Chapter 3

What do observations indicate about the changes of temperature in the atmosphere and at the surface since the advent of measuring temperatures vertically?

Convening Lead Author: John R. Lanzante

Lead Authors: Thomas C. Peterson, Frank J. Wentz and Konstantin Y.Vinnikov

Contributing Authors: Dian J. Seidel, Carl A. Mears, John R. Christy, Chris Forest, Russell S. Vose, Peter W. Thorne and Norman C. Grody

Key Findings

Observed Changes - Surface

Globally, as well as in the tropics, the temperature of the air near the Earth's surface has increased since 1958, with a greater rate of increase since 1979. All three surface temperature datasets are consistent in these conclusions.

- Globally, temperature increased at a rate of about 0.12ºC per decade since 1958, and about 0.16ºC per decade since 1979.
- In the tropics, temperature increased at a rate of about 0.11ºC per decade since 1958, and about 0.13ºC per decade since 1979.
- Most, if not all of the surface temperature increase since 1958 occurs starting around the mid-1970s, a time coincident with a previously identified abrupt climate regime shift. However, there does not appear to be a strong jump up in temperature at this time, rather the major part of the rise seems to occur in a more gradual fashion.

Observed Changes - Troposphere

Globally, as well as in the tropics, both balloon-based datasets dating back to 1958 agree that the tropospheric temperature has increased slightly more than that of the surface. Since 1979, due to the considerable disagreement among tropospheric datasets, it is not clear whether the temperature of the troposphere has increased more or less than that of the surface, both globally and in the tropics.

- Globally, temperature increased at a rate of about 0.14ºC per decade since 1958 according to the two balloon-based datasets. Since 1979, estimates of the increase from the two balloon and three satellite datasets range from about 0.10 to 0.20ºC per decade.
- In the tropics, temperature increased at a rate of about 0.13ºC per decade since 1958 according to the two balloon-based datasets. However, since 1979, estimates of the increase from the two balloon and three satellite datasets range from about 0.02 to 0.19ºC per decade.
- For the balloon-based estimates since 1958, the major part of the temperature increase appears in the form of an abrupt rise in the mid-1970s, apparently in association with a climate regime shift that occurred at this time.

Observed Changes - Lower Stratosphere

Globally, the temperature of the lower stratosphere has decreased both since 1958 and since 1979. The two balloon-based datasets yield reasonably consistent estimates of the rates of cooling for both time periods. However, since 1979 the two balloon datasets estimate a considerably greater rate of cooling than the two satellite datasets, which themselves disagree.

- Globally, the rate of cooling since 1958 is about 0.37ºC per decade based on the two balloon datasets. Since 1979, estimates of this decrease are about 0.65ºC per decade for the two balloon datasets, and from about 0.33 to 0.45º C per decade for the two satellite datasets.
- The bulk of the stratospheric temperature decrease occurred from about the late 1970s to the middle 1990s. It is unclear whether the decrease was gradual or occurred in abrupt steps in the first few years after each major volcanic eruption.

Chapter 3: Recommendations

- *Although considerable progress has been made in explaining the causes of discrepancies between upper-air datasets, both satellite and balloon-based, continuing steps should be taken to thoroughly assess and improve methods used to remove time-varying biases that are responsible for these discrepancies.*
- *New observations should be made available in order to provide more redundancy in climate monitoring. Activities should include both the introduction of new observational platforms as well as the necessary processing of data from currently under-utilized platforms. For example IR and GPS satellite observations have not been used to any great extent, the former owing to complications when clouds are present and the latter owing to a short period of record. Additionally, the introduction of a network of climate quality, reference stations, that include reference radiosondes, would place future climate monitoring on a firmer basis.*

1. Background

In this chapter we describe changes in temperature at the surface and in the atmosphere based on four basic types of products derived from observations: surface, radiosonde, satellite and reanalysis. However, we limit our discussion of reanalysis products given their more problematic nature for use in trend analysis (see Chapter 2); only a few trend values are presented for illustrative purposes.

Each of these four generic types of measurements consists of multiple datasets prepared by different teams of data specialists. The datasets are distinguished from one another by differences in the details of their construction. Each type of measurement system as well as each particular dataset has its own unique strengths and weaknesses. Because it is difficult to declare a particular dataset as being "the best," it is prudent to examine results derived from more than one "credible" dataset of each type. Also, comparing results from more than one dataset provides a better idea of the uncertainties or at least the range of results. In the interest of clarity and conciseness, we have chosen to display and perform calculations for a representative subset of all available datasets. We consider these to be the "state of the art" datasets of their type, based on our collective expert judgment.

In selecting datasets for use in this report, we limit ourselves to those products that are being actively updated and for which temporal homogeneity is an explicit goal in the construction, as these are important considerations for their use in climate change assessment. By way of a literature review, we discuss additional datasets not used in this report. Since some datasets are derivatives of earlier ones, we mention this where appropriate. One should not misconstrue the exclusion of a dataset from this report as an invalidation of that product. Indeed, some of the excluded datasets have proved to be quite valuable in the past and will continue to be so into the future.

Most of the analyses that we have performed involve data that were averaged over a large region, such as the entire globe or the tropics. The spatial averaging process is complicated by the fact that the locations (gridpoints or stations) at which data values are available can vary fundamentally by data type (see Chapter 2 for details) and, even for a given type, between data production teams. In an effort towards more consistency, the spatial averages we use represent the weighted average of zonal averages[1] (i.e., averages around an entire latitude line or zone), where the weights are the cosine of latitude[2]. This insures that the different latitude zones are given equal treatment across all datasets.

This chapter begins with a discussion of the four different data types, introducing some temperature datasets for each type, and then discussing their time histories averaged over the globe. Later we present more detail, concentrating on the analysis of temperature trends for two eras: (1) the period since the widespread availability of radiosonde observations in 1958, and (2) since the introduction of satellite data in 1979. We compare

[1] The zonal averages, which were supplied to us by each dataset production team, differ among datasets. We allowed each team to use their judgment as how to best produce these from the available gridpoint or station values in each latitude zone

[2] The cosine factor weights lower latitudes more than higher ones, to account for the fact that lines of longitude converge towards the poles. As a result, a zonal band in lower latitudes encompasses more area than a comparably sized band (in terms of latitude/longitude dimensions) in higher latitudes.

overall temperature trends from different measurement systems and then go into more detail on trend variations in the horizontal and vertical.

2. Surface Temperatures

2.1 Land-based temperature data

Over land, temperature data come from fixed weather observing stations with thermometers housed in special instrument shelters. Records of temperature from many thousands of such stations exist. Chapter 2 outlines the difficulties in developing reliable surface temperature datasets. One concern is the variety of changes that may affect temperature measurements at an individual station. For example, the thermometer or instrument shelter might change, the time of day when the thermometers are read might change, or the station might move. These problems are addressed through a variety of procedures (see Peterson et al., 1998 for a review) that are generally quite successful at removing the effects of such changes at individual stations (e.g., Vose et al., 2003 and Peterson, 2005) whether the changes are documented in the metadata or detected via statistical analysis using data from neighboring stations as well (Aguilar et al., 2003). Subtle or widespread impacts that might be expected from urbanization or the growth of trees around observing sites might still contaminate a dataset. These problems are addressed either actively in the data processing stage (e.g., Hansen et al., 2001) or through dataset evaluation to ensure as much as possible that the data are not biased[3]

[3] Changes in regional land use such as deforestation, aforestation, agricultural practices, and other regional changes in land use are not addressed in the development of these datasets. While modeling studies have

(e.g., Jones et al., 1990; Peterson, 2003; Parker, 2004; Peterson and Owen, 2005).

2.2 Marine temperature data

Data over the ocean come from moored buoys, drifting buoys, and volunteer observing ships. Historically, ships have provided most of the data, but in recent years an increasing number of buoys have been used, placed primarily in data-sparse areas away from shipping lanes. In addition, satellite data are often used after 1981. Many of the ships and buoys take both air temperature observations and sea surface temperature (SST) observations. Night marine air temperature (NMAT) observations have been used to avoid the problem that the Sun's heating of the ship's deck can make the thermometer reading greater than the actual air temperature. Where there are dense observations of NMAT and SST, over the long term they track each other very well. However, since marine observations in an area may only be taken a few times per month, SST has the advantage over air temperature in that water temperature changes much more slowly than that of air. Also, there are twice as many SST observations as NMAT from the same platforms as SSTs are taken during both the day and night and SST data are supplemented in data sparse areas by drifting buoys which do not take air temperature measurements. Accordingly, only having a few SST observations in a grid box for a month can still provide an accurate measure of the average temperature of the month.

suggested over decades to centuries these affects can be important on regional space scales (Oleson et al., 2004), we consider these effects to be those of an external forcing to the climate system and are treated as such by many groups in the simulation of climate using the models described in Chapter 5. To the extent that these effects could be large enough to have a measurable influence on global temperature, these changes will be detected by the land-based surface network.

2.3 Global surface temperature data

Currently, there are three main groups creating global analyses of surface temperature (see Table 3.1), differing in the choice of available data that are utilized as well as the manner in which these data are synthesized. Since the network of surface stations changes over time, it is necessary to assess how well the available observations monitor global or regional temperature. There are three ways in which to make such assessments (Jones, 1995). The first is using "frozen grids" where analysis using only those grid boxes with data present in the sparsest years is used to compare to the full dataset results from other years (e.g., Parker et al., 1994). The results generally indicate very small errors on multi-annual timescales (Jones, 1995). The second technique is subsampling a spatially complete field, such as model output, only where in situ observations are available. Again the errors are small (e.g., the standard errors are less than 0.06°C for the observing period 1880 to 1990; Peterson et al., 1998b). The third technique is comparing optimum averaging, which fills in the spatial field using covariance matrices, eigenfunctions or structure functions, with other analyses. Again, very small differences are found (Smith et al., 2005). The fidelity of the surface temperature record is further supported by work such as Peterson et al. (1999) which found that a rural subset of global land stations had almost the same global trend as the full network and Parker (2004) that found no signs of urban warming over the period covered by this report.

Table 3.1: Temperature datasets utilized in this report.

Our Name	*Name given by Producers*	*Producers*	*Web Page*
-- Surface –			
NOAA	ER-GHCN-ICOADS	NOAA's National Climatic Data Center (NCDC)	http://www.ncdc.noaa.gov/oa/climate/monitoring/gcag/gcag.html
GISS	Land+Ocean Temperature	NASA's Goddard Institute for Space Studies (GISS)	http://www.giss.nasa.gov/data/update/gistemp/graphs/
HadCRUT2v	HadCRUT2v	Climatic Research Unit of the University of East Anglia and the Hadley Centre of the UK Met Office.	http://www.cru.uea.ac.uk/cru/data/temperature
-- Radiosonde –			
RATPAC	RATPAC	NOAA's: Air Resources Laboratory (ARL), Geophysical Fluid Dynamics Laboratory (GFDL), and National Climatic Data Center (NCDC)	http://www.ncdc.noaa.gov/
HadAT2	HadAT2	Hadley Centre, UK	http://www.hadobs.org/
-- Satellite –			
Temperature of the Lower Troposphere			
T_{2LT-A}	TLT	University of Alabama in Huntsville (UAH)	http://vortex.nsstc.uah.edu/data/msu/t2lt
T_{2LT-R}	TLT	Remote Sensing System, Inc. (RSS)	http://www.remss.com/msu/msu_data_description.html

Temperature of the Middle Troposphere

$T_{2\text{-}A}$ TMT — University of Alabama in Huntsville (UAH)
http://vortex.nsstc.uah.edu/data/msu/t2

$T_{2\text{-}R}$ TMT — Remote Sensing System, Inc. (RSS)
http://www.remss.com/msu/msu_data_description.html

$T_{2\text{-}M}$ Channel 2 — University of Maryland and NOAA/NESDIS (U.Md.)

Temperature of the Middle Troposphere minus Stratospheric Influences

T^*_G (global) $T_{(850\text{-}300)}$ — University of Washington, Seattle (UW) and
T^*_T (tropics) — NOAA's Air Resources Laboratory (ARL)

Temperature of the Lower Stratosphere

$T_{4\text{-}A}$ TLS — University of Alabama in Huntsville (UAH)
http://vortex.nsstc.uah.edu/data/msu/t4

$T_{4\text{-}R}$ TLS — Remote Sensing System, Inc. (RSS)
http://www.remss.com/msu/msu_data_description.html

-- Reanalysis –

US NCEP50 — National Center for Environmental Prediction, NOAA and the National Center for Atmospheric Research
http://wesley.ncep.noaa.gov/reanalysis.html

European ERA40 — European Center for Medium Range Forecasting
http://www.ecmwf.int/research/era

2.3.1 NOAA

The National Oceanic and Atmospheric Administration (NOAA) National Climatic Data Center (NCDC) integrated land and ocean dataset (see Table 3.1) is derived from in situ data. The SSTs come from the International Comprehensive Ocean-Atmosphere Data Set (ICOADS) SST observations release 2 (Slutz et al., 1985; Woodruff et al., 1998; Diaz et al., 2002). Those that pass quality control tests are averaged into monthly 2^o grid boxes

(Smith and Reynolds, 2003). The land surface air temperature data come from the Global Historical Climatology Network (GHCN) (Peterson and Vose, 1997) and are averaged into 5° grid boxes. A reconstruction approach is used to create complete global coverage by combining together the faster and slower time-varying components of temperature (van den Dool et al., 2000; Smith and Reynolds, 2005).

2.3.2 NASA (GISS)

The NASA Goddard Institute for Space Studies (GISS) produces a global air temperature analysis (see Table 3.1) known as GISTEMP using land surface temperature data primarily from GHCN and the U.S. Historical Climatology Network (USHCN; Easterling, et al., 1996). The NASA team modifies the GHCN/USHCN data by combining at each location the time records of the various sources and adjusting the non-rural stations in such a way that their long-term trends are consistent with those from neighboring rural stations (Hansen et al., 2001). These meteorological station measurements over land are combined with in situ sea surface temperatures and Infrared Radiation (IR) satellite measurements for 1982 to the present (Reynolds and Smith, 1994; Smith et al., 1996) to produce a global temperature index (Hansen et al., 1996).

2.3.3 UK (HadCRUT2v)

The UK global land and ocean dataset (HadCRUT2v, see Table 3.1) is produced as a joint effort by the Climatic Research Unit of the University of East Anglia and the Hadley Centre of the UK Meteorological (Met) Office. The land surface air temperature data are from Jones and Moberg (2003) of the Climatic Research Unit. The global SST fields are produced by the Hadley Centre using a blend of COADS and Met Office data bank in situ observations (Rayner, et al., 2003). The integrated dataset is known as HadCRUT2v (Jones and Moberg, 2003)[4]. The temperature anomalies were calculated on a 5^ox5^o grid box basis. Within each grid box, the temporal variability of the observations has been adjusted to account for the effect of changing the number of stations or SST observations in individual grid-box temperature time series (Jones et al., 1997, 2001). There is no reconstruction of data gaps because of the problems of introducing biased interpolated values.

2.3.4 Synopsis of surface datasets

Since the three chosen datasets utilize many of the same raw observations, there is a degree of interdependence. Nevertheless, there are some differences among them as to which observing sites are utilized. An important advantage of surface data is the fact that at any given time there are thousands of thermometers in use that contribute to a global or other large-scale average. Besides the tendency to cancel random errors, the large number of stations also greatly facilitates temporal homogenization since a given station may have several “near-neighbors” for “buddy-checks.” While there are fundamental differences in the methodology used to create the surface datasets, the differing techniques with the same data produce almost the same results (Wuertz et al., 2005). The

[4] Although global and hemispheric temperature time series created using a technique known as optimal averaging (Folland et al., 2001a; Parker et al., 2004), which provides estimates of uncertainty in the time series, including the effects of data gaps and uncertainties related to bias corrections or uncorrected biases, are available, we have used the data in their more basic form, for consistency with the other datasets.

small differences in deductions about climate change derived from the surface datasets are likely to be due mostly to differences in construction methodology and global averaging procedures.

2.4 Global surface temperature variations and differences between the datasets

Examination of the three global temperature anomaly time series (T_{Sfc}) from 1958 to the present shown in Figure 3.1 reveals that the three time series have a very high level of agreement. They all show some temperature decrease from 1958 to around 1976, followed by a strong increase. That most of the temperature change occurs after the mid 1970s has been previously documented (Karl et al., 2000; Folland et al., 2001b; Seidel and Lanzante, 2004). The variability of the time series is quite similar as are their trends. The signature of the El Niño-Southern Oscillation (ENSO), whose origin is in the tropics, is responsible for many of the prominent short-term (several year) up and down swings of temperature (Trenberth et al., 2002). The strong El Niño of 1997-98 stands out as an especially large warm event within an overall upward trend.

Figure 3.1 - Time series of globally averaged surface temperature (T_S) for NOAA (violet), GISS (black), and HadCRUT2v (green) datasets. All time series are 7-month running averages (used as a smoother) of original monthly data, which were expressed as a departure (ºC) from the 1979-97 average.

3 RADIOSONDE TEMPERATURES

3.1 Balloon-borne temperature data

Since the beginning of the radiosonde era, several thousand sites have been used to launch balloons. However, many of these were in operation for only short periods of time. One approach has been to use a fixed station network consisting of a smaller number of stations having long periods of record. A complimentary approach is to grid the data, using many more stations, allowing stations to join or drop out of the network over the course of time. Since each approach has advantages and disadvantages, we utilize both. A further complication is that changes over time in instruments and recording practices have imparted artificial changes onto the temperature records. Some groups have developed methods that try to remove these artificial effects as much as possible. We employ two radiosonde datasets (see Table 3.1), one station-based and one gridded. Both datasets have been constructed using homogeneity adjustments in an attempt to minimize the effects of artificial changes.

3.2 Radiosonde temperature datasets

3.2.1 NOAA (RATPAC)

For several decades the 63 station dataset of Angell (Angell and Korshover, 1975) was the most widely used station-based radiosonde temperature dataset for climate monitoring. Recently, due to concerns regarding the effects of inhomogeneities, that network shrank to 54 stations (Angell, 2003). To better address these concerns, LKS (Lanzante, Klein, Seidel) (Lanzante et al., 2003a,b) built on the work of Angell by applying homogeneity adjustment to the time series from many of his stations, as well as several dozen additional stations, to create better regional representation via a network of 87 stations. However, because of the labor-intensive nature of the homogenization process on these 87 stations, extension of the LKS dataset beyond 1997 is impractical. Instead, the adjusted LKS dataset is being used as the basis for a new product (see Table 3.1), Radiosonde Air Temperature Products for Assessing Climate (RATPAC), that will be updated regularly (Free et al., 2003; Free et al., 2005). A NOAA group (a collaboration between the Air Resources Laboratory, the Geophysical Fluid Dynamics Laboratory, and NCDC) is responsible for the creation of RATPAC.

The RATPAC product consists of two parts: RATPAC-A and RATPAC-B[5], both of which use the adjusted LKS data, supplemented by an extension up to present using data from the Integrated Global Radiosonde Archive (IGRA). The IGRA data used in RATPAC are based on individual soundings that have been quality controlled and then averaged into monthly station data (Durre, 2005). In this report we use RATPAC-B. Generally speaking, based on data averaged over large regions such as the globe or tropics, trends from RATPAC-A and RATPAC-B are closer to one another than they are to the unadjusted (IGRA) data (Free et al., 2005).

3.2.2 UK (HadAT2)

For several decades the Oort (1983) product was the most widely used gridded radiosonde dataset. With the retirement of Abraham Oort, and cessation of his product, the dataset produced at the Hadley Centre, UK Met Office, HadRT (Parker et al., 1997) became the most widely used gridded product. Because of concern about the effects of artificial changes, this product incorporated homogeneity adjustments, although they were somewhat limited[6]. As a successor to HadRT, the Hadley Centre has created a new product (HadAT2, see Table 3.1) that uses all available digital radiosonde data for a larger network of almost 700 stations having relatively long records[7]. Identification and

[5] RATPAC-A uses the adjusted LKS data up through 1997 and provides an extension beyond that using a different technique to reduce the impact of inhomogeneities (Peterson et al., 1998). However, the RATPAC-A methodology can only be used to derive homogenized temperature averaged over many stations, and thus cannot be used to homogenize temperature time series at individual stations. RATPAC-B consists of the LKS adjusted station time series that have been extended beyond 1997 by appending (unadjusted) IGRA data.

[6] Adjustments were made to upper levels only (300 hPa and above), and since they were based on satellite data, only since 1979.

[7] High quality small station subsets, such as Lanzante et al. (2003a) and the Global Climate Observing System Upper Air Network, were used as a skeletal network from

adjustment of inhomogeneities was accomplished by way of comparison of neighboring stations.

3.2.3 Synopsis of radiosonde datasets

The two chosen datasets differ fundamentally in their selection of stations in that the NOAA dataset uses a relatively small number of highly scrutinized stations, while the UK dataset uses a considerably larger number of stations. Compared to the surface, far fewer thermometers are in use at any given time (hundreds or less) so there is less opportunity for random errors to cancel, but more importantly, there are far fewer suitable "neighbors" to aid in temporal homogenization. While both products incorporate a common building-block dataset (Lanzante et al., 2003a), their methods of construction differ considerably. Any differences in deductions about climate change derived from them could be attributed to both the differing raw inputs as well as differing construction methodologies. Concerns about poor temporal homogeneity are much greater than for surface data.

3.3 Global radiosonde temperature variations and differences between the datasets

3.3.1 Troposphere

Figure 3.2a displays $T_{(850\text{-}300)}$ time series for the RATPAC and HadAT2 radiosonde

---

which to define a set of adequately similar station series used in homogenization. The dataset is designed to impart consistency in both space and time and, by using radiosonde neighbors rather than satellites or reanalyses, minimizes the chances of introducing spurious changes related to the introduction of satellite data and their subsequent platform changes (Thorne et al., 2005).

datasets. Several noteworthy features are common to both. First, just as for the surface, ENSO signatures are clearly evident. Second, there is an apparent step-like rise of temperature around 1976-77 associated with the well-documented climate regime shift (Trenberth, 1990). Third, there is a long-term rise in temperatures, although a considerable amount of it may be due to the step-like change (Seidel and Lanzante, 2004). To a first approximation, both datasets display these features similarly and there is very little systematic difference between the two. Although a major component of the RATPAC product is used in the construction of the HadAT2 dataset, it should be kept in mind that the former utilizes a much smaller network of stations, although the length of the station records tends to be relatively long. If the good agreement is not fortuitous, this suggests that the reduced RATPAC station network provides representative spatial sampling[8].

[8] This result is consistent with the relatively large spatial scales represented by a single radiosonde station at this level on an annual time scale demonstrated by Wallis (1998) and Thorne et al. (2005).

Figure 3.2a – Bottom: Time series of globally averaged tropospheric temperature ($T_{(850-300)}$) for RATPAC (violet) and HadAT2 (green) radiosonde datasets. All time series are 7-month running averages (used as a smoother) of original monthly data, which were expressed as a departure (ºC) from the 1979-97 average.

Figure 3.2b – Top: Time series of globally averaged stratospheric temperature ($T_{(100-50)}$) for RATPAC (violet) and HadAT2 (green) radiosonde datasets. All time series are 7-month running averages (used as a smoother) of original monthly data, which were expressed as a departure (ºC) from the 1979-97 average.

3.3.2 Lower Stratosphere

Figure 3.2b displays global temperature anomaly time series of $T_{(100-50)}$ from the RATPAC and HadAT2 radiosonde datasets. Several noteworthy features are common to both datasets. First is the prominent signature of three climatically important volcanic eruptions: Agung (March 1963), El Chichon (April 1982), and Pinatubo (June 1991). Temperatures rise rapidly as volcanic aerosols are injected into the stratosphere and remain elevated for about 2-3 years before diminishing. There is some ambiguity as to whether the temperatures return to their earlier values or whether they experience step-like falls in the post-volcanic period for the latter two volcanoes, particularly Pinatubo (Pawson et al., 1998; Lanzante et al., 2003a; Seidel and Lanzante, 2004). Second, there are small amplitude variations associated with the tropical quasi-biennial oscillation (QBO) with a period of ~ 2-3 years (Seidel et al., 2004). Third, there is a downward trend, although there is some doubt as to whether the temperature decrease is best described by a linear trend over the period of record. For one thing, the temperature series prior to about 1980 exhibits little or no decrease in temperature. After that, the aforementioned step-like drops represent a viable alternative to a linear decrease (Seidel and Lanzante, 2004).

In spite of similarities among datasets, closer examination reveals some important differences. There is a rather large difference between RATPAC and HadAT2 time series for the peak volcanic warming associated with Agung in 1963. This may be a reflection of differences in spatial sampling because the horizontal pattern of the response is not uniform (Free and Angell, 2002). More noteworthy for estimates of climate change are some subtle systematic differences between the two datasets that vary over time. A closer examination reveals that the RATPAC product tends to have higher temperatures than the HadAT2 product from approximately 1963-85, with the RATPAC product having lower values before and after this time period[9]. As we will see later, this yields a slightly greater decreasing trend for the RATPAC product. Poorer agreement between the RATPAC and HadAT2 products in the stratosphere compared to the troposphere is not unexpected because of the fact that artificial jumps in temperature induced by changes in radiosonde instruments and measurement systems tend to increase in magnitude from the near-surface upwards (Lanzante et al., 2003b). More details on this issue are given in Chapter 4, Section 2.1.

4 SATELLITE-DERIVED TEMPERATURES

4.1 Microwave satellite data

Three groups, employing different methodologies, have developed satellite Microwave Sounding Unit (MSU) climate datasets (see Table 3.1). We do not present results from a fourth group (Prabhakara et al., 2000), which developed yet another methodology, since they are not continuing to work on MSU climate analyses and are not updating their time series. One of the main issues that is addressed differently by the groups is the inter-calibration between the series of satellites, and is discussed in Chapters 2 and 4.

4.2 Microwave Satellite Datasets

[9] It is worth noting that prominent artificial step-like drops, many of which were associated with the adoption of a particular type of radiosonde (Vaisala), were found in stratospheric temperatures at Australian and western tropical Pacific stations in the mid to late 1980s by Parker et al. (1997), Stendel et al. (2000), and Lanzante et al. (2003a Differences in consequent homogeneity adjustments around this time could potentially explain a major part of the difference between the NOAA and UK products, although this has not been demonstrated.

4.2.1 University of Alabama In Huntsville (UAH)

The first group to produce MSU climate products, by adjusting for the differences between satellites and the effects of changing orbits (diurnal drift), was UAH (A). Their approach (Christy et al., 2000; Christy et al., 2003) uses both an offset adjustment to allow for the systematic average differences between satellites and a non-linear hot target temperature[10] calibration to create a homogeneous series. The UAH dataset has products corresponding to three temperature measures: T_{2LT}, T_2, and T_4 (see Chapter 2 for definitions of these measures). In this report we use the most up to date versions available to us at the time, which is version 5.1 of the UAH dataset for T_2, and T_4, and version 5.2 for T_{2LT}[11].

4.2.2 Remote Sensing Systems (RSS)

After carefully studying the methodology of the UAH team, another group, RSS (R) created their own datasets for T_2 and T_4 using the same input data but with modifications to the adjustment procedure (Mears et al., 2003), two of which are particularly noteworthy: (1) the method of inter-calibration from one satellite to the next and (2) the computation of the needed correction for the daily cycle of temperature. While the second modification has little effect on the overall global trend differences between the two teams, the first is quite important in this regard. Recently the RSS team has created their own version of T_{2LT} (Mears and Wentz, 2005) and in doing so discovered a

[10] In fact, two targets are used, both with temperatures that are presumed to be well known. These are *cold* space, pointing away from the Earth, Moon, or Sun, and an onboard *hot* target.

[11] The version number for T_{2LT} differs from that for T_2, and T_4 because an error, which was found to affect the former (and was subsequently corrected), does not affect the latter two measures.

methodological error in the corresponding temperature measure of UAH. The UAH T_{2LT} product used in this report is based on their corrected method. In this report we use version 2.1 of the RSS data.

4.2.3 University of Maryland (U.Md.)

A very different approach (Vinnikov et al., 2004) was developed by a team involving collaborators from the University of Maryland and the NOAA National Environmental Satellite, Data, and Information Service (NESDIS) and was used to estimate globally averaged temperature trends (Vinnikov and Grody, 2003). After further study, they developed yet another new method (Grody et al., 2004). As done by the other two groups, the U.Md. (M) team's methodology also recalibrates the instruments based on overlapping data between the satellites. However, the manner in which they perform this recalibration differs. Also, they do not adjust for diurnal drift directly, but average the data from ascending and descending orbits. In their second approach, they substantially altered the manner in which target temperatures are used in their recalibration to a scheme more consistent with that of the other two groups (UAH and RSS). The effect of their revision was to reduce the global temperature trends derived from their data from 0.22-0.26 to 0.17 ºC/decade. Very recently they have revised their method to produce a third version of their dataset, which we use in this report, whose trends differ only slightly with those from the second version. In this most recent version they apply the nonlinear adjustment of Grody et al. (2004) and estimate the diurnal cycle as described in Vinnikov et al. (2005). The U.Md. group produces only a measure of T_2, hence there is no stratospheric product (T_4) or one corresponding to the lower troposphere (T_{2LT}).

4.3 Synopsis of satellite datasets

The relationship among satellite datasets is fundamentally different from that for surface or radiosonde products. For satellites, different datasets use virtually the same raw inputs so that any differences in derived measures are due to construction methodology. The excellent coverage provided by the orbiting sensors, more than half the Earth's surface daily, is a major advantage over in situ observations. The disadvantage is that while in situ observations rely on data from many hundreds or thousands of individual thermometers every day, providing a beneficial redundancy, the satellite data typically come from only one or two instruments at a given time. Therefore, any problem impacting the data from a single satellite can adversely impact the entire climate record. The lack of redundancy, compounded by occasional premature satellite failure that limits the time of overlapping measurements from successive satellites, elevates the issue of temporal homogeneity to the overwhelming explanation for any differences in deductions about climate change derived from the three datasets.

4.4 Global satellite temperature variations and differences between the datasets

4.4.1 Temperature of the Troposphere

Two groups (UAH and RSS) produce lower tropospheric temperature datasets, T_{2LT} (see Chapter 2 for definition of this and related temperature measures) directly from satellite measurements. Their time series are shown in Figure 3.3a along with an equivalent measure constructed from the HadAT2 radiosonde dataset (see Box 2.2 for an explanation as to how these equivalent measures were generated).The three temperature series have quite similar behavior, with ENSO-related variations accounting for much of the up and down meanderings, for example the historically prominent 1997-1998 El Niño. But over the full period of record, the amount of increase indicated by the datasets varies considerably. A closer look reveals that as time goes on, the RSS product indicates a noticeably greater increase of temperature than the other two. For comparison purposes, in Figure 3.3b we show an alternate measure of lower tropospheric temperatures, T^*_G , derived from products produced by the same three groups. From comparison of Figures 3.3a and 3.3b we see that both measures of lower tropospheric temperature agree remarkably well, even with regard to the more subtle differences relating to the longer-term changes. We will return to the issue of agreement between T_{2LT} and T^*_G later when we discuss trends (section 6).

Figure 3.3a– Bottom: Time series of globally averaged lower tropospheric temperature (T_{2LT}) as follows: UAH (blue) and RSS (red) satellite datasets, and HadAT2 (green) radiosonde data. All time series are 7-month running averages (used as a smoother) of original monthly data, which were expressed as a departure (ºC) from the 1979-97 average.

Figure 3.3b– Third: Time series of globally averaged middle tropospheric temperature (T^*_G) as follows: UAH (blue) and RSS (red) satellite datasets, and HadAT2 (green) radiosonde data. All time series are 7-month running averages (used as a smoother) of original monthly data, which were expressed as a departure (ºC) from the 1979-97 average.

Figure 3.3c – Second: Time series of globally averaged upper middle tropospheric temperature (T_2) as follows: UAH) (blue), RSS (red), and U.Md. (black) satellite datasets, and HadAT2 (green) radiosonde data. All time series are 7-month running averages (used as a smoother) of original monthly data, which were expressed as a departure (ºC) from the 1979-97 average.

Figure 3.3d – Top: Time series of globally averaged lower stratospheric temperature (T_4) as follows: UAH (blue) and RSS (red) satellite datasets, and HadAT2 (green) radiosonde data. All time series are 7-month running averages (used as a smoother) of original monthly data, which were expressed as a departure (ºC) from the 1979-97 average.

Time series corresponding to the temperature of the upper middle troposphere (T_2) are shown in Figures 3.3c. The products represented in this figure are the same as for the lower troposphere, except that an additional product, that from the U.Md. group is available. Again, all of the time series have similar behavior with regard to the year to year variations. However, closer examination shows that two of the products (U.Md. and RSS satellite data) indicate considerable temperature increase over the period of record, whereas the other two (UAH satellite and HadAT2 radiosonde) indicate slight warming only. A more detailed discussion of the differences between the various products can be found in Chapter 4.

We note that all of the curves for the various tropospheric temperature series (Figures 3.3a-c) exhibit remarkably similar shape over the period of record. For the common time period, the satellite measures are similar to the tropospheric layer-averages computed from radiosonde data. The important differences between the various series are with regard to the more subtle long-term evolution over time, which manifests itself as differences in linear trend, discussed later in more detail.

4.4.2 Temperature of the Lower Stratosphere

Figure 3.3d shows the temperature of the lower stratosphere (T_4); note that there is no product from the U.Md. team for this layer. The dominant features for this layer are the major volcanic eruptions: El Chichon in 1982 and Pinatubo in 1991. As discussed above, the volcanic aerosols tend to warm the stratosphere for about 2-3 years before diminishing. In contrast, ENSO events have little influence on the stratospheric temperature. Both products show that the stratospheric temperature has decreased considerably since 1979, as compared to the lesser amount of increase that is seen in the troposphere. The T_{4-R} product shows somewhat less overall decrease than the T_{4-A} product, in large part as a result of the fact that the former increases relative to the latter from about 1992-94. As was the case for the troposphere, the radiosonde series show a greater decrease than the satellite data. Again, the satellite and radiosonde series for the lower-stratosphere exhibit the same general behavior over time.

5 REANALYSIS TEMPERATURE "DATA"

A number of agencies from around the world have produced reanalyses based on different schemes for different time periods. We focus on two of the most widely referenced, which cover a longer time period than the others (see Table 3.1). The U.S. reanalysis represents a collaborative effort between NOAA's National Center for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR). For U.S. reanalysis, gridded air temperatures at the surface and aloft are available from 1958 to present. Using a completely different system, the European Center for Medium-Range Weather Forecasts (ECMWF) has produced similar gridded data from September 1957 to August 2002 (Europe). Reanalyses are “hybrid products,” utilizing raw input data of many types, as well as complex mathematical models to combine these data. For more detailed information on the reanalyses, see Chapter 2. As the reanalysis output does not represent a different observing platform, a separate assessment of reanalysis data will not be made.

6. COMPARISONS BETWEEN DIFFERENT LAYERS AND OBSERVING PLATFORMS

6.1 During the radiosonde era, 1958 to the present

6.1.1 Global

As shown in earlier sections, globally averaged temperature time series indicate increasing temperature at the surface and in the troposphere with decreases in the stratosphere over the course of the last several decades. It is desirable to derive some estimates of the magnitude of the rate of these changes. The widely-used, least-squares, linear trend technique is adopted for this purpose with the explicit caveat that long-term changes in temperature are not necessarily linear, as there may be departures in the form of periods of enhanced or diminished change, either linear or nonlinear, as well as abrupt, step-like changes[12]. While it has been shown that such constructs are plausible, it is nevertheless difficult to prove that they provide a better fit to the data, over the time periods addressed in this report, than the simple linear model (Seidel and Lanzante, 2004). Additional discussion on this topic is given in the Statistical Appendix.

Trends computed for the radiosonde era are given in Table 3.2 for the surface as well as various tropospheric and stratospheric layer averages[13]. The surface products are quite consistent with one another, as are the radiosonde products in the troposphere. In the stratosphere, the radiosonde products differ somewhat, although there is an inconsistent relationship involving the two stratospheric measures ($T_{(100\text{-}50)}$ and T_4) regarding which product indicates a greater decrease in temperature[14]. The reanalysis products, which are "hybrid-measures," agree better with the "purer" surface and radiosonde measures at and near the surface. Agreement degrades with increasing altitude such that the reanalyses indicate more tropospheric temperature increase and considerably less stratospheric decrease than do the radiosonde products. The disparity between the reanalyses and other products is not surprising given the suspect temporal homogeneity of the reanalyses (see Chapter 2, Section 1c).

[12] For example, the tropospheric linear trends in the periods 1958-1979 and 1979-2003 were shown to be much less than the trend for the full period (1958-2003), based on one particular radiosonde dataset (Thorne et al., 2005), due to the abrupt rise in temperature in the mid 1970s.

[13] Note that it is instructive to examine the behavior of radiosonde and reanalysis temperatures averaged in such a way as to correspond to the satellite layers (T_{2LT}, T^*_G, T_2, and T_4) even though there are no comparable satellite measures prior to 1979.

[14] The reason for this inconsistency is that the HadAT2 product records data at fewer vertical levels than the RATPAC product, so the comparison is not one-to-one.

Table 3.2 - Global temperature trends in ºC per decade from 1958 through 2004 (except for European which terminates September 2001) calculated for the surface or atmospheric layers by data source. The trend is shown for each, with the approximate 95% confidence interval (2 sigma) below in parentheses. The levels/layers, from left to right, go from the lowest to the highest in the atmosphere. Bold values are estimated to be statistically significantly different from zero (at the 5% level). A Student's t-test, using the lag-1 autocorrelation to account for the non-independence of residual values about the trend line, was used to assess significance (see Appendix for discussion of confidence intervals and significance testing).

Perhaps the most important result shown in Table 3.2 is that both the radiosonde and reanalysis trends indicate that the tropospheric temperature has increased as fast as or faster than the surface over the period 1958 to present. For a given dataset, the 3 measures (T_{2LT}, $T_{(850\text{-}300)}$, and T^*_G) always indicate more increase in the troposphere than at the surface, although this is usually not true when the T_2 measure is considered. The reason for the inconsistency involving T_2 is because of contributions to the layer that it measures from stratospheric cooling, an effect first recognized by Spencer and Christy (1992) (see discussion of this issue in Chapters 2 and 4). The development of T^*_G as a global measure, and its counterpart, T^*_T for the tropics (Fu et al., 2004; Fu and Johanson, 2005; Johanson and Fu, 2005) was an attempt to remove the confounding effects of the stratosphere using a statistical approach (see Chapter 2).

6.1.2 Land vs. ocean

Most of the land and ocean surface temperature increased during the radiosonde era, with the exception of parts of the North Atlantic Ocean, the North Pacific Ocean, and a few smaller areas. With a few exceptions, such as the west coast of North America, trends in land air temperature in coastal regions are generally consistent with trends in SST over neighboring ocean areas (Houghton et al., 2001). Because bias adjustments are performed separately for land and ocean areas, before merging to create a global product, it is unlikely that the land-ocean consistency is an artifact of the construction methods used in the various surface analyses. However, land air temperatures did increase somewhat more rapidly than SSTs in some regions during the past two decades. Possibly related to this is the fact that since the mid-1970s, El Niño has frequently been in its "warm" phase, which tends to bring higher than normal temperatures to much of North America, among other regions, which have had strong temperature increases over the past few decades (Hurrell, 1996). Also, when global temperatures are rising or falling, the global mean land temperature tends to both rise and fall faster than the ocean, which has a tremendous heat storage capacity (Waple and Lawrimore, 2003).

6.1.3 Marine air vs. sea surface temperature

In ocean areas, it is natural to consider whether the temperature of the air and that of the ocean surface (SST) increases or decreases at the same rate. Several studies have examined this question. Overall, on seasonal and longer scales, the SST and marine air temperature generally move at about the same rate globally and in many ocean basin scale regions (Bottomley et al., 1990; Parker et al., 1995; Folland et al., 2001b; Rayner et al., 2003). Differences between SST and marine air temperature in some regions were first noted by Christy et al. (1998) and then examined in more detail by Christy et al. (2001). The latter study found that in the tropics, SST increased more than NMAT from 1979 –1999 derived from the Tropical Atmosphere Ocean (TAO) array of tropical buoys and transient marine ship observations. But this difference may be related to changes in surface fluxes associated with ENSO and the interdecadal Pacific oscillation (Folland et al., 2003). Consistent results were found using two datasets, one with more widespread observations from ships, and another, which sampled a more limited number of locations using moored buoys. There were some indications that the accelerated increase of SST compared to air temperature may have been concentrated in two periods: the early 1980s and mid 1990s. So over the satellite era, there are some unexplained differences in these trends that were also noted by Folland et al. (2003) in parts of the tropical south Pacific using the Rayner et al. (2003) NMAT dataset which incorporates new corrections for the effect on NMAT of increasing deck (and hence measurement) heights.

6.1.4 Minimum vs. maximum temperatures over land

Daily minimum temperature increased about twice as fast as daily maximum temperature over global land areas during the radiosonde era (Karl et al., 1993; Easterling et al., 1997; Folland et al, 2001b). However, a closer look at recent years has found that during the satellite era, maximum and minimum temperature have been rising at the same rate (Vose et al, 2005). Daily minimum temperature increased in virtually all areas except eastern Canada, Eastern Europe, and other scattered regions often near coasts. Most regions also witnessed an increase in the daily maximum, but over the longer time frame the rate of increase was generally smaller, and decreasing trends were somewhat more common (e.g., in eastern Canada, the southern United States, southern China, eastern Europe, and portions of South America). The causes of this asymmetric warming are still debated, but many of the areas with greater increases of minimum temperatures correspond to those where cloudiness appears to have increased over the period as a whole (Dai et al., 1999; Henderson-Sellers, 1992; Sun and Groisman, 2000; Groisman et al., 2004). This makes physical sense since clouds tend to cool the surface during the day by reflecting incoming solar radiation, and warm the surface at night by absorbing and reradiating infrared radiation back to the surface.

6.2 During the satellite era, 1979 to the present

6.2.1 Global

A comparable set of global trends for the satellite era is given in Table 3.3. Comparison between Tables 3.2 and 3.3 reveals that some of the relationships between levels and layers, as well as among datasets, are different during the two eras. Comparing satellite era trends with the radiosonde era trends for datasets that have both periods in common, it is clear that the surface temperature increase (see Figure 3.1) has accelerated in recent decades while the tropospheric increase (see Figure 3.2a) has decelerated. Since most of the stratospheric decrease has occurred since 1979 (see Figure 3.2b) the rate of temperature decrease there is close to twice as large as during the full radiosonde era. Thus, care must be taken when interpreting results from only the most recent decades. Agreement among different surface and radiosonde datasets is reasonable and about as good as during the longer radiosonde era. The reanalysis datasets show poorer agreement with surface data and especially with stratospheric radiosonde data for the European product.

Table 3.3 - Global temperature trends in ºC per decade from 1979 through 2004 (except for European which terminates September 2001) calculated for the surface or atmospheric layers by data source. The trend is shown for each, with the approximate 95% confidence interval (2 sigma) below in parentheses. The levels/layers, from left to right, go from the lowest to the highest in the atmosphere. Bold values are estimated to be statistically significantly different from zero (at the 5% level). A Student's t-test, using the lag-1 autocorrelation to account for the non-independence of residual values about the trend line, was used to assess significance (see Appendix for discussion of confidence intervals and significance testing).

numerical trend values found in Table 3.3, which are also graphed in Figure 3.4a. While the tropospheric satellite products from the UAH team have trends that are not too dissimilar from the corresponding radiosonde trends, the two other satellite datasets show a considerably greater increase in tropospheric temperature. In the stratosphere, there is a large disagreement between satellite and radiosonde products, with the latter indicating much greater decreases in temperature. Here too, the reanalyses are quite inconsistent, with the European product closer to the satellites and the U.S. product closer to the radiosondes.

Figure 3.4a (top) – Global temperature trends (ºC/decade) for 1979-2004 from Table 3.3 plotted as symbols. See figure legend for definition of symbols. Filled symbols denote trends estimated to be statistically significantly different from zero (at the 5% level). A Student's t-test, using the lag-1 autocorrelation to account for the non-independence of residual values about the trend line, was used to assess significance (see Appendix for discussion of confidence intervals and significance testing).

Figure 3.4b (bottom) – Tropical (20ºN-20ºS) temperature trends (ºC/decade) for 1979-2004 from Table 3.4 plotted as symbols. See figure legend for definition of symbols. Filled symbols denote trends estimated to be statistically significantly different from zero (at the 5% level). A Student's t-test, using the lag-1 autocorrelation to account for the non-independence of residual values about the trend line, was used to assess significance (see Appendix for discussion of confidence intervals and significance testing).

Perhaps the most important issue is the relationship between trends at the surface and in the troposphere. As shown in Table 3.3 and Figure 3.4a, both radiosonde datasets as well as the UAH satellite products indicate that, in contrast to the longer radiosonde era, during the satellite era the temperature of the surface has increased more than that of the troposphere. However, tropospheric trends from the RSS satellite dataset, based on both measures of temperature having little or no stratospheric influence (T_{2LT} and T^*_G) yield an opposing conclusion: the tropospheric temperature has increased as much or more than the surface. For the third satellite dataset, comparisons with surface temperature are complicated by the fact that the U.Md. team produces only T_2, which is influenced by stratospheric cooling (see Chapter 2). Nevertheless, we can infer that it too suggests more of a tropospheric temperature increase than that at the surface[15].

Since global change theory suggests more warming of the troposphere than the surface only in the tropics (see Chapter 1), much of the interest in observed trends has been in this region. Therefore, to compliment the global trends (Figure 3.4a and Table 3.3), we present a similar plot of tropical trends in Figure 3.4b (with corresponding trend values in

[15] The difference in trends, T^*_G minus T_2, for the UAH and RSS datasets is about 0.06 to 0.08 ºC/decade. Adding this amount to the U.Md. T_2 trend (0.20 ºC/decade) yields an estimate of the U.Md. trend in T^*_G of about 0.26 to 0.28 ºC/decade. In this calculation we are assuming that the effects of the stratospheric cooling trend on the U.Md. product are the same as from the UAH and RSS datasets.

Table 3.4). Compared to the global trends, the tropical trends show even more spread among datasets, particularly in the lower stratosphere[16]. The result of the greater spread is that the range of plausible values for the difference in trends between the surface and troposphere is larger than that for the globe as a whole. Similar to the global case, in the tropics the UAH satellite plus the two radiosonde datasets (RATPAC and HadAT2) suggest more warming at the surface than in the troposphere, while the opposite conclusion is reached based on the other two satellite products (RSS and U.Md.). Resolution of this issue would seem to be of paramount importance in the interpretation of observed climate change central to this synthesis assessment.

Table 3.4 – Tropical (20ºN-20ºS) temperature trends in ºC per decade from 1979 through 2004 (except for European which terminates September 2001) calculated for the surface or atmospheric layers by data source. The trend is shown for each, with the approximate 95% confidence interval (2 sigma) below in parentheses. The levels/layers, from left to right, go from the lowest to the highest in the atmosphere. Bold values are estimated to be statistically significantly different from zero (at the 5% level). A Student's t-test, using the lag-1 autocorrelation to account for the non-independence of residual values about the trend line, was used to assess significance (see Appendix for discussion of confidence intervals and significance testing).

[16] The larger spread may be partially an artifact of the fact that when averaging over a smaller region, there is less cancellation of random variations. In addition, the fact that the networks of in situ observations are much sparser in the tropics than in the extratropics of the Northern Hemisphere may also contribute.

RATPAC	**0.13** (0.068)	0.08 (0.119)	0.06 (0.136)	0.07 (0.153)	0.00 (0.140)	**-0.75** (0.362)	**-0.69** (0.289)
HadAT2	**0.15** (0.115)	0.05 (0.152)	0.03 (0.164)	0.02 (0.176)	-0.04 (0.170)	**-0.66** (0.304)	**-0.64** (0.307)
Satellite:							
UAH		0.05 (0.176)		0.09 (0.191)	0.05 (0.167)		**-0.37** (0.281)
RSS		0.15 (0.192)		0.18 (0.196)	0.14 (0.175)		-0.29 (0.303)
U.Md.					**0.19** (0.159)		
Reanalyses:							
US	0.03 (0.163)	0.05 (0.172)	0.04 (0.173)	-0.03 (0.183)	-0.10 (0.166)	**-0.89** (0.405)	**-0.83** (0.340)
European	0.03 (0.211)	0.00 (0.234)	-0.03 (0.249)	0.06 (0.255)	0.05 (0.232)	-0.03 (0.453)	-0.05 (0.423)

6.2.2 Latitude bands

Globally averaged temperatures paint only part of the picture. Different layers of the atmosphere behave differently depending on the latitude. Furthermore, even the processing of the data can make for latitudinal difference in long-term trends. Figure 3.5 shows the trends in temperature for different datasets and levels averaged over latitude bands. Each of these trends was created by making a latitudinally averaged time series of monthly anomalies and then fitting that time series with a standard least-squares linear regression slope.

Figure 3.5 -- Temperature trends for 1979-2004 (ºC/decade) by latitude.
Left: stratospheric temperature (T_4) based on RSS (red) and UAH (blue) satellite datasets, and RATPAC (violet) and HadAT2 (green) radiosonde datasets.
Middle: mid-tropospheric temperature (T_2) based on U.Md. (orange), RSS (red) and UAH (blue) satellite datasets, and RATPAC (violet) and HadAT2 (green) radiosonde datasets; and surface temperature (T_S) from NOAA data (black).
Right: surface temperature (T_S) from NOAA data (black) and lower tropospheric temperature (T_{2LT}) from RSS (red) and UAH satellite data (blue), and from RATPAC (violet) and HadAT2 (green) radiosonde data. Filled circles denote trends estimated to be statistically significantly different from zero (at the 5% level). A Student's t-test, using the lag-1 autocorrelation to account for the non-independence of residual values about the trend line, was used to assess significance (see Appendix for discussion of confidence intervals and significance testing).

In the stratosphere (left panel of Figure 3.5), trend profiles for the two satellite datasets are fairly similar, with a greater temperature decrease everywhere according to $T_{4\text{-}A}$ than $T_{4\text{-}R}$. Some of the largest temperature decrease occurs in the South Polar Region, where ozone depletion is largest. A broad region of weaker decrease occurs in the deep tropics. By contrast, the RATPAC and HadAT2 radiosonde datasets are quite different from the satellite products, with much flatter profiles. It is worth noting that there is a fundamental disagreement between the radiosonde and satellite products. Except for the mid-latitudes of the Northern Hemisphere[17], at most other latitudes the radiosonde products show more of a temperature decrease than the satellite products, with the largest discrepancy in the tropics[18].

For the middle troposphere (middle panel of Figure 3.5) there is general agreement among the radiosonde and satellite datasets in depicting the same basic structure. The largest temperature increase occurs in the extratropics of the Northern Hemisphere, with a smaller increase or slight decrease in the tropics, and even lesser increase or more decrease in the extratropics of the Southern Hemisphere. At most latitudes, $T_{2\text{-}M}$ indicates the most increase (least decrease), followed next by $T_{2\text{-}R}$, then $T_{2\text{-}A}$, and finally the radiosonde products with the least increase (most decrease).

[17] The apparently better radiosonde-satellite agreement in the midlatitudes of the Northern Hemisphere may be the result of spurious stratospheric warming at stations located in countries of the former Soviet Union, offsetting the more typical spurious cooling bias of radiosonde temperatures (Lanzante et al., 2003a,b).
[18] We note that in the tropics, where the radiosonde and satellite products differ the most, abrupt artificial drops in temperature appear to be particularly problematic for radiosonde data (Parker et al., 1997; Lanzante et al., 2003a,b). Other studies (Sherwood et al., 2005; Randel and Wu, 2005) also suggest spurious cooling for radiosonde temperatures, especially in the tropics.

For the lower troposphere and surface (right panel of Figure 3.5) the profiles are roughly similar in shape to those for the middle troposphere with one major exception: the higher-latitude temperature increase of the Northern Hemisphere is more pronounced compared to the other regions. Comparing the surface temperature trend profile (black) with that from the various tropospheric products in the middle and right panels of Figure 3.5 suggests that the sign and magnitude of this difference is highly dependent upon which tropospheric measure is used.

6.2.3 Maps

Trend maps represent the finest spatial granularity with which different levels/layers and observing platforms can be compared. However, since maps may not be the optimal way in which to examine trends[19], we present only a limited number of such maps for illustrative purposes. Figure 3.6 presents maps of trends for the surface (bottom), lower troposphere (second from bottom), upper middle troposphere (second from top), and stratosphere (top). The surface map is based on the NOAA dataset[20] while those for the troposphere and stratosphere are based on the RSS satellite dataset[21]. In examining these

[19] Averaging over space (e.g., over latitudes, the tropics or the globe, as presented earlier) tends to reduce noise that results from the statistical uncertainties inherent to any observational measurement system. Furthermore, models that are used to study climate change have limited ability to resolve the smallest spatial scales and therefore there is little expectation of detection at the smallest scales (Stott and Tett, 1998). The formal methodology that is used to compare models with observations ("fingerprinting," see Chapter 5) concentrates on the larger-scale signals in both models and observations in order to optimize the comparisons.

[20] Trend maps from other surface datasets (not shown) tend to be fairly similar to that of the NOAA map, differing mostly in their degree of spatial smoothness, which is a function of dataset construction methodology.

[21] A comparison between UAH and RSS trend maps for tropospheric layers is given in Chapter 4.

maps it should be kept in mind that based on theory *we expect* the difference in trend between the surface and troposphere to vary by location. For example, as shown in Chapter 1, climate model projections typically indicate that human induced changes should lead to more warming of the troposphere than the surface in the tropics, but the opposite in the Arctic and Antarctic.

Figure 3.6 – Temperature trends for 1979-2004 (^{o} C /decade).
Bottom (d): NOAA surface temperature (T_{S-N}).
Third (c): RSS lower tropospheric temperature (T_{2LT-R}).
Second (b): RSS upper middle tropospheric temperature (T_{2-R}).

Top (a): RSS lower stratospheric temperature ($T_{4\text{-}R}$).

The trend maps indicate both similarities and differences between the surface and tropospheric trend patterns. There is a rough correspondence in patterns between the two. The largest temperature increase occurs in the extratropics of the Northern Hemisphere, particularly over landmasses. A decreases or smaller increase is found in the high latitudes of the Southern Hemisphere as well as in the eastern tropical Pacific. Note the general correspondence between the above noted features in Figures 3.6c,d and the zonal trend profiles (middle and right panels of Figure 3.5). Note that the upper middle tropospheric temperature is somewhat of a hybrid measure, being affected most strongly by the troposphere, but with a non-negligible influence by the stratosphere.

In contrast to the surface and troposphere, a temperature decrease is found almost everywhere in the stratosphere (Figure 3.6a). The largest decrease is found in the midlatitudes of the Northern Hemisphere and the South Polar Region, with a smaller decrease in the tropics. Again note the correspondence between the main features of the trend map (Figure 3.6a) and the corresponding zonal trend profiles (left panel of Figure 3.5).

7. CHANGES IN VERTICAL STRUCTURE

7.1 Vertical profiles of trends

Up to this point, our vertical comparisons have contrasted trends of surface temperature with trends based on different layer-averaged temperatures. Layers are useful because the averaging process tends to reduce noise. The use of layer-averages is also driven by the limitations of satellite measurement systems that are unable to provide much vertical detail. However, as illustrated in Chapter 1, changes in various forcing agents can lead to more complex changes in the vertical. Radiosonde data, because of their greater vertical resolution, are much better suited for this than currently available satellite data.

Figure 3.7 shows vertical profiles of trends from the RATPAC and HadAT2 radiosonde datasets for temperature averaged over the globe (top) or tropics (bottom) for the radiosonde (left) and satellite (right) eras. The trend values of Figure 3.7 are also given in Table 3.5. Each graph has profiles for the two radiosonde datasets. The tropics are of special interest because many climate models suggest that under global warming scenarios trends should increase from the lower troposphere upwards, maximizing in the upper troposphere (see Chapters 1 and 5).

Table 3.5 – Temperature trends in ºC per decade from the RATPAC and HadAT2 radiosonde datasets corresponding to the plots in Figure 3.7 (see figure caption for further details). Global and tropical trends are given for 1958 through 2004 and 1979 through 2004 (except for European which terminates September 2001). The HadAT2 dataset does not have temperatures for some of the levels, hence the empty table cells. The trend is shown for each vertical level (hPa), with the approximate 95% confidence interval (2 sigma) below in parentheses. Bold values are estimated to be statistically significantly different from zero (at the 5% level). A Student's t-test, using the lag-1 autocorrelation to account for the non-independence of residual values about the trend line, was used to assess significance (see Appendix for discussion of confidence intervals and significance testing).

Figure 3.7 -- Vertical profiles of temperature trend (ºC/decade) as a function of altitude (i.e., pressure in hPa) computed from the RATPAC (violet) and HadAT2 (green) radiosonde datasets. Trends (which are given in Table 3.5) have been computed for 1958-2004 (left) and 1979-2004 (right) based on temperature that has been averaged over the globe (top) or the tropics, 20^{o}N-20^{o}S (bottom). Surface data for the HadAT2 product is taken from HadCRUT2v since the HadAT2 dataset does not include values at the surface; the surface values have been averaged so as to match their observing locations with those for the radiosonde data. By contrast, the surface temperatures from the RATPAC product are those from the RATPAC dataset, which are surface station values reported with the radiosonde data. Note that these differ from the NOAA surface dataset values (ER-GHCN-ICOADS) as indicated in Table 3.1. Filled symbols denote trends estimated to be statistically significantly different from zero (at the 5% level). A Student's t-test, using the lag-1 autocorrelation to account for the non-independence of residual values about the trend line, was used to assess significance (see Appendix for discussion of confidence intervals and significance testing).

For the globe, the figure indicates that during the longer period the tropospheric temperature increased slightly more than that of the surface. By contrast, for the globe during the satellite era, the surface temperature increased more than that of the troposphere. Both datasets agree reasonably well in these conclusions. For the tropics, the differences between the two eras are more pronounced. For the longer period there is good agreement between the two datasets in that the temperature increase is smaller at the surface and maximized in the upper troposphere. The largest disagreement between datasets and least amount of tropospheric temperature increase is seen in the tropics during the satellite era. For the RATPAC product, the greatest temperature increase occurs at the surface with a slight increase (or decrease) in the lower and middle troposphere followed by somewhat larger increase in the upper troposphere. The HadAT2 product also shows largest increase at the surface, with a small increase in the troposphere, however, it lacks a distinct return to increase in the upper troposphere. In summary, the two datasets have fairly similar profiles in the troposphere with the exception of the tropics during the satellite era[22]. For the stratosphere, the decrease in temperature is noticeably greater for both the globe and the tropics during the satellite than radiosonde era as expected (see Figure 3.2b). Some of the largest discrepancies between datasets are found in the stratosphere.

7.2 Lapse rates

Temperature usually decreases in the troposphere going upward from the surface. Lapse

[22] However, the differences between datasets may not be meaningful since they are small compared to the statistical uncertainty estimates (see Table 3.5 and discussion in the Appendix).

rate is defined as the rate of decrease in temperature with increasing altitude and is a measure of the stability of the atmosphere[23]. Most of the observational work to date has not examined lapse rates themselves, but instead has used an approximation in the form of a vertical temperature difference[24]. This difference has taken on the form of the surface temperature minus some tropospheric temperature, either layer-averaged (in the case of satellite data) or at some specific pressure level (in the case of radiosonde data)[25].

Much of the interest in lapse rate variations has focused on the tropics. Several studies (Brown et al., 2000; Gaffen et al., 2000; Hegerl and Wallace, 2002; Lanzante et al., 2003b) present time series related to tropical lapse rate based on either satellite or radiosonde measures of tropospheric temperature. As examples, we present some such time series in Figure 3.8, based on measures of lower tropospheric temperature from three different datasets. Some essential low-frequency characteristics are common to all. A considerable proportion of the variability of the tropical lapse rate is associated with ENSO[26], a manifestation of which are the up and down swings of about 3-7 years in the

[23] A larger lapse rate implies more unstable conditions and a greater tendency towards vertical mixing of air.

[24] The reasons for this are two-fold: (1) satellite measurement systems are only able to resolve temperatures in deep layers rather than at specific levels, and (2) radiosonde measurements are consistently recorded at a fixed number of constant pressure rather than height levels.

[25] When constant pressure level data from radiosondes are used, the resulting lapse rate quantity may be influenced by changes in the thickness (i.e., average temperature) of the layer. However, some calculations by Gaffen et al. (2000) suggest that thickness changes do not have very much influence. Therefore, we consider vertical temperature differences to be a suitable approximation of lapse rate

[26] Lapse rate changes occur about five to six months after a particular change in ENSO (Hegerl and Wallace, 2002; Lanzante et al., 2003b). During a tropical warming event (El Niño) the tropical troposphere warms relative to the surface; the opposite is true during a tropical cooling event (La Niña).

series shown in Figure 3.8. Another feature evident in the four studies cited above, and seen in Figure 3.8 as well, is an apparent strong association with the climate regime shift that occurred ~1976-77 (Trenberth and Hurrell, 1994). There is a rather sharp drop in tropical lapse rate at this time[27], coincident with an abrupt change in a measure of convective stability (Gettelman et al. 2002). Overall, the variation in tropical lapse rate can be characterized as highly complex, with rapid swings over a few years, superimposed upon persistent periods of a decade or more, as well as longer-term drifts or trends evident during some time periods.

Figure 3.8 - Time series of vertical temperature difference (surface minus lower troposphere) for the tropics (20°N-20°S). NOAA surface temperatures (T_{S-N}) are used in each case to compute differences with lower tropospheric temperature (T_{2LT}) from three different groups: HadAT2 radiosonde (green), RSS satellite (red), and UAH satellite (blue). All time series are 7-month running averages (used as a smoother) of original monthly data, which were expressed as a departure (°C) from the 1979-97 average.

[27] Lanzante et al. (2003b) also noted an apparent decrease in the amplitude of ENSO-related tropical lapse rate variations after the ~1976-77 regime shift.

The feature of the tropical lapse rate series that has drawn the most interest is the linear trend component during the satellite era. From a long historical perspective (see also Figure 3.8), this trend is a rather subtle feature, being overshadowed by both the ENSO-related variations as well as the regime shift of the late 1970s. Several studies (Brown et al., 2000; Gaffen et al., 2000; Hegerl and Wallace, 2002; Lanzante et al., 2003b) have estimated trends in lower tropospheric lapse rate while another (Christy et al., 2001) has estimated trends in the difference between SST and surface air temperature.

The different trend estimates vary considerably among the above-cited studies, being dependent upon the details of the calculations[28]. From the cited studies, satellite-era trends in lapse rate based on temperatures averaged over the tropics range from nearly zero (no change) to about 0.20ºC/decade (surface warms more than the troposphere). The time series of Figure 3.8 also exhibit a wide range of satellite-era trends[29]. During the longer radiosonde era, the various studies found trends of opposite sign (i.e., air temperature at the surface increases more slowly than that of air aloft) and show less

[28]These details include: time period, latitude zone, datasets utilized, station network vs. grid, time of day of observations, use of homogeneity adjustment, and whether or not measurements in the troposphere and surface were taken from the same locations. Particularly noteworthy is the fact that Lanzante et al. (2003b) found that during the satellite era, use of homogenized data could, depending the other details of the analysis, either halve or eliminate the positive tropical lapse rate trend found using the unadjusted data.

[29] Trends from 1979 to 2004 (ºC /decade) for the three time series in Figure 3.8 are: 0.11 (HadAT2 radiosonde), 0.08 (UAH satellite), and -0.02 (RSS satellite). While the first two of these trends are statistically significant at the 5% level, the third is not (see Appendix for discussion of significance testing).

sensitivity, with a range of values of near-zero to about -0.05ºC/decade[30].

Spatial variations in lapse rate trends have also been examined. During the satellite era, some have found predominantly increasing trends in the tropics (Gaffen et al., 2000; Brown et al., 2000) while others have found a greater mixture, with more areas of negative trends (Hegerl and Wallace, 2002; Lanzante et al., 2003b). Outside of the tropics, both Hegerl and Wallace (2002) and Lanzante et al. (2003b) found complex spatial patterns of trend. Lanzante et al. (2003b) also found considerable local sensitivity to homogeneity adjustment in the tropics and even more so over the extratropics of the Southern Hemisphere, which is quite sparsely sampled.

[30] The trend from 1958 to 2004 for the HadAT2 radiosonde series shown in Figure 3.8 is -0.02 ºC/decade. This trend is not statistically significant at the 5% level (see Appendix for discussion of significance testing).

REFERENCES

Aguilar, E., I. Auer, M. Brunet, T. C. Peterson and J. Wieringa, 2003: *Guidelines on Climate Metadata and Homogenization, WCDMP-No. 53, WMO-TD No. 1186.* World Meteorological Organization, Geneva, 55 pp.

Andersson, E., and Coauthors, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). Part III: Experimental results. *Quarterly Journal of the Royal Meteorological. Society*, **124**, 1831-1860.

Angell, J. K., and J. Korshover, 1975: Estimate of the global change in tropospheric temperature between 1958 and 1973. *Monthly Weather Review*, **103**, 1007-1012.

Angell, J. K., 2003: Effect of exclusion of anomalous tropical stations on temperature trends from a 63-station radiosonde network, and comparison with other analyses. *Journal of Climate*, **16**, 2288-2295.

Bottomley, M., Folland, C.K., Hsiung, J., Newell, R.E. and D.E. Parker, 1990: *Global Ocean Surface Temperature Atlas "GOSTA"*. Joint project of Met Office and Massachusetts Institute of Technology supported by US Dept. of Energy, US National Science Foundation and US Office of Naval Research. HMSO, London, 20pp+iv, 313 plates.

Brown, S., D. Parker, C. Folland, and I. Macadam, 2000: Decadal variability in the lower-tropospheric lapse rate. *Geophysical Research Letters*, **27**, 997-1000.

Christy, J., D. Parker, S. Brown, I. Macadam, M. Stendel, and W. Norris, 2001: Differential trends in tropical sea surface and atmospheric temperature since 1979. *Geophysical Research Letters*, **28(1)**, 183-186.

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**, 2016-2041.

Christy, J. R., R. W. Spencer, and W. D. Braswell, 2000: MSU tropospheric temperatures: dataset construction and radiosonde comparisons. *Journal of Atmospheric and Oceanic Technology*, **17**, 1153-1170.

Christy, J. R., R. W. Spencer, W. B. Norris, W. D. Braswell, and D. E. Parker, 2003: Error estimates of version 5.0 of MSU-AMSU bulk atmospheric temperatures. .

Journal of Atmospheric and Oceanic Technology, **20**, 613-629.

Dai, A., K.E. Trenberth, and T.R. Karl, 1999: Effects of clouds, soil moisture, precipitation, and water vapor on diurnal temperature range. *Journal of Climate*, **12**, 2451-2473.

Diaz, H.F**.,** Folland, C.K., Manabe, T., Parker, D.E., Reynolds, R.W. and Woodruff, S.D., 2002: Workshop on Advances in the Use of Historical Marine Climate Data (Boulder, Co., USA, 29th Jan - 1st Feb 2002). *WMO Bulletin*. **51**, 377-380

Durre, I., R. S. Vose, and D. B. Wuertz, 2005: Overview of the Integrated Global Radiosonde Archive. *Journal of Climate,* submitted.

Easterling, D.R., T.R. Karl, E.H. Mason, P.Y. Hughes, and D.P. Bowman, 1996: *United States Historical Climatology Network (U.S. HCN), monthly temperature and precipitation data*. (Environmental Sciences Division Publication no. 4500; ORNL/CDIAC 87; NDP-0l9/R3). Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, 263 pp.

Easterling, D. R., B. Horton, P. D. Jones, T. C. Peterson, T. R. Karl, D. E. Parker, M. J. Salinger, V. Razuvayev, N. Plummer, P. Jamason, and C. K. Folland, 1997: Maximum and minimum temperature trends for the globe. *Science*, **277**, 364-367.

Folland, C.K., Rayner, N.A., Brown, S.J., Smith, T.M., Shen, S.S.P., Parker, D.E., Macadam, I., Jones, P.D., Jones, R.N., Nicholls, N. and Sexton, D.M.H., 2001a: Global temperature change and its uncertainties since 1861. *Geophysical Research Letters,* **28**, 2621-2624.

Folland, C.K., T.R. Karl, J.R. Christy, R.A. Clarke, G.V. Gruza, J. Jouzel, M.E. Mann, J. Oerlemans, M.J. Salinger and S-W. Wang, and 142 other authors, 2001b: Observed Climate Variability and Change. In: *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change* [Houghton, J. T., Y Ding, D. J. Griggs, M. Noguer, P. van der Linden, X. Dai, K. Maskell, and C. I. Johnson (eds.)]. Cambridge University Press, 99-181.

Folland, C.K., M. J. Salinger, N. Jiang, and N.A. Rayner, 2003: Trends and variations in South Pacific Island and ocean surface temperatures. *Journal of Climate*, **16**, 2859-2874.

Free, M., and J. K. Angell, 2002: Effect of volcanoes on the vertical temperature profile in radiosonde data. *Journal of Geophysical Research,* **107**, 4101, doi: 10.1029/2001JD001128.

Free, M., J. K. Angell, I. Durre, J. R. Lanzante, T. C. Peterson and D. J. Seidel, 2003: Using first differences to reduce inhomogeneity in radiosonde temperature datasets. *Journal of Climate*, **21**, 4171-4179.

Free, M., I. Durre, E. Aguilar, D. J. Seidel, T. C. Peterson, R. E. Eskridge, J. K. Luers, D. Parker, M. Gordon, J. R. Lanzante, S. A. Klein, J. R. Christy, S. Schroeder, B. J. Soden, and L. M. McMillin, 2002: CARDS workshop on adjusting radiosonde temperature data for climate monitoring: Meeting summary, *Bulletin of the American Meteorological Society*, **83**, 891-899.

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

Fu, Q., and C. M. Johanson, 2005: Satellite-Derived Vertical Dependence of Tropical Tropospheric Temperature Trends. *Geophysical Research Letters,* **32**, L10703, doi:10.1029/2004GL022266.

Fu, Q., C. M. Johanson, S. Warren, and D. Seidel, 2004: Contribution of stratospheric cooling to satellite-inferred tropospheric temperature trends. *Nature*, **429**, 55-58.

Gaffen, D, B. Santer, J. Boyle, J. Christy, N. Graham, and R. Ross, 2000: Multi-decadal changes in the vertical temperature structure of the tropical troposphere. *Science*, **287**, 1239-1241.

Gettelman, A., D. J. Seidel, M. C. Wheeler and R. J. Ross, 2002: Multidecadal trends in tropical convective available potential energy. *Journal of Geophysical Research*, **107**, 4606, doi:10.1029/2001JD001082.

Grody, N. C., K. Y. Vinnikov, M. D. Goldberg, J. T. Sullivan, and J. D. Tarpley, 2004: Calibration of Multi-Satellite Observations for Climatic Studies: Microwave Sounding Unit (MSU). *Journal of Geophysical Research*. **109**, D24104, doi:10.1029/2004JD005079.

Groisman, P.Y., R.W. Knight, T.R. Karl, D.R. Easterling, B. Sun, and J.H. Lawrimore, 2004: Contemporary changes of the hydrological cycle over the contiguous United States, trends derived from in situ observations. *Journal of Hydrometeorology*, **5** 64-85.

Hansen, J., R. Ruedy, M. Sato, and R. Reynolds 1996. Global surface air temperature in 1995: Return to pre-Pinatubo level. *Geophysical Research Letters,* **23**, 1665-1668.

Hansen, J., R. Ruedy, M. Sato, M. Imhoff, W. Lawrence, D. Easterling, T. Peterson, and T. Karl, 2001: A closer look at United States and global surface temperature change. *Journal of Geophysical Research,* **106**, 23,947-23,963.

Hegerl, G., and J. Wallace, 2002: Influence of patterns of climate variability on the difference between satellite and surface temperature trends. *Journal of Climate*, **15**, 2412-2428.

Henderson-Sellers, A., 1992: Continental cloudiness changes this century. *GeoJournal*, **27**, 255–262.

Houghton, J. T., Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, and C. A. Johnson, Eds., 2001: Climate Change 2001: The Scientific Basis. Cambridge University Press, 881 pp.

Hurrell, J.W., 1996: Influence of variations in extratropical wintertime teleconnections on Northern Hemisphere temperature. *Geophysical Research Letters,* **23**, 665-668.

Johanson, C. M., and Q. Fu, 2005: Robustness of tropospheric temperature trends from MSU Channels 2 and 4. *Journal of Climate*, submitted.

Jones, P.D., 1995: Land surface temperatures - is the network good enough? *Climatic Change* **31**, 545-558.

Jones, P. D., P. Ya. Groisman, M. Coughlan, N. Plummer, W-C. Wang, and T. R. Karl, 1990: Assessment of urbanization effects in time series of surface air temperature over land. *Nature*, **347**, 169-172.

Jones, P.D., Osborn, T.J. and Briffa, K.R., 1997: Estimating sampling errors in large-scale temperature averages. *Journal of Climate,* **10**, 2548-2568.

Jones, P.D., New, M., Parker, D.E., Martin, S. and Rigor, I.G., 1999: Surface air temperature and its changes over the past 150 years. *Reviews of Geophysics*, **37**, 173-199.

Jones, P.D., Osborn, T.J., Briffa, K.R., Folland, C.K., Horton, B., Alexander, L.V., Parker, D.E. and Rayner, N.A., 2001: Adjusting for sampling density in grid-box land and ocean surface temperature time series. *Journal of Geophysical Research* **106**, 3371-3380.

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**, 206-223.

Karl, T. R, Jones, P. D, Knight, R. W, Kukla, G., Plummer, N., Razuvayev, V., Gallo, K. P., Lindseay, J., Charlson, R. J, and T. C. Peterson, 1993: A new perspective on recent global warming: Asymmetric trends of daily maximum and minimum temperature. *Bulletin of the American Meteorological Society*, **74**, 1007-1024.

Karl, T. R., R. W. Knight, and B. Baker, 2000: The record breaking global temperatures of 1997 and 1998: Evidence for an increase in the rate of global warming? *Geophysical Research Letters,* **27**, 719-722.

Kistler, R., and Coauthors, 2001: The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bulletin of the American Meteorological Society*, **82**, 247-267.

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**, 224-240.

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**, 241-262.

Mears, C. A., M. C. Schabel, and F. J. Wentz, 2003: A Reanalysis of the MSU channel 2 tropospheric temperature record. *Journal of Climate*, **16**, 3650-3664.

Mears, C. A., and F. J. Wentz, 2005: The effect of diurnal correction on satellite-derived lower tropospheric temperature. *Science*, submitted.

Oleson, K. W., G. B. Bonan, S. Levis and M. Vertenstein, 2004: Effects of land use change on North American climate: impact of surface datasets and model biogeophysics. *Climate Dynamics*, **23**,117 - 132.

Oort, A. H. (1983): Global Atmospheric Circulation Statistics, 1958-1973. NOAA Prof. Pap 4. US Government Printing Office, Washington, D.C., 180 pp.

Parker D.E., Folland, C.K. and M. Jackson, 1995: Marine surface temperature: observed variations and data requirements. *Climatic Change*, 31, 559-600, and in: *Long-term climate Monitoring by the Global Climate Observing System*, Ed: T. Karl, pp429-470, Kluwer, Dordrecht.

Parker, D. E., P. D. Jones, C. K. Folland and A. Bevan, 1994: Interdecadal changes of surface temperature since the late nineteenth century. *Journal of Geophysical Research***,** **99**, 14,373-14,399.

Parker, D. E., M. Gordon, D. P. N. Cullum, D. M. H. Sexton, C. K. Folland, and N. Rayner, 1997: A new global gridded radiosonde temperature data base and recent temperature trends. *Geophysical Research Letters***,** **24**, 1499-1502.

Parker, D.E., Alexander, L.V. and Kennedy, J., 2004: Global and regional climate in 2003. *Weather*, **59**, 145-152.

Parker, D.E., 2004: Large-scale warming is not urban, *Nature*, **432**, 290-291, 10.1038/432290b.

Pawson, S., K. Labitzke, and S. Leder, 1998: Stepwise changes in stratospheric temperature. *Geophysical Research Letters*, **25**, 2157-2160.

Peterson, Thomas C., 2003: Assessment of urban versus rural in situ surface temperatures in the contiguous U.S.: No difference found. *Journal of Climate*, **18**, 2941-2959.

Peterson, Thomas C., 2005: Examination of potential biases in air temperature caused by poor station locations. *Bulletin of the American Meteorological Society*, submitted.

Peterson, T. C., D. R. Easterling, T. R. Karl, P. Ya. Groisman, N. Nicholls, N. Plummer, S. Torok, I. Auer, R. Boehm, D. Gullett, L. Vincent, R. Heino, H. Tuomenvirta, O. Mestre, T. Szentimre, J. Salinger, E. Førland, I. Hanssen-Bauer, H. Alexandersson, P. Jones, D. Parker, 1998: Homogeneity adjustments of in situ atmospheric climate data: A review. *International Journal of Climatology*, **18**, 1493-1517.

Peterson, T. C., T. R. Karl, P. F. Jamason, R. Knight, and D. R. Easterling, 1998b: The first difference method: Maximizing station density for the calculation of long-term global temperature change. *Journal of Geophysical Research***,** **103**, 25,967-25,974.

Peterson, T. C., K. P. Gallo, J. Lawrimore, T. W. Owen, A. Huang, and D. A. McKittrick, 1999: Global rural temperature trends. *Geophysical Research Letters*, **26**, 329-332.

Peterson, T.C., T.W. Owen, 2005. Urban heat island assessment: Metadata are important. *Journal of Climate*, in press.

Peterson, Thomas C. and Russell S. Vose, 1997: An overview of the Global Historical Climatology Network temperature data base. *Bulletin of the American Meteorological Society*, **78**, 2837-2849.

Prabhakara, C., R. Iacovazzi Jr, J.-M. Yoo, and G. Dalu., 2000: Global warming: estimation from satellite observations. *Geophysical Research Letters*, **27**, 3517-3520.

Randel, W. J. and F. Wu, 2005: Biases in stratospheric and tropospheric temperature trends derived from historical radiosonde data. *Journal of Climate*, accepted.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research*, **108**, 4407, doi:10.1029/2002JD002670.

Reynolds, R. W. and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. *Journal of Climate*, **7**, 929-948.

Santer, B. D., and Coauthors, 2004: Identification of anthropogenic climate change using a second-generation reanalysis. *Journal of Geophysical Research***,** **109**, D21104, doi:10.1029/2004JD005075.

Seidel, D. J., J. K, Angell, M. Free, J. Christy, R. Spencer, S. A. Klein, J. R. Lanzante, C. Mears, M. Schabel, F. Wentz, D. Parker, P. Thorne, and A. Sterin, 2004: Uncertainty in signals of large-scale climate variations in radiosonde and satellite upper-air temperature datasets. *Journal of Climate*, **17**, 2225-2240.

Seidel, D.J., and J.R. Lanzante, 2004: An assessment of three alternatives to linear trends for characterizing global atmospheric temperature changes. *Journal of Geophysical Research*, **109**, doi: 10.1029/2003JD004414.

Sherwood, S., J. R. Lanzante, and C. Meyer, 2005: Radiosonde daytime biases and late-20th Century warming. *Science*, **309**, 1556-1559.

Slutz, R. J., S. J. Lubker, J. D. Hiscox, S. D. Woodruff, R. L. Jenne, D. H. Joseph, P. M. Steurer, and J. D. Elms, 1985: *COADS: Comprehensive Ocean-Atmosphere Data Set. Release 1*, 262 pp. [Available from Climate Research Program, Environmental Research Laboratories, 325 Broadway, Boulder, CO 80303.]

Smith, T. M., R. W. Reynolds, R. E. Livezey and D. C. Stokes, 1996: Reconstruction of historical sea surface temperatures using empirical orthogonal functions. *Journal of Climate*, **9**, 1403-1420.

Smith, T.M., and R.W. Reynolds, 2003: Extended reconstruction of global sea surface temperatures based on COADS Data (1854-1997). *Journal of Climate*, **16**, 1495-1510.

Smith, T.M. and R. W. Reynolds, 2005: A global merged land and sea surface temperature reconstruction based on historical observations (1880-1997). *Journal of Climate*, **18**, 2021-2036.

Smith, T. M., T. C. Peterson, J. Lawrimore, and R. W. Reynolds, 2005: New surface temperature analyses for climate monitoring. *Geophysical Research Letters*, submitted.

Spencer, R.W. and J. R. Christy, 1990: Precise monitoring of global temperature trends from satellites. *Science,* **247**, 1558-1562.

Spencer, R.W. and J. R. Christy, 1992: Precision and radiosonde validation of satellite gridpoint temperature anomalies. Part I: MSU channel 2. *Journal of Climate*, **5**, 847-857.

Stendel, M., J. Christy, and L. Bengtsson, 2000: Assessing levels of uncertainty in recent temperature time series. *Climate Dynamics*, **16**, 587-601.

Stott, P. A., and S. F. B. Tett, 1998: Scale-dependent detection of climate change. *Journal of Climate*, **11**, 3282-3294.

Sun, B. and P.Ya. Groisman, 2000: Cloudiness variations over the former Soviet Union. *International Journal of Climatology*, **20**, 1097–1111.

Thorne, P.W., D. E. Parker, S. F. B. Tett, P. D. Jones, M. McCarthy, H. Coleman, and P. Brohan, 2005: Revisiting radiosonde upper-air temperatures from 1958 to 2002. *Journal of Geophysical Research,* **110**, D18105, doi:10.1029/2004JD00575.

Trenberth, K. E., 1990: Recent observed interdecadal climate changes in the Northern Hemisphere. *Bulletin of the American Meteorological Society,* **71**, 988-993.

Trenberth, K., and J. Hurrell, 1994: Decadal atmosphere-ocean variations in the Pacific. *Climate Dynamics*, **9**, 303-319.

Trenberth, K.E., Carron, J.M., Stepaniak, D.P. and S. Worley, 2002: Evolution of the El Nino-Southern Oscillation and global atmospheric surface temperatures. *Journal of Geophysical Research,* **107**, D8, 10.1029/2000JD000298.

Van den Dool, H. M., S. Saha, and A. Johansson, 2000: Empirical orthogonal teleconnections. *Journal of Climate*, **13**, 1421-1435.

Vinnikov, K. Y., and N. C. Grody, 2003: Global warming trend of mean tropospheric temperature observed by satellites. *Science*, **302**, 269-272.

Vinnikov, K. Y., A. Robock, N. C. Grody, and A. Basist, 2004: Analysis of diurnal and seasonal cycles and trends in climatic records with arbitrary observation times. *Geophysical Research Letters*, **31**, L06205, doi:10.1029/2003GL019196, 2004.

Vinnikov, K. Y., N. C. Grody, A. Robock, R. J. Stouffer, P. D. Jones, and M. D. Goldberg, 2005: Temperature Trends at the Surface and in the Troposphere. *Journal of Geophysical Research*, submitted.

Vose, R. S., C. N. Williams, T. C. Peterson, T. R. Karl, and D. R. Easterling, 2003: An evaluation of the time of observation bias adjustment in the U.S. historical climate network. *Geophysical Research Letters*, **30**, doi:10.1029/2003GL018111.

Vose, R.S., D.R. Easterling, and B. Gleason, 2005: Maximum and minimum temperature trends for the globe: An update through 2004. *Geophysical Research Letters,* accepted.

Wallis, T.W.R., 1998: A subset of core stations from the Comprehensive Aerological Reference Data Set (CARDS). *Journal of Climate*, **11**, 272-282.

Waple, A. M. and J. H. Lawrimore, eds, 2003: Climate of 2002. *Bulletin of the American Meteorological Society*, S1-S68.

Woodruff, S. D., H. F. Diaz, J. D. Elms, and S. J. Worley, 1998: COADS Release 2 data and metadata enhancements for improvements of marine surface flux fields. *Physics and Chemistry of the Earth,* **23**, 517-527.

Wuertz, D., R. S.Vose, T. C. Peterson and P. D. Jones, 2005: The GHCN-Jones comparison. *Geophysical Research Letters*, submitted.