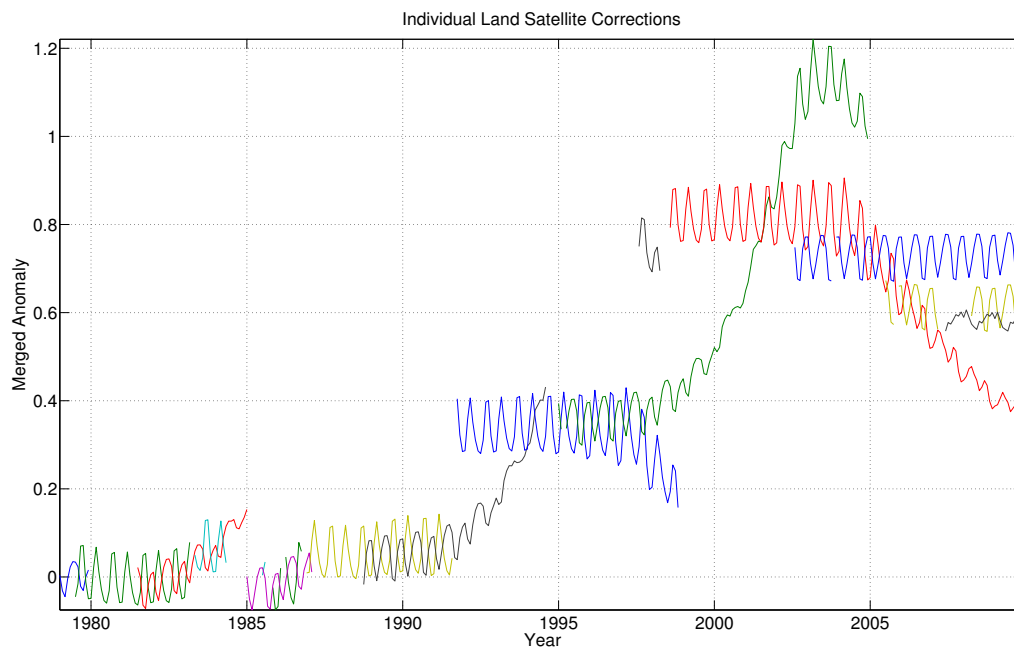


Figure C.3: Estimated tropical ( $30^{\circ}\text{S} - 30^{\circ}\text{N}$ ) diurnal corrections for each satellite. We've offset the corrections such that the mean difference between pairs of satellites approaches zero.

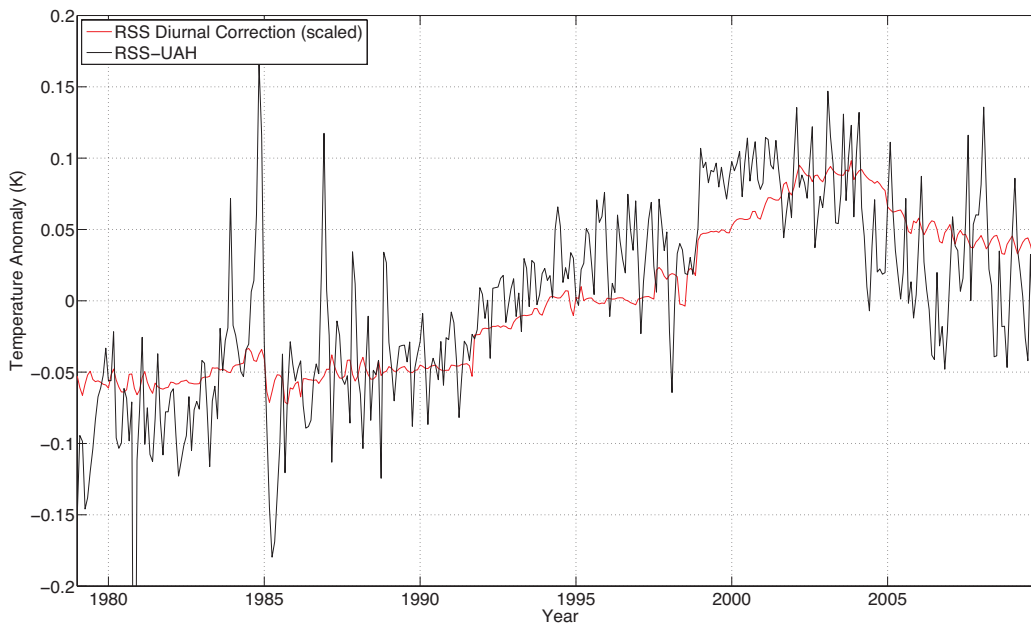


Figure C.4: RSS - UAH TLT difference series in the tropics ( $30^{\circ}\text{S} - 30^{\circ}\text{N}$ ) along with the scaled mean of the diurnal corrections estimated from Figure C.3.

There may be two directions for validating the diurnal correction for the different MSU/AMSU groups. We see from Figure C.5 that the diurnal correction is very sensitive to the phase of the applied diurnal correction. We also know from Dai and Trenberth (2004) that the phase is not necessarily correct in global circulation models. It may be possible to use geostationary satellites to constrain the phase of the clear sky microwave diurnal cycle. Similarly, over periods of large differences between MSU/AMSU groups and large diurnal corrections (such as near 2005 when NOAA-15 has a large diurnal drift) it may be possible to use radio occultation data (Anthes et al., 2008; Ho et al., 2009) to evaluate spurious drifts in the various MSU/AMSU time series (this data is available starting in the early 2000s).

Even though these suggested analyses could shed light on differences between MSU/AMSU group's diurnal cycles, differences in the treatment of the instrument body effect could be a complicating factor because the instrument body effect is related to the

drift of the satellite through the diurnal cycle. Regardless, scrutiny of these records with an array of complementary datasets can help shed light on the accuracy of our long-term temperature trends in the troposphere.

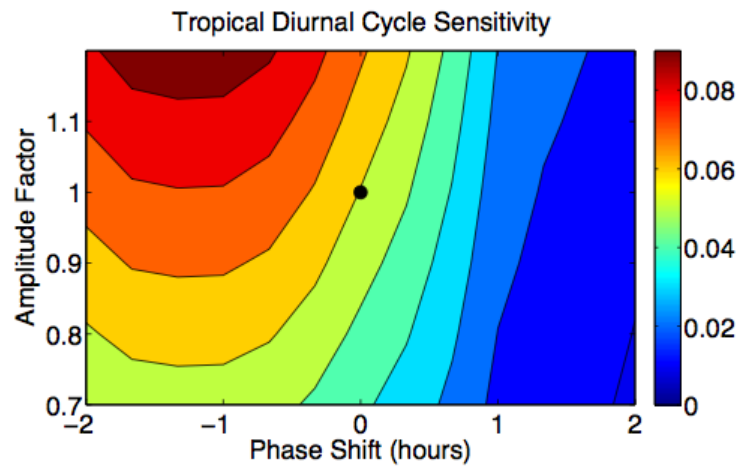


Figure C.5: Estimated sensitivity of the tropical diurnal cycle corrections to the phase and amplitude of the applied diurnal cycle.