Ground Truth Verification of RSS Lower Troposphere Temperature Anomalies Using HadAT2 Radiosonde Data

Randy Miller

GEOG 417: Satellite Climatology

4/29/13

Abstract

Creating reliable upper-air climate records is a challenge for scientists faced with the issue of understanding the climate and climate variability. Radiosonde networks provide a dense but regionally dependent climatology, thus satellite measurements are an important tool for filling in data voids. However, satellite data need "ground" truth verification before they can be used reliably. Hadley Centre's radiosonde temperature anomalies (HadAT2) are used as verification for the Remote Sensing Systems' (RSS) Microwave Sounding Units (MSU)/Advance Microwave Sounding Units (AMSU) temperature anomaly dataset. Several analyses are used, including a comparison of the data series and their trends, the calculation of several statistical variables and the Brier Skill Scores (BSS), and a comparison of the difference between the normalized data series to see if there are atmospheric processes causing noise in the data. It is found that with the exception of the southern hemisphere extratropics the RSS data are in high agreement with the HadAT2 data. There is evidence of underlying atmospheric processes that affect these results, but they are not yet understood. The MSU/AMSUs are important and reliable tools for understanding climate, and satellites should continue to be used and improved upon in the future.

Table of Contents

Introduction 1

Data 3

RSS MSU/AMSU Temperature Anomalies 4

HadAT2 Radiosonde Temperature Anomalies 6

Methods 7

Results 10

Discussion 12

Summary 14

Conclusion 15

References 17

List of Figures

Figure 1 Weightings Used in Scans to Construct the MSU/AMSU Products 5

Figure 2 Weighting Functions for the MSU/AMSU Atmospheric Levels 8

Figure 3 Temperature Anomalies and Trends 10

Figure 4 Temperature Anomaly Difference between HadAT2 data and RSS Data 12

List of Tables

Table 1 Statistical Analysis Results 11

Introduction

In order to understand Earth's climate and climate variability, there needs to be reliable climatological records to analyze. One method of building a climatological record is by taking in situ measurements of the upper atmospheric temperatures using radiosondes. Radiosondes are instruments that are attached to weather balloons and transmit atmospheric measurements, such as temperature and humidity, back to a fixed receiver. This allows scientists to get atmospheric profiles that would be impractical to obtain through other means, such as measurements by aircraft. Thanks to radiosondes, there are many quality upper atmosphere datasets available for scientists to use in their research.

There are limitations to the radiosondes' ability to provide an accurate climate record. Because they are attached to weather balloons, they must be manually launched in order to get upper atmosphere measurements. This means the overwhelming majority of radiosonde measurements are over land, and only a few are taken over the ocean. However, the Earth's surface is primarily (~70%) ocean, leaving most of it unsampled by the radiosonde networks. Furthermore, a robust climatology requires frequent sampling over the course of an hour to have the most accurate results. The nature of radiosonde measurements – the manpower needed, the cost of equipment, the time involved in preparing, launching, and recovering the radiosondes – makes it impractical to take more than two measurements a day. Thus, scientists do their best to make use of what they have.

The Met Office Hadley Centre for Climate Change is dedicated to understanding climatological processes, and to providing quality research products. One of their products is the globally gridded radiosonde temperature anomalies (HadAT2). The HadAT2 project analyses several radiosonde sources attempting to minimize bias caused by the uneven distribution of radiosonde launch sites (Thorne et al. 2005). While the resulting product is very good quality, it still suffers from limitations inherent to radiosonde measurements. However, with the advent of satellites and instruments for measuring microwave radiation from space, it is possible to significantly expand coverage both spatially and temporally.

The National Oceanic and Atmospheric Administration (NOAA) launched satellites with Microwave Sounding Units (MSUs) since 1978, and Advanced Microwave Sounding Units (AMSUs) since 1988 (Enloe 2011). These instruments are capable of measuring radiances of broad vertical layers of the atmosphere, depending on the frequency of microwave radiation used. The measured radiances can then be mathematically converted to atmospheric temperatures. Because the measurements are automated and the satellites revolve around the planet on a polar orbit, they provide a much more homogeneous distribution of measurements. However, satellite measurements have limitations as well.

One problem with using satellites for temperature records is that satellites suffer from orbital decay over time. This leads to spurious results in the temperature record (Wentz and Schabel 1998). Furthermore, measurements are taken by different satellites at different times, and the amount and types of satellites have changed over time. This inevitably leads to biases and noise in the data. Finally, the satellite MSU record only goes back to 1979. With such a short record, it is difficult to draw accurate conclusions about the climate trends. Fortunately, the errors involved in satellite measurement can be reduced with ground truth verification.

Ground truth verification means to verify satellite measurements by comparing them with "ground" in situ measurements. Only through verification can the satellite record be corrected to provide a reliable climate history. Some research has already been done regarding NOAA's

MSU/AMSU satellite records by using Hadley Centre's radiosonde datasets for ground truth verification (Mears and Wentz 2009). However, since the publishing of that research the Hadley Centre has updated their product. Thus, it is useful to reanalyze the MSU/AMSU satellite temperature anomaly record with the updated HadAT2 temperature anomalies.

In this paper, I will determine whether MSU/AMSU satellite temperature anomalies are a reliable means for constructing a climatological record. To accomplish this, I will compare the temperature anomaly trends inferred from Remote Sensing Systems' (RSS) MSU/AMSU measurements with HadAT2's temperature anomalies. Furthermore, I will calculate and analyze the correlation coefficients, the Brier Skill Score (BSS), and the mean anomalies between the datasets. Finally, I will normalize the data and subtract them to potentially discover any patterns that might be causing inaccuracies. The goal is to demonstrate that MSU/AMSU satellites are a reliable means for constructing climatological records.

Data

Two datasets are needed for ground truth verification: the satellite derived dataset in question, and another dataset for comparison. The satellite data are temperature anomalies found on the RSS website[1]. They are designated as Temperature Lower Troposphere (TLT), and are longitudinally average by month over several latitudinal bends. The data used for verification are HadAT2 temperature anomalies from radiosonde networks[2]. There are several versions of the

[1] http://www.ssmi.com/msu/msu_introduction.html

[2] http://www.metoffice.gov.uk/hadobs/hadat/index.html

data, but the one used for this analysis will be the monthly and longitudinally averaged anomalies. The following will discuss both datasets in detail.

RSS MSU/AMSU Temperature Anomalies

The MSU/AMSU instruments measure microwave radiance at several different frequencies or "channels," ranging from 51.3 to 57.95 GHz ("Description of MSU and AMSU Data Products" 2012). Oxygen molecules absorb and emit strongly at these frequencies, thus it is effective for deriving atmospheric temperatures. The temperatures are representative of thick layers of the atmosphere. TLT uses channels 2 and 5, and it represents the bottom-most layer of the troposphere. It is important to note that the MSU/AMSUs do not directly measure the radiance of the lower troposphere. Instead, weighted differences between measurements made at different angles are used to extrapolate lower troposphere measurements (Figure 1). Unfortunately, this extrapolation is a source of uncertainty and noise in the data. However, the advantage to using this method is the lower troposphere is not influenced by the stratospheric cooling, which means the temperature anomalies are more representative of temperature trends observed near the surface.

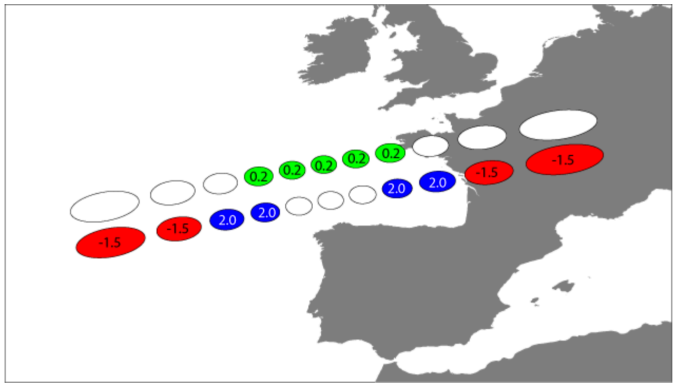

Figure 1 Weightings Used in Scans to Construct the MSU/AMSU Products. The top is the weighting values used for near-nadir measurements for the layers above the lower troposphere. The bottom is the weighting values for scans for the lower troposphere temperatures. The blue are positive weights for close to nadir measurements, while the red are negative weights for the near limb view. ("Description of MSU and AMSU Data Products" 2012)

The MSU/AMSU instruments are located on NOAA polar-orbiting satellites. The MSUs began operation in 1978, and the last one ceased operating in 2005. The AMSUs began operation in 1988, and continue to this day. Each satellite measures most of the surface area of the earth every few days. However, because the microwave wavelengths are low-energy radiation, they suffer from low resolution. The resulting grid of brightness temperatures has a 2.5° resolution. These instruments are also meant for weather forecasting, as apposed to climate records. However, with careful intercalibration between instruments, climate quality datasets can be extracted.

The RSS's website has many datasets available for use, but the dataset that is of interest for this paper is the zonally averaged temperature anomalies. Of the six latitudinal bands available in the dataset, the follow four will be used: -70° to 82.5° (global), -20° to 20° (tropics), 20° to 82.5° (northern hemisphere extratropics (NHE)), and -70° to -20° (southern hemisphere extratropics (SHE)). The data are averaged monthly beginning in January 1979, and ending in December 2012. The anomalies are based off the 1979-1988 period. Finally, because radiosonde data are limited primary to land measurements, only the land-based temperature anomalies will be used.

HadAT2 Radiosonde Temperature Anomalies

Unlike MSU/AMSUs, radiosondes measure upper-atmospheric temperatures directly, or in situ. They are attached to weather balloons and released typically twice a day. The measurements they take represent point measurements, as opposed to the pixel-averaged measurements of MSU/AMSUs. The radiosondes can be programmed to report temperatures at certain atmospheric levels as they rise through the atmosphere. The disadvantage to radiosonde measurements is they are limited primarily to land-based measurements. Furthermore, the distribution of stations using radiosondes is denser in the northern hemisphere, primarily in North America and Eurasia. This lack of homogeneity can lead to temperature biases and noise in the data. Finally, the methods and instruments used in the radiosonde measurements have varied over the record, making it difficult to construct a consistent dataset. HadAT2 data reduce this problem by rejecting much of the data that do not follow a consistent pattern.

The Hadley Centre creates several radiosonde temperature anomaly products. The dataset that will be used for the ground truth verification is the zonally average monthly temperature anomalies. The data are separated into 5° latitudinal bands that cover most of the surface area from -87.5° to 87.5°. Furthermore, there are measurements at each of the following pressure levels: 850, 700, 500, 300, 200, 150, 100, 50, and 30 hPa. The data begin in 1958 and continue to the present, but only January 1979 through December 2012 will be used in the analysis. The HadAT2 datasets are constructed from several available radiosonde networks by analyses and composites of neighboring radiosondes. Finally, the anomalies are based off the 1966 to 1995 climatology.

Methods

Before beginning the analyses, the HadAT2 dataset is processed and made compatible with the RSS dataset. First, the anomalies are averaged over the latitudinal bands of the RSS data. This is accomplished by taking the cosine of the latitudes within each band, and using the result as the weight for calculating the latitudinal average. This is the same procedure outlined by Carl A Mears and Frank J. Wentz (2009) in their analysis of previous HadAT datasets. An assumption made with this method is that the zonal averages are good representations of the anomalies at that latitude, which may not be the case. For example, -42.5° will have the same weight as 42.5°; however, there are far fewer radiosonde stations at -42.5°. Next, the column average anomaly is calculated using the pressure level weightings and method available though the RSS website (Figure 2). The weights are given for many more pressure levels than the nine available in the radiosonde data, thus the average weight is calculated for each pressure level using the following equation:

$$Level\ weight = 0.5\ x\ (weight_i + \ weight_{i-1})\ x\ (height_i - \ height_{i-1}), \quad (1)$$

where i is the top of the level, i-1 is the bottom of the level, weight is the weight of the level based on the weighting function, and height is the height of the level in kilometers. After the weights for each level are calculated, they are used to find the column-averaged anomaly. Because the HadAT2 and RSS anomalies are based off two different climatological periods, the difference between these periods must be removed. The average temperatures for these periods are not provided, so the average anomalies are calculated for the two periods using the HadAT2 dataset, and the difference is subtracted from the HadAT2 dataset. Both datasets are then averaged annually to smooth out the data.

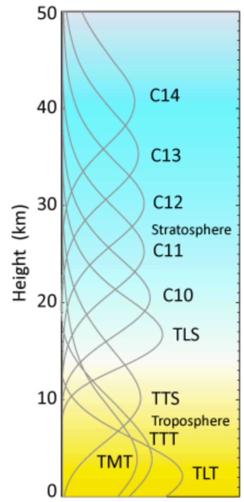

Figure 2 Weighting Functions for the MSU/AMSU Atmospheric Levels. TLT is the function for the lowest region of the troposphere, which is the region used for this paper. ("Description of MSU and AMSU Data Products" 2012)

Using a MATLAB function, the linear trend lines are then calculated for each dataset, and each latitudinal band. Four plots are created for comparison: one plot for each latitudinal band. Within each plot, the RSS and HadAT2 temperature anomalies are plotted as a function of time, as well as their respective trend lines. This allows for observational analysis of the agreement between the datasets.

Statistical analyses include calculating the correlation coefficient, the difference between the mean anomalies (the bias), and the Brier Skill Score (BSS) of each latitude band. The correlation coefficients give a sense of how correlated the two datasets are. Because they are temperature anomaly records over the same upper-atmospheric level, they should have a high correlation coefficient. The mean anomaly difference gives a sense to the bias of each dataset. Ideally, there will not be a bias; however, because the average temperatures were not provided for each anomaly dataset, there may be a bias. The BSS is a way to evaluate a forecast against the observations. It is a useful analysis tool for weather forecasts, but it can also give an objective measure of the "skill" of the satellite-derived climatology as well. The RSS data can be treated as the forecast and compared to the HadAT2 data, or the observations. The equation for calculating the BSS is

$$BSS = r^2 - \left(r - \frac{s_f}{s_o}\right)^2 - \left(\frac{\bar{f} - \bar{o}}{s_o}\right)^2, \quad (2)$$

where r is the correlation coefficient, s_f and s_o are the standard deviations of the RSS and HadAT2 temperature anomalies, respectively, and $\bar{f}$ and $\bar{o}$ are the mean anomalies for the RSS and HadAT2 data, respectively. The ideal score is 1, and the farther below 1 the score is (including negative numbers), the worse the score.

The final analysis normalizes the datasets by removing their trends. Then, the RSS dataset is subtracted from the HadAT2 dataset. This is repeated for all four latitudinal bands, and the results are averaged between the datasets for each year. Finally, the five time series are plotted together. If there are any background patterns causing inaccuracies between the datasets, the plotted differences may provide insight as to what they are.

Results

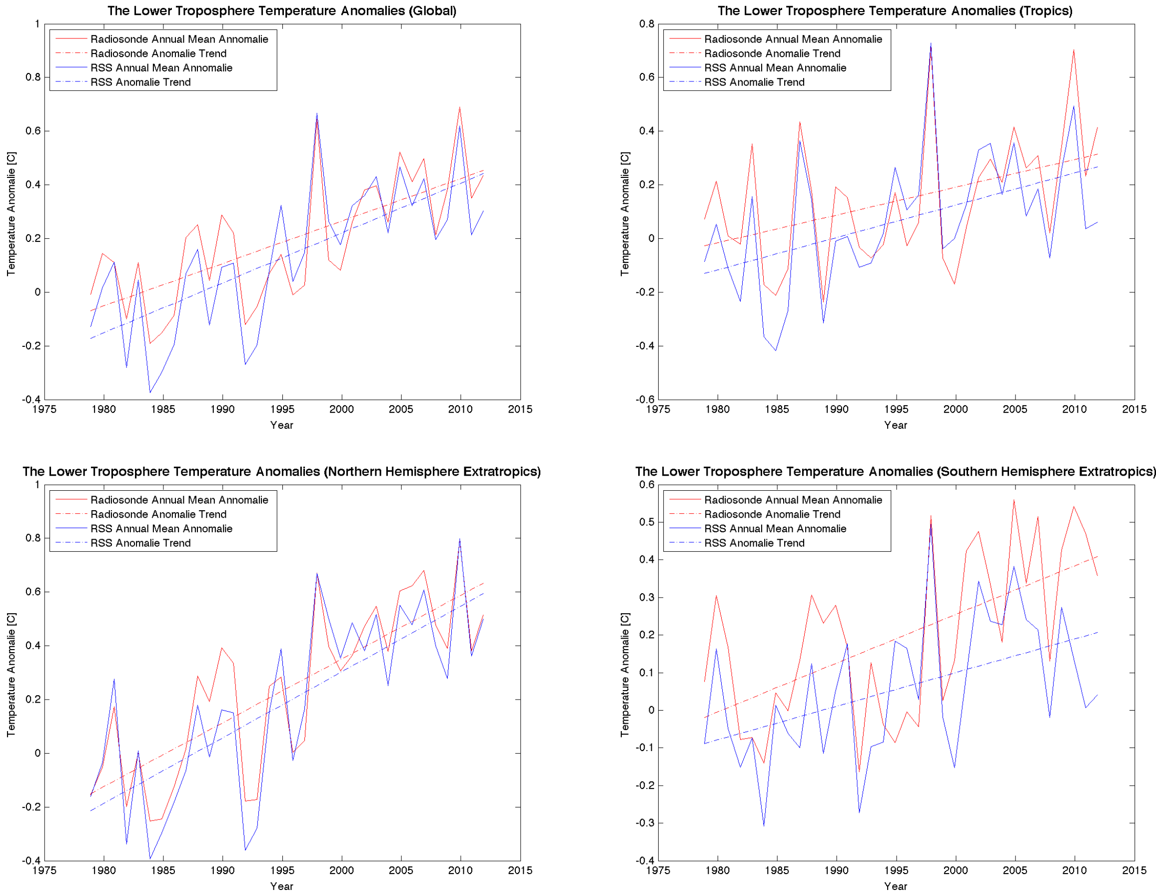

Figure 3 Temperature Anomalies and Trends. Top-left: global anomalies. Top-right: tropical anomalies. Bottom-left: northern hemisphere extratropics. Bottom-right: southern hemisphere extratropics. Red lines represent HadAT2 data while blue lines represent RSS data. Trends and means appear to agree in all regions except the southern hemisphere extratropics.

Figure 3 shows the plotted temperature anomalies and trends for the four latitude bands. The difference between the two climate periods that was subtracted out of the HadAT2 data is 0.12°C. With the exception of the SHE, there is a lot of agreement between the datasets. The NHE has the best agreement of all the latitude ranges; the difference in trends is indistinguishable when plotted. The tropical and global anomalies look similar in overall accuracy; however, the global data look better due to the slightly smaller bias. Furthermore, the signatures of large global events such as the Mt. Pinatubo eruption in 1992 and the strong El Niño of 1998 are also noticeable in the time series. The HadAT2 dataset shows a warm bias in all four plots. The four datasets agree with their counterpart for the 1998 El Niño event. All of the latitude bands also showed a warming trend over the period of analysis.

	Latitude Band	Trend (°C per year)	Standard Deviation (°C)	Mean Anomaly (°C)	Correlation Coefficient	BSS
RSS	Global	0.019	0.259	0.135	0.924	0.741
	Tropics	0.012	0.246	0.069	0.868	0.624
	N.H. Extratropics	0.024	0.324	0.190	0.956	0.871
	S.H. Extratropics	0.009	0.184	0.059	0.708	0.089
HadAT2	Global	0.016	0.226	0.192		
	Tropics	0.010	0.236	0.143		
	N.H. Extratropics	0.024	0.301	0.241		
	S.H. Extratropics	0.013	0.217	0.195		

Table 1 Statistical Analysis Results. The statistical variables are listed across the top, and the datasets and latitude bands are labeled on the left. The correlation coefficients and BSSs are not listed for HadAT2 data because they are calculated using both datasets, thus the values listed in the RSS columns represent both datasets.

Table 1 summarizes the results from the statistical analyses. The NHE have the largest temperature anomaly trend of 0.024°C per year, and it is the only latitude band where both datasets have the same trend. The SHE have the largest disagreement in temperature anomaly trends between the datasets, with a difference of 0.004°C per year. The NHE (SHE) have the highest (lowest) standard deviation of ~0.31°C (~0.20°C). The SHE HadAT2 data have the strongest temperature bias, with a value of 0.14°C. The NHE HadAT2 data have the lowest bias (0.05°C), and the global bias is only slightly higher (0.06°C). The NHE have to highest correlation coefficient (0.96) and the SHE have the lowest (0.71). Finally, the NHE have the highest BSS (0.87) as opposed to the SHE with the lowest BSS (0.09).

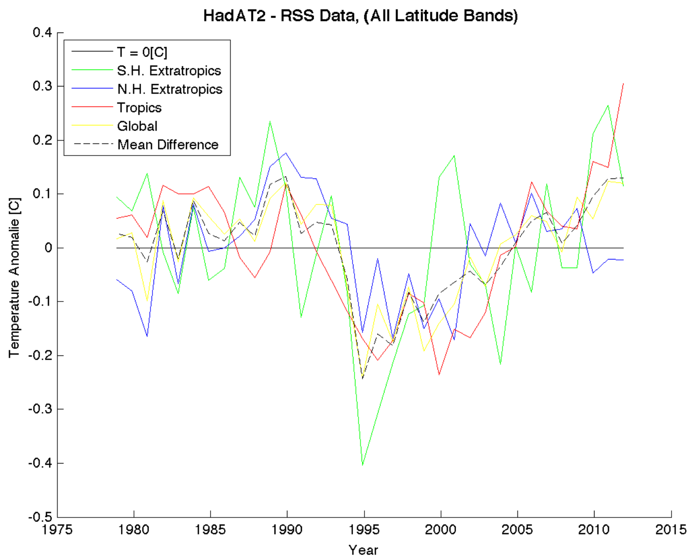

Figure 4 Temperature Anomaly Difference between HadAT2 data and RSS Data. The data were first normalized by removing the trend, and then the RSS data were subtracted from the HadAT2 data. A distinctive almost sinusoidal pattern is noticeable over the period; however, asymmetry prevents it from being a true sinusoidal pattern.

Figure 4 shows the result of the normalized anomaly differencing. If there are no underlying processes affecting the datasets then the different plots would more or less fluctuate about the T = 0°C line without a noticeable pattern; however, this is not the case. A very distinctive pattern starts out slightly positive, increases a little, then quickly dips after 1992, then increases for the remainder of the period. This pattern is noticeable across all of the latitude bands, with the global anomalies following it most closely. Furthermore, it does not seem to correlate with any single synoptic-dynamic climatological processes, or the solar cycle. It is difficult to draw an immediate conclusion from this result.

Discussion

There is overwhelming agreement between the two datasets, with a few exceptions. One issue is the temperature anomaly biases, with the RSS data consistently having a cool bias. As previously mentioned the average temperatures used to calculate the temperature anomalies are not listed on the dataset provider's websites. To substitute for the missing averages, the HadAT2 dataset was used to calculate the mean anomaly for both periods, and then the difference was subtracted from the data. While this reduces the overall bias, it is not an accurate solution. This trick could be the reason there is a bias, or it may be hiding a larger bias in the data. No conclusion can be drawn about the bias with certainty.

The next issue is the consistently poor evaluations of the SHE. Most radiosondes are launched from the land, and only a small portion of the surface in the SHE is land. Furthermore, there are much less stations providing radiosonde measurements in the SHE. Because of these issues, the temperature anomalies probably do not provide an accurate representation of the zonal averages in the SHE, and the data are more sensitive to noise. It will be useful if in future studies the data in the SHE are considered regionally, although that will not solve all of the inaccuracies. Considering the strong agreement of the global and NHE bands, the inaccuracies in the SHE are likely due to gaps in HadAT2 data, not the RSS measurements.

The NHE and the Global latitude bands performed exceptionally well over all of the analyses. Considering the dense coverage by radiosonde networks in the northern hemisphere, this serves to confirm that MSU/AMSU can be used to construct climatological records. The RSS datasets will probably improve as more accurate sounding units are put into orbit, and as the dataset grows larger and more consistent. Carl A. Mears' and Frank J. Wentz's (2009) results are similar to this study. Their derived trend lines are comparable to the values in Table 1; however, their research has a noticeably smaller bias. They also use an area averaging along with the zonal-band averaging, with better agreement between radiosonde and MSU/AMSU measurements.

The final discussion point is the peculiar pattern observed in the plot of temperature anomaly differences. The plot can be divided into three sections: the period from 1979 to 1992 when the mean's trend was slightly positive and HadAT2 anomalies were higher than RSS anomalies, the period from 1992 to 1995 when the mean's trend was sharply negative, and the 1995 to 2012 period where the mean's trend was steady and somewhere in-between the absolute values of the other two slopes. When compared to some of the dominant atmospheric and ocean oscillations (the El Niño Southern Oscillation, for example) there does not seem to be a noticeable connection. Furthermore, the timeframe of the pattern is longer than any single atmospheric process. It is possible that the pattern is the result of several interacting processes. The rapid decline after 1992 could be related to the Mt. Pinatubo eruption; however, the atmosphere usually recovers more quickly from volcanic eruptions than what is observed. Future research should further the investigation into how atmospheric processes may interact with and influence satellite measurements of TLT.

Summary

Maintaining accurate climate records is essential to our understanding of climate and climate variability. Radiosonde networks are great tools for monitoring upper-air temperature anomalies, among other variables, but they are severely limited in what they can cover over spatial and temporal scales. Satellites help fill in the voids, but need verification to ensure their results are reliable. Thus, radiosonde networks are often used as ground truth for verifying MSU/AMSU instruments. Two popular sources for radiosonde and MSU/AMSU data are the Hadley Centre and Remote Sensing Systems, respectively. The RSS datasets take advantage of NOAA's extensive MSU/AMSU coverage of lower troposphere temperatures since 1979. The

Hadley Centre combines several radiosonde networks data together to create a reliable temperature anomaly record of the upper atmosphere. Thus, it makes sense to use the Hadley Centre's datasets as verification for the RSS datasets.

Various statistical analyses can be used for verification, such as comparing biases, trends, or skill scores. It is useful to plot each time series with their trends as well so that patterns in the data can be observationally analyzed. Another analysis, a normalized differencing of the datasets, can reveal the signature of atmospheric patterns that are difficult to notice otherwise. Overall, there is large agreement between the datasets with a notable exception in the SHE. The uncertainties associated with the SHE should not discourage the use of satellite-derived datasets in the southern hemisphere. While there is possibly a scientific reason for the errors, it is more likely due to the lack of coverage in the radiosonde network in the global south. The normalized differencing gives insight to underlying processes that affect the agreement between measurements, but more research needs to be conducted in order to determine what those processes are. The satellite-derived climatology is a reliable record, and could be used for research where the radiosonde networks are inadequate, such as over the oceans.

Conclusion

Humankind is in a precarious place where for the first time in history they are impacting climate on a global scale. Anthropogenic climate change has large implications on many aspects of life, such as human health and food production. It is especially important that we continue to increase our understanding of climate and climate variability in order to adapt to or even mitigate the impacts. The only way to increase our knowledge of the climate system is through reliable measurement of variables. Satellites became an invaluable tool for atmospheric measurements as they provide near-global coverage at much higher sampling rates than radiosondes. Therefore, it is imperative that satellite technology continues improving. The next generation of microwave sounders, the Advanced Technology Microwave Sounders (ATMS), is already beginning as ATMS satellites began preliminary operation in October 2011. I hope that this will improve upon the already reliable AMSU, because we cannot ignore the need for accurate information to prepare for the climate challenges that lie ahead.

References

"Description of MSU and AMSU Data Products." 2012. *Remote Sensing Systems*. August 17. http://www.ssmi.com/msu/msu_data_description.html.

Enloe, Jesse. 2011. "Upper Atmospheric Temperatures." *National Climatic Data Center*. December 21. http://www.ncdc.noaa.gov/temp-and-precip/msu/.

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

Thorne, Peter W, David E Parker, Simon FB Tett, Phil D Jones, Mark McCarthy, Holly Coleman, and Philip Brohan. 2005. "Revisiting Radiosonde Upper Air Temperatures from 1958 to 2002." *Journal of Geophysical Research: Atmospheres (1984–2012)* 110 (D18). http://onlinelibrary.wiley.com/doi/10.1029/2004JD005753/full.

Wentz, Frank J., and Matthias Schabel. 1998. "Effects of Orbital Decay on Satellite-derived Lower-tropospheric Temperature Trends." *Nature* 394 (6694) (August 13): 661–664. doi:10.1038/29267.