A global comparison of surface and free-air temperatures at high elevations

N. C. Pepin[1]

Department of Geography, University of Portsmouth, Hants, UK

Dian J. Seidel

Air Resources Laboratory, NOAA, Silver Spring, Maryland, USA

[1] Surface and free-air temperature observations from the period 1948–2002 are compared for 1084 surface locations at high elevations (>500 m) on all continents. Mean monthly surface temperatures are obtained from two homogeneity adjusted data sets: Global Historical Climate Network (GHCN) and Climatic Research Unit (CRU). Free-air temperatures are interpolated both vertically and horizontally from the National Centers for Environmental Prediction/National Center for Atmospheric Research Reanalysis R1 $2.5°$ grids at given pressure levels. The compatibility of surface and free-air observations is assessed by examination of the interannual variability of both surface and free-air temperature anomalies and the surface/free-air temperature difference (ΔT). Correlations between monthly surface and free-air anomalies are high. The correlation is influenced by topography, valley bottom sites showing lower values, because of the influence of temporally sporadic boundary layer effects. The annual cycle of the derived surface/free-air temperature difference (ΔT) demonstrates physically realistic variability. Cluster analysis shows coherent ΔT regimes, which are spatially organized. Temporal trends in surface and free-air temperatures and ΔT are examined at each location for 1948–1998. Surface temperatures show stronger, more statistically robust and widespread warming than free-air temperatures. Thus ΔT is increasing significantly at the majority of sites (>70%). A sensitivity analysis of trend magnitudes shows some reliance on the time period used. ΔT trend variability is dominated by surface trend variability because free-air trends are weak, but it is possible that reanalysis trends are unrealistically small. Results are sensitive to topography, with mountaintop sites showing weaker ΔT increases than other sites (although still positive). There is no strong relationship between any trend magnitudes and elevation. Since ΔT change is dependent on location, it is clear that temperatures at mountain sites are changing in ways contrasting to free air.

1. Introduction

[2] Much recent research has involved the investigation of global temperature trends over the 20th century, in particular the concern over whether an anthropogenic signal can be observed in the surface, radiosonde and satellite observations. Analyses of variability and trends in these observations do not always agree [*National Research Council*, 2000]. At least part of the difference can be accounted for because of observational uncertainties in all three types of data, and part because the three types of observing system are not measuring the same temperatures,

either temporally or spatially. Satellite data yield a smoothed view of the vertical temperature profile, and radiosonde data are point measurements at constant pressure levels (not constant elevation or position), while surface data are anchored in space at a variety of fixed elevations. A new approach is taken in this study, minimizing such sampling differences by comparing surface and free-air temperatures from different data sets at fixed locations, mountain sites.

[3] This paper compares surface temperatures from 1084 high-elevation surface sites from the Global Historical Climate Network (GHCN) [*Peterson and Vose*, 1997] and Climatic Research Unit (CRU) [*Jones et al.*, 1999; *Jones and Moberg*, 2003] surface data sets with free-air equivalent temperatures interpolated from the National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) reanalysis data set [*Kalnay et al.*, 1996; *Kistler et al.*, 2001]. Free-air temperatures are defined

---

[1]Also at Air Resources Laboratory, NOAA, Silver Spring, Maryland, USA.

in this paper as temperatures in the free atmosphere that are not substantially influenced by surface boundary layer effects. At high-elevation sites the temperature recorded at the surface at screen level is often different from the free atmosphere because the mountain surface sets up its own atmosphere, especially under calm conditions [see *Barry*, 1992].

[4] It is not clear whether temperature trends at mountain sites more closely approximate trends in the free-atmosphere at the same level, or trends at surface sites at lower elevations [*Seidel and Free*, 2003]. It is expected that mean climatological differences between surface and equivalent free-air temperatures should decrease in magnitude with elevation because boundary layer effects are reduced. This does not mean that instantaneous differences will always decrease because the intense radiation at high elevations and increased transmissivity can create large surface to free-air temperature gradients when advection is weak. Any systematic long-term differences in surface and free-air temperature trends could be accompanied by attendant changes in the mean energy balance of high-elevation areas. They would also be of importance because of the following:

[5] 1. Most GCMs (general circulation models) and assimilation schemes used to develop scenarios of future climate represent surface mountain conditions poorly, because of poor surface representation (lack of sufficient spatial resolution) and problems in relating free-air response to that at the surface. The atmosphere above mountains in GCMs can be made to respond to radiative forcing in contrasting ways to lowland areas [*Giorgi et al.*, 1997] through complex feedback processes, but model surface temperatures can still be very different from those observed. Comparison of surface and free-air observations is therefore helpful.

[6] 2. It could help explain the apparent discrepancy between rapid glacier retreat and lack of strong free-air warming in the mid troposphere. It may be that the surface is warming at a faster rate than the free-air in many mountain locations, but it could also be that factors other than temperature change are more influential in causing glacier retreat [*Kaser et al.*, 2004].

2. Past Studies

[7] Mean surface air temperature data show strong warming trends, especially over the last 2 decades [*Folland et al.*, 2001; *Jones et al.*, 1999; *Jones and Moberg*, 2003; *Peterson et al.*, 1998]. Other studies have suggested that surface temperature trends are increasing more rapidly at higher elevations [*Beniston et al.*, 1997; *Diaz and Bradley*, 1997], but this is not universally true [*Vuille and Bradley*, 2000; *Pepin*, 2000]. Many glaciers have retreated during this time period, especially in the tropics [*Thompson*, 2000], but it is not clear whether this is an indication of increased temperature change at high elevations since glaciers can respond in a complex manner to climate.

[8] Free-air trends as measured by radiosonde data have been widely reported [*Diaz and Graham*, 1996; *Gaffen et al.*, 2000]. Over large areas, significantly different trends have been identified over the last 2 decades at the surface and at higher levels in the troposphere, especially in the tropics. During 1979–1997 the mid troposphere cooled slightly, while the surface temperatures increased [*Gaffen et al.*, 2000], meaning an increase in lower tropospheric lapse rates. This discrepancy disappears when a longer time period is considered (1960–1997). More recent analyses, which have taken into account some of the inhomogeneities in the radiosonde record [*Lanzante et al.*, 2003a, 2003b], again show that the lapse rate trend is extremely sensitive to the time period chosen, and over a long time period (1959–1997) is dwarfed by a downward step-like change in the mid-1970s.

[9] A further study [*Diaz et al.*, 2003] examined free-air freezing level changes in selected high-elevation regions over 1956–2000 and for various subperiods using reanalysis data. Rising heights were identified in most areas. Most trend figures quoted were regional aggregates.

[10] Another source of free-air temperature data is that from satellite observations [*Christy et al.*, 2000; *Mears et al.*, 2003]. Differences in trends among analyses arise because of different corrections made by different research groups to account for factors such as radiative errors, orbital decay and overlapping satellites. It is difficult to compare such measurements with radiosonde data, because of the need to develop a vertical weighting function to facilitate comparison of satellite channels and radiosonde levels [*Lanzante et al.*, 2003b]. Lapse rate change cannot easily be identified using satellite data.

[11] A few studies have compared long-term records of surface and free-air temperatures (at the same elevation), at least locally. *Pepin and Losleben* [2002] show for the Colorado Front Range that divergent trends (surface cooling/free-air warming) result in the surface becoming an increasing heat sink relative to the free air. Possible explanations include increased atmospheric instability leading to higher snowfall selectively at high elevations, depressing surface temperatures [*Barry*, 1990]. Preliminary analysis in other mountain locations [*Pepin and Losleben*, 2001] fails to replicate this pattern, illustrating the localised nature of such changes.

[12] A more extensive study [*Seidel and Free*, 2003] examines the contrast in surface and free-air equivalent temperatures using 26 pairs of radiosonde stations from the Comprehensive Aerological Reference Data Set (CARDS). Comparing trends for pairs of sites (high-elevation surface versus free air above adjacent lowland site) showed contrasting patterns, with a tendency toward increased warming rates at high-elevation surface sites relative to the paired lowland site, more commonly in the tropics. The results from the Seidel and Free study have the advantage of being based on one high-quality data set, reducing data compatibility problems. However, it is debatable as to how representative the changes identified are of high elevations and mountain summits. Many "high-elevation" radiosonde sites are at relatively low elevations (e.g., Mexico City and Denver) and/or in suburban airport locations in valley bottoms subject to strong boundary layer effects, rather than in true mountain locations.

3. Data and Method

[13] The surface GHCN and CRU data sets were chosen for their spatial and temporal coverage, and for their mature consideration of homogeneity issues [*Peterson et*
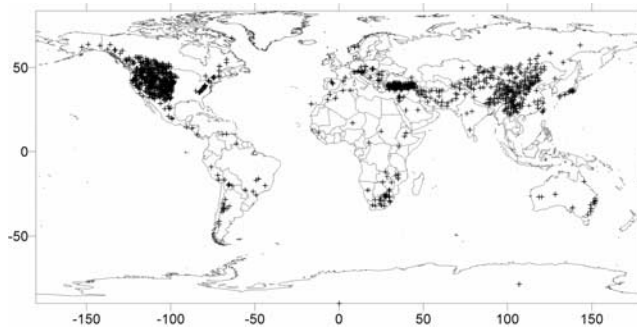
Figure 1. Distribution of 1084 Global Historical Climate Network/Climatic Research Unit (GHCN/CRU) surface sites used in this analysis.

al., 1998]. Adjustments have been made for changes in instrumentation, observing practice, site, land use around the site, and other issues. The interested reader is referred to *Peterson and Vose* [1997] and *Jones and Moberg* [2003] for details of these procedures. The homogeneity adjustment method is different for the U.S. stations which were originally from the U.S. HCN (Historical Climate Network). The GHCN version 2 data set has comprehensive global coverage for 1948–1997, including the high-elevation areas of Asia and North America, but data from a few sites finish as early as 1990. The CRU data set on the other hand has been updated to 2002 at some sites. All available mean monthly temperatures between 1948 and 2002 were used in this study, although the main trend analyses were restricted to 1948–1998. Using mean monthly temperatures maximizes the number of stations because although mean monthly maximum and minimum temperatures exist for some stations, data availability is much more limited.

[14] The NCEP/NCAR reanalysis R1 [*Kistler et al.*, 2001] (http://www.cdc.noaa.gov) was chosen because of its comprehensive coverage, with little missing data. It is a combination of free-air data, including radiosonde and satellite data, with model output. It does not include surface observations. Free-air temperatures are recorded on a 2.5° latitude/longitude grid, four times daily (0, 6, 12 and 18 Z) at pressure levels 500, 600, 700, 850, 925 and 1000 mbar. R1 was obtained from the Climate Diagnostics Center in Boulder, Colorado. The reanalysis starts in 1948. Surface (skin) temperatures, derived from the model, and used by *Kalnay and Cai* [2003] in an analysis of urban effects, are

not used here, because of uncertainties in interpretation [*Vose et al.*, 2004; *Trenberth*, 2004] and the documented snow cover problem in R1 which also influenced near surface temperature (2 m). For details of this and other known errors in R1 (which are not relevant to the present analysis) see *Kanamitsu et al.* [2002].

[15] There are homogeneity concerns in the reanalysis, particularly because of a time change of observations in 1958 and the introduction of satellite data in 1979. The potential effects of this are examined. The reanalysis temperature observations over land areas (predominately Northern Hemisphere) investigated in this study are regarded as of relatively high quality and the model output is strongly biased toward the raw observations. Nevertheless, there must remain some healthy skepticism as to what the NCEP/NCAR reanalysis actually represents, because of the inclusion of a wide range of data sources and a modeling component. Overall, the advantages of spatial and temporal comprehensiveness, especially in comparison with homogenized radiosonde networks, were thought to outweigh these limitations.

[16] Since the focus of this study is "mountain" sites, surface sites over 500 m above sea level in both GHCN and CRU data sets were considered for inclusion. This liberal threshold allows the sensitivity of results to elevation and topography to be examined. More than 300 months of data in the 360 month period (1961–1990) were required, leading to 1456 potential stations (923 GHCN and 533 CRU). Temperatures were converted to anomalies with respect to 1961–1990. There is considerable overlap in data set coverage, with many stations in both. Unfortunately, the official WMO station numbers did not always agree despite similar location and vice versa (similar station numbers with different locations), so merging the two data sets was not trivial.

[17] For stations that were thought to be the same in both data sets, individual monthly temperature anomalies for all years (up to 360 values) were correlated. If the correlation was below 0.95 both time series were dropped. In a few cases this was because the stations were indeed different locations (or it is suspected that this is the case), but in most cases, particular homogeneity adjustments had been made in one data set but not in the other, leading to distinct banding of anomalies. Such adjustments would have an influence on trends. Instead of attempting to decide which data set was "correct" such controversial stations were dropped. For pairs of stations for which the anomaly correlation was greater than 0.95, the station data set with the longest period

Table 1. Number of Stations in Each Continent and Summary Characteristics

| Continent | Number of Stations | Mean | | | Stations | | | |
		Elevation, m	Latitude, deg	Longitude, deg	Elevation >2000 m	Elevation >3000 m	Polar[a]	Tropical[a]
North America	552	1107.7	43.1	−108.2	34	0	10	10
South America	33	1299.8	−19.9	−64.3	6	4	0	22
Europe	162	1206.5	39.8	31.8	13	1	3	4
Africa	41	1203.2	−18.9	26.3	1	0	0	36
Asia	280	1645.3	37.1	100.5	72	38	1	62
Australia	14	674.1	−29.4	143.3	0	0	0	8
Antarctica	2	3136.5	−84.3	53.5	2	1	2	0

[a]Polar is defined as above 60° latitude, and tropical is defined as below 30° latitude.
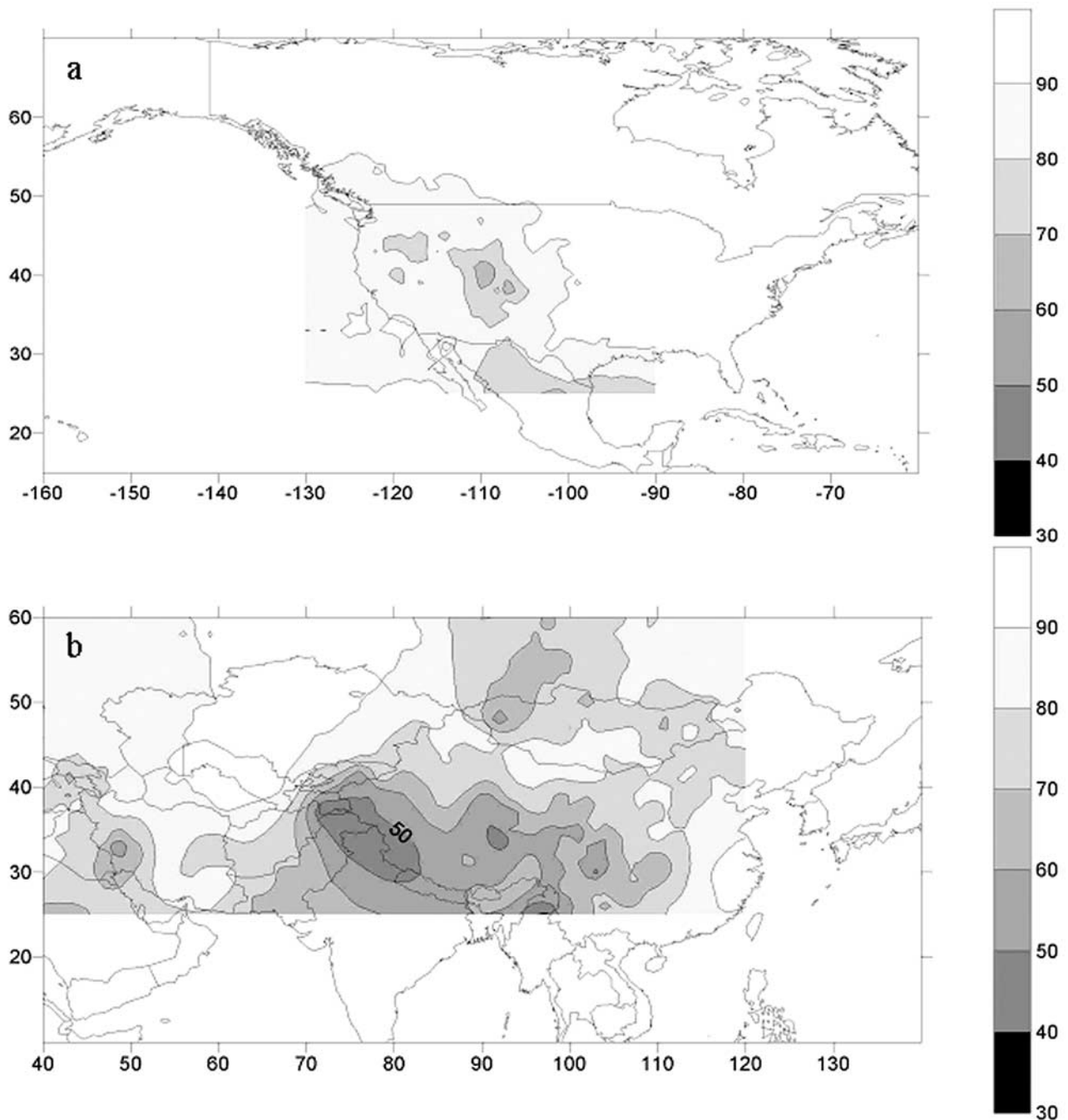
Figure 2. Monthly surface temperature/free-air temperature anomaly correlation fields for (a) western North America and (b) Asia. Correlations are interpolated using kriging over regions with numerous stations. Low-elevation areas and oceanic areas with sparse data coverage are left blank.

of record was retained. No merging of station records to obtain longer records was performed, because it would involve potentially controversial extra adjustments. In many cases the CRU data were chosen because of the more up-to-date status of the data set.

[18] A map of the final stations (695 GHCN and 389 CRU) and their continental and elevational distribution is outlined in Figure 1 and Table 1. Stations are concentrated in North America and Asia. The mean elevation in each continent is remarkably similar (~1200 m) apart from in Asia where the Tibetan plateau is well represented, and in Antarctica (only two stations). Tropical stations occur in most continents and are well spread around the globe, even though their absolute numbers are relatively small. There is a positively skewed distribution of station elevations with only 17 sites above 4000 m. Although mountain areas are well represented, the area covered by these high-elevation sites in global terms is only about 20% of the Earth's surface
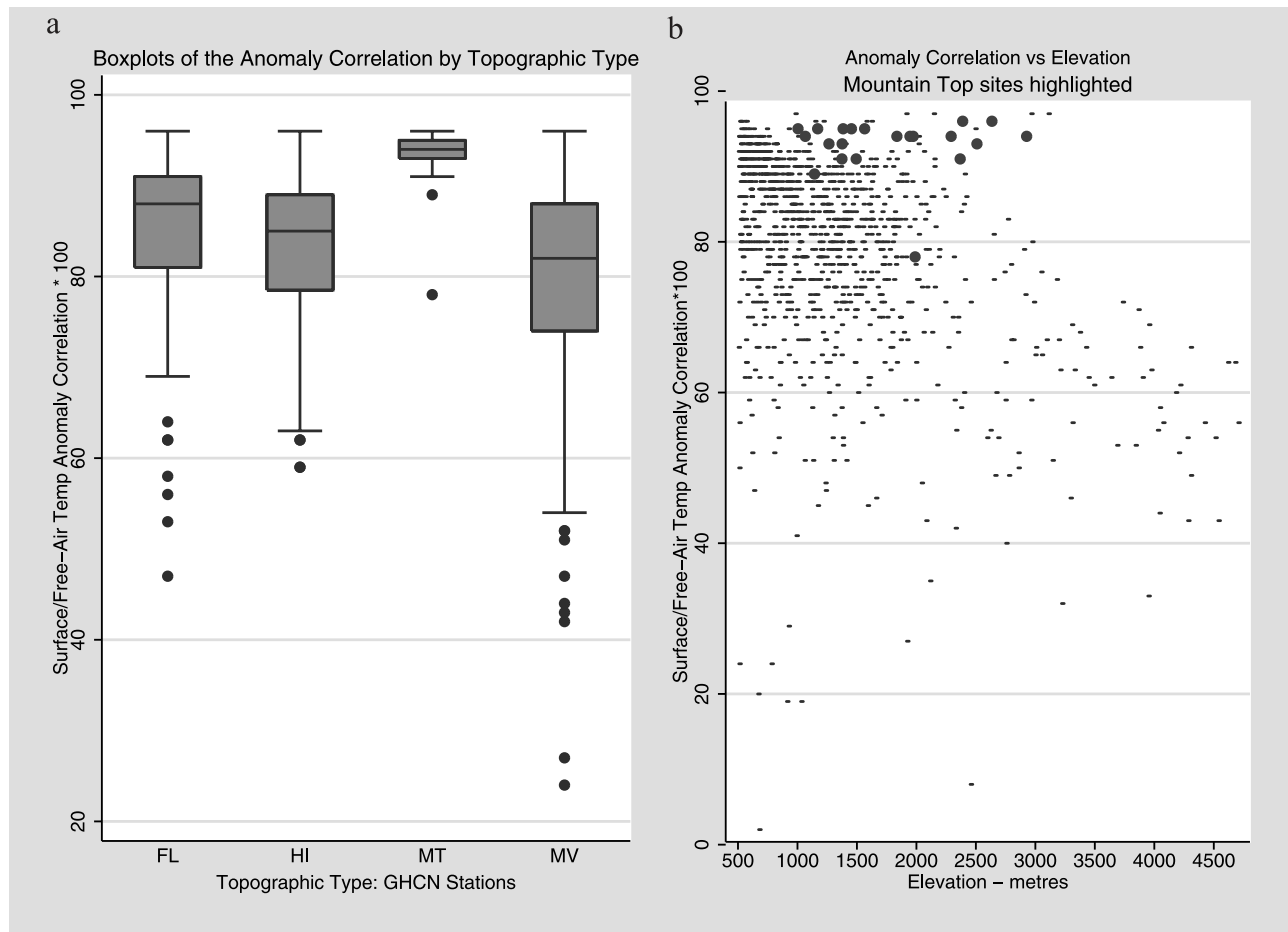
Figure 3. (a) Analysis of variance of the surface/free-air temperature anomaly correlation by topographic class for GHCN stations. Correlations are multiplied by 100. Topographic types are FL, flat; HI, hilly; MT, mountaintop; and MV, mountain valley. (b) Anomaly correlation (multiplied by 100) versus station elevation for all sites (GHCN and CRU). Mountaintop sites are highlighted as large circles.

area, so a comparison of these results with global average studies could be misleading.

[19] To make free-air temperatures from the reanalysis comparable in position with the surface data, a monthly mean free-air equivalent temperature (Ta) was created for each surface site by interpolating between reanalysis pressure levels using the individual mean monthly temperature and pressure height fields. The vertical interpolation was done first for the four nearest grid points based on a linear lapse rate between the two nearest pressure levels. This ignores the possibility of inversions in the free air. Since surface sites are invariably not at a 2.5° intersection of the NCEP/NCAR grid, horizontal interpolation was done after the vertical interpolation. On the scale of 2.5° grid boxes most high-elevation surface sites are on higher land than the surface elevation at the four grid points used for the interpolation. Thus the interpolation usually uses free-air values above the surface rather than extrapolated subsurface temperatures, but this is not always the case. Sensitivity of results to this is examined.

[20] To accept the idea of reanalysis interpolation one must accept that the free atmosphere is less complex than the Earth's surface and that the free-air temperature field is

therefore less variable (smaller rates of change) and more regular. The opposite approach of interpolating the surface data to a 2.5° grid would be unsuitable in mountainous areas, where local-scale variability is large.

[21] A new variable was created representing the difference between the free-air equivalent temperature (Ta) and surface temperatures (Ts), to be referred to as the free-air/surface temperature difference or ΔT (Ts − Ta). This was obtained by subtracting the instantaneous mean monthly free-air equivalent temperature from the surface temperature. If ΔT is positive, this represents a warm surface in comparison with the free air and vice versa. Time series were created of monthly anomalies for Ts, Ta and ΔT.

4. Data Evaluation

[22] For each location, the correlation between monthly surface and free-air temperature anomalies was calculated. Selected correlation maps are shown in Figure 2 for the United States and Asia (the two continents with the largest number of stations). Globally, at 1054 out of 1084 sites the correlation is above 0.5, and the mean value is 0.801. Nearly 60% of sites have correlations exceeding 0.8. Of
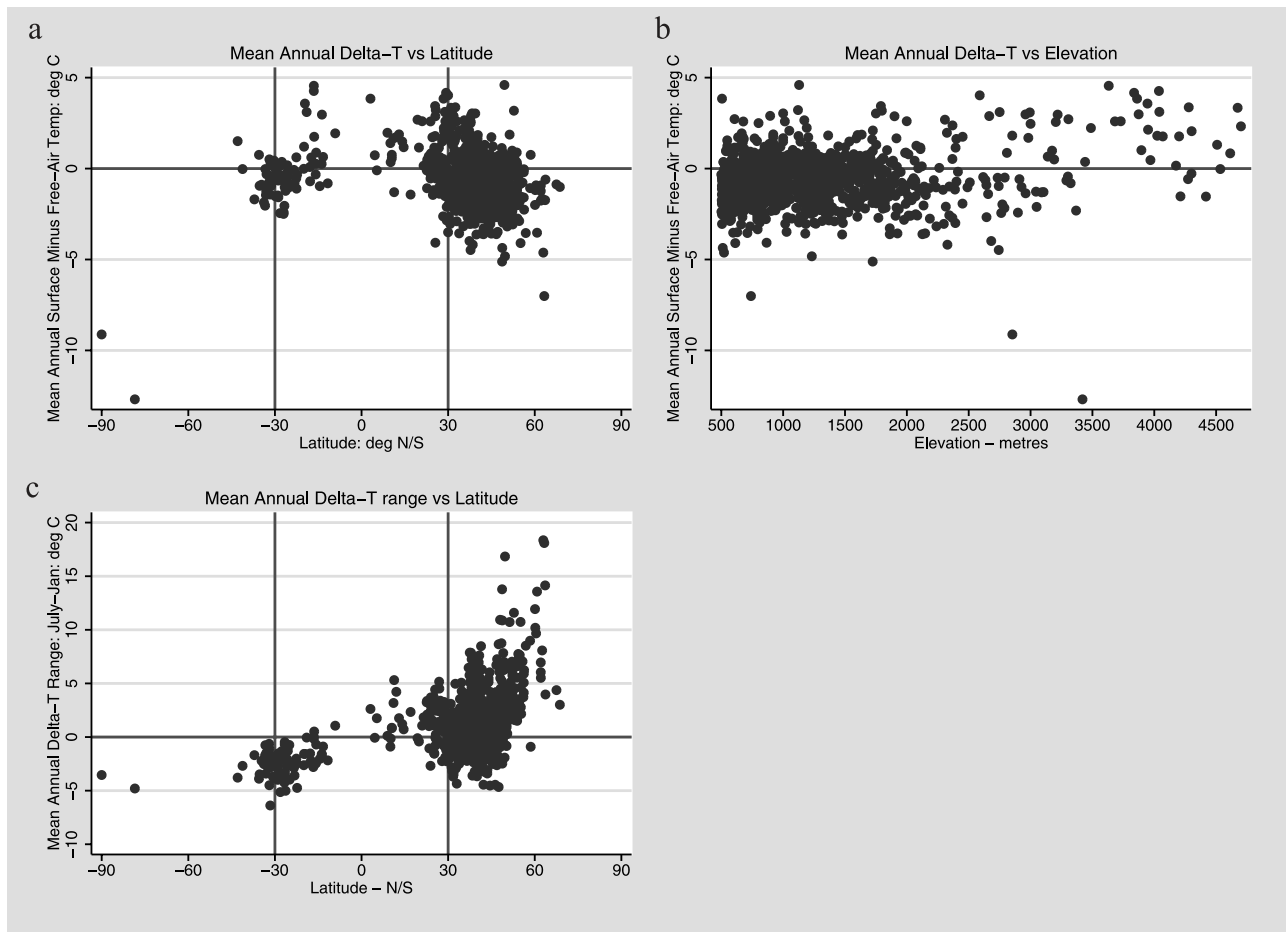
Figure 4. Mean annual surface/free-air temperature difference (ΔT) (°C) versus (a) latitude and (b) elevation and (c) July minus January ΔT versus latitude.

the 30 sites with poor correlations (<0.5), most are in the tropics in South America or Africa. There is an improvement of anomaly correlations at higher latitudes.

[23] The anomaly correlation offers an indication of the quality of both the reanalysis and the surface data. A good correlation increases our confidence in both types of data (since they are independent), but one must be careful in assuming that low correlations automatically mean data inadequacy, since surface stations will exhibit boundary layer effects which are not included in the reanalysis.

[24] The GHCN stations had been classified by topography as part of the preparation of the GHCN data set [*Peterson and Vose*, 1997]. Classification was subjective, based on examination of 1:1,000,000 Operational Navigation Charts and included four categories (FL, flat; HI, hilly; MT, mountain summit; and MV, mountain valley). An analysis of variance for the anomaly correlation by topographic class yields highly significant results (Figure 3a, $p < 0.001$) showing that mountain summits (MT) show much higher anomaly correlations than the other three categories, especially mountain valley (MV) sites. Most of the low correlations are from deeply incised mountain valley locations where sporadic boundary layer effects (especially inversions) would decouple the surface temperature variation from that of the free air. A plot of

anomaly correlation against elevation (Figure 3b) shows a weak decrease at higher elevations (although the worst values are at low elevations). It is tempting to suggest that this is because the reanalysis performs badly in mountain areas, but paradoxically most of the surface sites at the highest elevations are mountain valley sites (on the Tibetan plateau). It is most likely the incised topography and complex relief and not elevation per se, which causes the lower correlation. There is no decrease in correlation at higher elevations for the mountain summit sites (large circles on Figure 3b), at least up to around 3000 m (as high as the summit sites go). Surface temperatures at mountain summit sites have more in common with free-air temperatures as measured by the NCEP reanalysis, as do the other surface sites.

[25] GHCN stations were also classified as rural, suburban or urban according to population estimates [see *Peterson and Vose*, 1997]. It is important to note that sites classified as urban may not have been so at the start of the record, although rural sites are unlikely to have changed. An analysis of variance of the anomaly correlation by the degree of urbanization (not shown) also yields significant results ($p < 0.001$), with urban areas showing lower correlations than small town areas, in turn having lower correlations than rural areas. This suggests an additional (temporally variable)
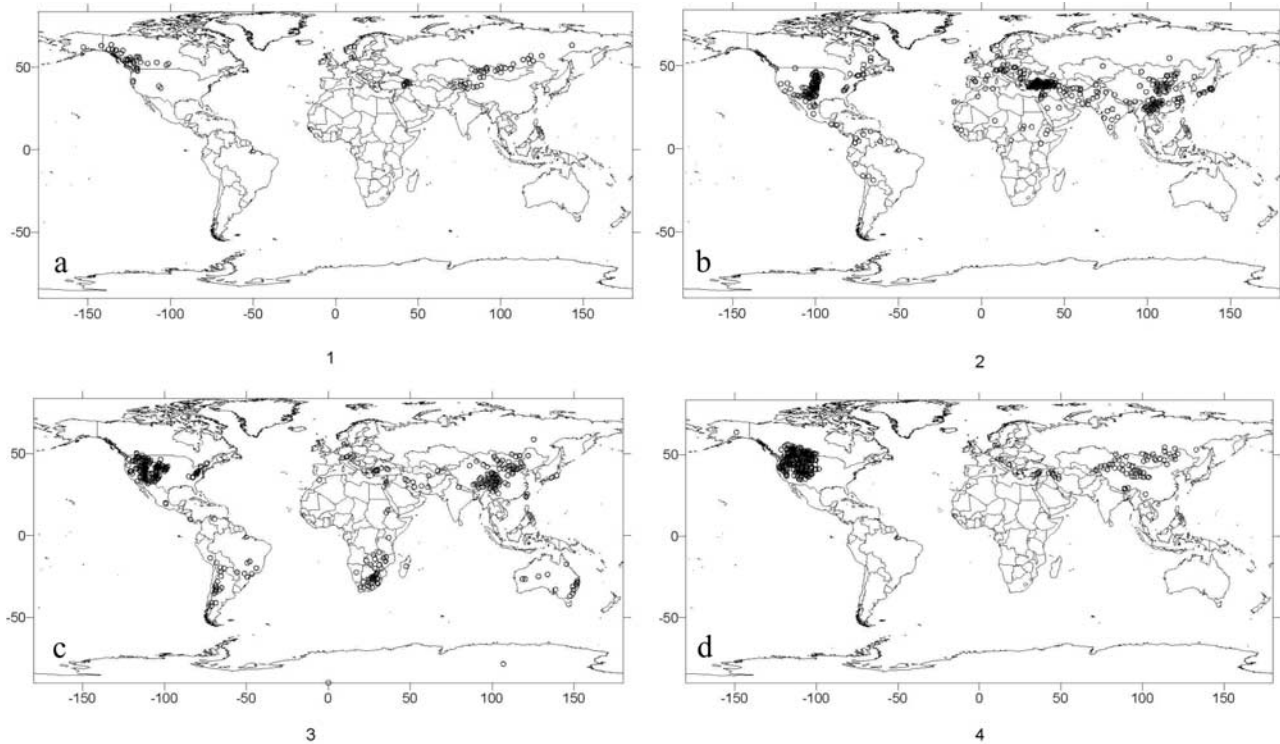
Figure 5. Geographic location of the four types of annual ΔT regime as identified by the K means cluster analysis on the 12 monthly ΔT anomalies: (a) type 1, (b) type 2, (c) type 3, and (d) type 4. For a discussion of the characteristics of the four regime types see Table 2 and text.

climate influence in urban areas. However, since urban areas tend to be in valleys (i.e., the topographic and urban/rural classifications are not independent) one must be careful in this interpretation.

[26] The mean values of ΔT at each site were also examined to cast light on the differences between the two data sets. Overall patterns of the mean annual difference are intuitive (Figure 4a) with more highly positive values (Ts > Ta) in the tropics where the net radiation balance is often positive and a decrease with increasing latitude in both hemispheres, especially the Northern Hemisphere. The lowest values ($<-6°C$) are recorded in Antarctica and at Ojmjakan in Siberia (well known for its intense surface inversions in winter). Overall there is a weak increase in ΔT with elevation (Figure 4b), which occurs in many continents (including Asia (r = 0.42, p < 0.01), South America (r = 0.85, p < 0.01) and Australia (r = 0.50, p < 0.01)) but not others (notably Europe). It is suggested that this increase

Table 2. Different Annual ΔT Regime Types as Identified by Cluster Analysis[a]

Regime Type[b]	Month With Most Positive Mean ΔT	Month With Most Negative Mean ΔT	Number of Stations	Common Location(s)
1	March (+3.4),	Dec. (−5.6)	77	higher northern latitudes
2	July (+0.9)	Jan. (−1.0)	349	midlatitudes (lower elevation sites)
3	March (+1.0)	Aug. (−1.1)	358	Southern Hemisphere, high elevation
4	April (+1.8), Oct. (+0.2)	Jan. (−2.0)	300	Northern Hemisphere Continents

Regime Type[b]	Percentage of Stations Within Elevation Bands[c]			
	<1000 m	1000−2000 m	2000−3000 m	>3000 m
1	67.6	26.1	5.4	0.9
2	49.0	41.0	6.9	3.1
3	36.0	46.0	9.3	8.7
4	49.1	42.4	6.1	2.4

[a]When there are 2 months in the year with strongly positive or negative ΔT values in comparison with preceding and following months, both are listed. The values in parentheses give the mean ΔT values (°C) in the highest and lowest months.
[b]Type 1 represents extreme annual signal, with strong winter surface cooling below the free air, and a strong spring peak (surface warms more rapidly in spring in response to increased solar input). Type 2 represents weaker annual cycle, with maximum in summer and minimum in midwinter. Type 3 represents Southern Hemisphere type with minimum during July/August. Type 4 represents subdued version of type 1, with maximum in spring and minimum in winter. There is often a secondary maximum in autumn.
[c]For the percentage of stations within each elevation band, rows add up to 100.

Table 3. Two-Way Tabulation Between Topographic Type and ΔT Regime Type[a]

Regime Type	Topographic Type				
	FL	HI	MT	MV	Total
1	10	4	0	22	36
2	72	77	17	95	261
3	31	59	3	114	207
4	55	72	1	73	201
Total	168	212	21	304	705

[a]Topographic types are defined as FL, flat; HI, hilly; MT, mountaintop; and MV, mountain valley. A chi-square analysis shows significant differences in regime types between topographic classes (p < 0.001).

could result from the increased radiative input at higher elevations if skies are clear, as ΔT should show strong relationships with the local radiation balance. Instantaneous values of ΔT on mountains are more positive in daytime and in summer at most locations [*Samson*, 1965; *Richner and Phillips*, 1984]. Europe, which is quite a cloudy continent, does not show such an increase with elevation.

[27] Mean ΔT was also calculated for each month. The difference between July and January values (July minus January) tends to be negative in the Southern Hemisphere and positive in the northern (increasing with latitude), as would be expected (Figure 4c). Thus the two data sets demonstrate the expected seasonal contrast in surface heating, surface sites being relatively warm in comparison with the reanalysis in summer and cool in winter, especially in continental areas. A K means cluster analysis [*Everitt et al.*, 2001] was attempted on the 12 monthly ΔT anomalies, to classify stations with similar monthly ΔT regimes. The number of classes was set to 4. Maps of the distribution of the four regimes show strong spatial clustering (Figure 5). There are four main types, their main features being listed in Table 2. Types 1 and 4 are restricted to the middle- and high-latitude Northern Hemisphere and are dominated by strong seasonal fluctuations, type 4 having a double peak in spring and fall. Types 2 and 3 are more subdued, with type 3 being the Southern Hemisphere regime. A two-way tabulation of regime type versus topographic type (Table 3) for GHCN stations shows significant interdependence. Mountaintop sites nearly always show regime type 2 (a small summer/winter contrast) and the highest-elevation sites (usually mountain valleys) often show regime type 3.

[28] All the above analyses increase confidence in both data sets since the differences between them (ΔT) and their temporal coherence (as measured by the anomaly correlation) follow logical patterns. Section 5 examines temporal trends in the surface data, NCEP reanalysis and the derived ΔT values.

5. Temporal Trends

[29] Statistically derived trends are dependent on the period of record used, and the methodologies used in order to derive them. The trend in a difference series (e.g., ΔT) is not the same as the difference between two trends calculated for the individual series (e.g., Ts and Ta). Trends here are derived using monthly anomalies with respect to 1961–1990, based on least squares. Mean/median trends and confidence intervals are calculated based on all trends,

whether significant or not. Significance of trends is assessed by using an adjusted sample size, standard error and degrees of freedom to take into account the temporal autocorrelation in temperatures [*Santer et al.*, 2000]. This is a stricter test than standard p values. In the following discussion numbers of significant trends are quoted at the 5% level. Trends are reported for the period 1948–1998.

5.1. Surface Temperatures (Ts)

[30] Maps of Ts trends for 1948–1998 are shown in Figure 6 and aggregate trends are summarized by continent (using all sites) with 95% confidence intervals in Table 4 (this also shows Ta and ΔT trends). Because of space restrictions Figure 6 only shows whether trends are positive, negative or insignificant (rather than their magnitude); 444 out of the 493 sites with significant trends show warming. On a continent-wide basis the mean trend is significantly greater than zero in all cases, except Antarctica (cooling) and Europe. Overall trend magnitudes at individual sites range from $-1.1°$ to $1.0°$C/decade, but the median trend is only $+0.13°$C/decade. There is some spatial clustering in trend sign with areas of warming being concentrated in Alaska and Canada, parts of Brazil, South Africa, and northern and western parts of Asia. Cooling is concentrated in the central United States, Iran, and central parts of China. There is no significant relationship between trend magnitude and elevation for the significant trends (Figure 7), despite the findings of some other studies [*Diaz and Bradley*, 1997; *Beniston et al.*, 1997]. South America shows a decrease in trend magnitude with elevation. Trend variability decreases at the highest elevations, but there are fewer stations.

[31] The influence of the degree of urbanization and latitude on trends was examined through individual analyses of variance. There is no significant difference in surface temperature trends between rural, suburban and urban sites [see *Jones et al.*, 1990; *Peterson et al.*, 1999]. Although there has long been an assertion that urbanization has contributed to an increased rate of surface warming in urban areas [*Cayan and Douglas*, 1984; *Kalnay and Cai*, 2003], this study fails to substantiate this [see also *Peterson*, 2003]. Extreme caution must be advised here since the subset of GHCN sites chosen in this study was not designed to sample the urban effect and the sample of large urban areas is poor. Topographic type also has no significant influence on mean trend magnitudes, although the most extreme cooling and warming rates are shown at mountain valley sites. To examine the influence of latitude, sites were classified into one of six $30°$ wide latitudinal bands ($>60°$N, $60°–30°$N, $30°–0°$N, $0°–30°$S, $30°–60°$S, $>60°$S). Mean warming (as measured by mean trend magnitude) is significantly higher at high latitudes of the Northern Hemisphere, in agreement with past global analyses [*Jones and Moberg*, 2003].

5.2. Free-Air Temperatures (Ta)

[32] Trends in free-air equivalent temperatures were calculated from the reanalysis (1948–1998). The trend maps (Figure 8) show that the spatial variation in trend magnitude is much smoother than for the surface, which is unsurprising since the reanalysis is one coherent data set rather than hundreds of independent stations. Out of 326 sites with
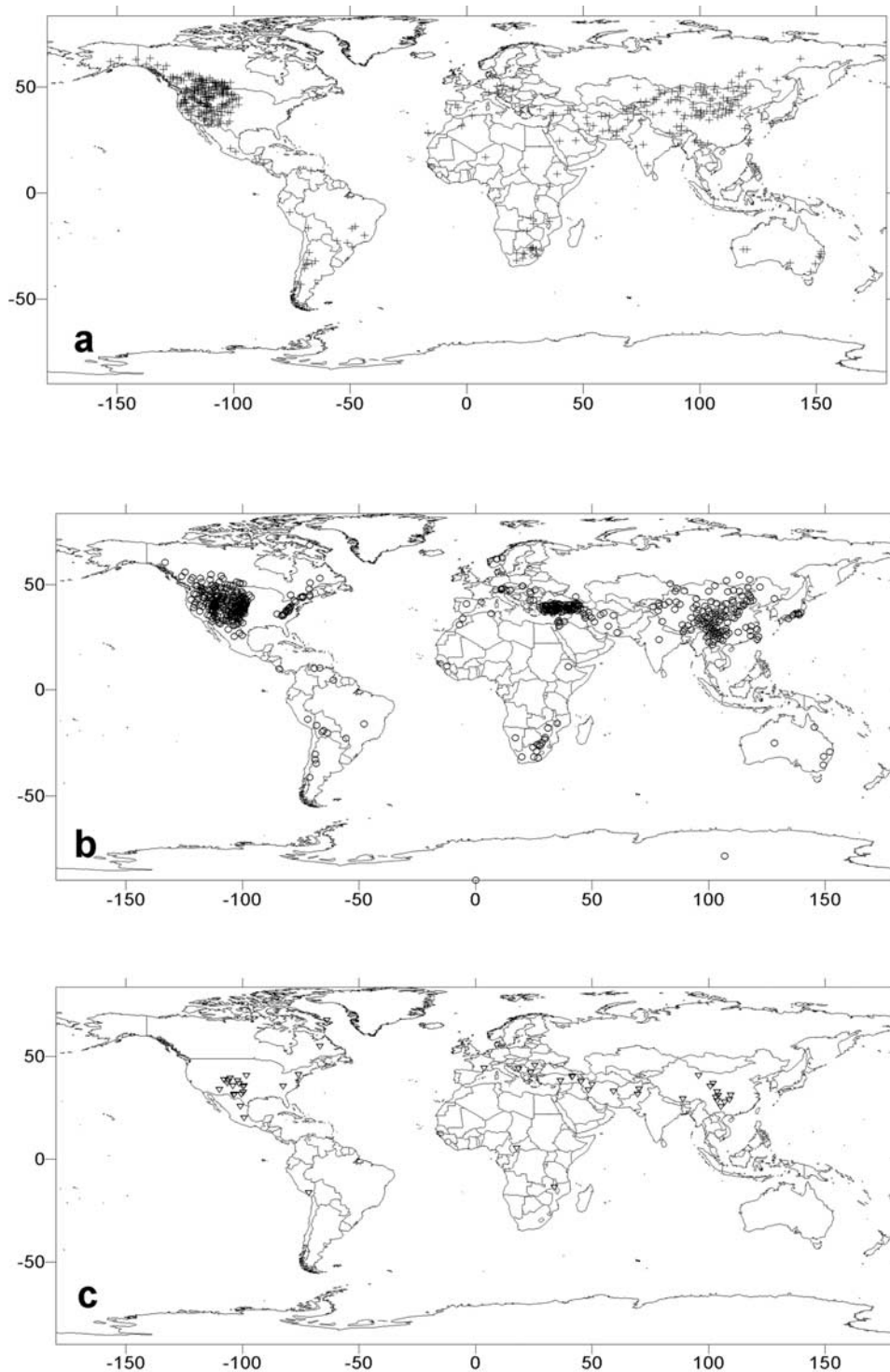
Figure 6. Maps of temporal trend in mean monthly surface temperature anomalies (1948–1998) showing stations with (a) significant warming, (b) no trend, and (c) significant cooling. Significance (p < 0.05) is assessed taking temporal autocorrelation into account.

significant trends, only 68 show warming. Significant rates of warming range from 0.08 to 0.48°C/decade. Significant cooling rates range from −0.40 to −0.10°C/decade. The median value is −0.05°C/decade. Thus the rates of change are much slower than for the surface temperatures and much

fewer sites show significant change. Areas of significant free-air cooling include the central United States, eastern Turkey, parts of South Africa, and China (concentrated in Sichuan and western Xinjiang provinces). Again, there is no significant increase in trend magnitude with elevation

Table 4. Mean Surface, Free-Air, and ΔT Trends (1948–1998) for Each Continent Based on All Stations[a]

Continent	Number of Sites	Surface Trend	Free-Air Trend	ΔT Trend
North America	552	0.14 ± 0.01[b]	−0.08 ± 0.009[b]	0.24 ± 0.01[b]
South America	33	0.13 ± 0.05[b]	0.24 ± 0.04[b]	−0.07 ± 0.07[c]
Europe	162	0.03 ± 0.04	−0.06 ± 0.04[b]	0.02 ± 0.05
Africa	41	0.14 ± 0.04[b]	−0.03 ± 0.05	0.08 ± 0.07[c]
Asia	280	0.15 ± 0.03[b]	0.006 ± 0.02	0.16 ± 0.03[b]
Australia	14	0.16 ± 0.06[b]	−0.005 ± 0.04	0.17 ± 0.09[b]
Antarctica	2	−0.07 ± 0.21	−0.08 ± 0.33	−0.07 ± 0.53
Total	1084	0.13 ± 0.01[b]	−0.04 ± 0.008[b]	0.17 ± 0.01[b]

[a]All trends are expressed in °C/decade ±1.96 standard errors (95% confidence interval). Not all the stations in each continent show significant trends.
[b]Trend is significantly different from zero (1%).
[c]Trend is significantly different from zero (5%).

(Figure 9), although the vast majority of trends at very high elevations are weakly positive. There is more variability in trend magnitude at lower elevations. South America appears anomalous with a concentration of free-air warming trends, although this agrees with the findings of *Diaz et al.* [2003], who analyzed free-air freezing level trends in the American Cordillera.

5.3. Surface/Free-Air Temperature Difference: ΔT

[33] Monthly ΔT anomalies were calculated (from raw ΔT values) and examined for trends. A different result would be obtained if the difference between the raw surface and free-air anomalies was calculated (not done).

[34] Maps of ΔT trends (Figures 10a–10c) show significant increases at 706 stations, with decreases at 94 sites. Significant increases range from 0.06 to 1.00°C/decade, and decreases range from −1.3 to −0.04°C/decade. The median value is +0.19°C/decade. Decreases in ΔT are restricted to parts of the southern and eastern United States, the Andes, western Turkey, Iran, the coast of South Africa, and small areas in China (notably the area northwest of Beijing). At about 70% of all locations the surface is warming at a rate faster than the free-air meaning a systematic increase in ΔT. There is no relationship between trend magnitude and elevation (Figure 11).

[35] The range of ΔT trend magnitudes is very similar to that in the surface trends. Much of the variability in ΔT trends is thus accounted for by variation in the surface temperature trends (correlation between the two is 0.696). Free-air trends are much weaker and thus contribute relatively little to the ΔT trend variation. In some ways this is reassuring since confidence is much higher in the homogeneity adjusted surface data set trends than the reanalysis data trends. In the same way that the diurnal and annual temperature signals are much greater on the surface than in the free air, it appears that the spatial variation in trend variability is much more pronounced in the surface data.

[36] As a whole one cannot assume that trends in surface temperatures at high-elevation sites are representative of trends in the free atmosphere at the same elevation, even if monthly surface anomalies at individual sites show a high degree of correlation with free-air anomalies. In most locations the surface is warming at a more rapid rate than the free air (on average, there is a threefold greater warming of surface than free-air temperatures), but it is dangerous to generalize. The reverse result occurs at a substantial minor-

ity of locations (~10%) and at yet more locations (~20%) there is no trend in ΔT.

6. Sensitivity of Trends to Spatial and Temporal Sampling

[37] Analyses were performed to assess the influence of sampling decisions on results. The choice of time period relates to concerns about the homogeneity of the reanalysis, particularly changes in radiosonde reporting times in 1958 and the introduction of satellite data in 1979. The choice of sites relates to concerns about data source and local topographical effects.

6.1. Time Period
6.1.1. Extension to 2002

[38] Sites in the GHCN data set end in 1998, whereas in the CRU data set some report additional data up to 2002. An assessment of the effect of this 4 year update on trend magnitudes was made through recalculation of trends for the 194 sites with this extra information. Changes were small for Ts, Ta and ΔT. The correlation between ΔT trends for the two periods is 0.947 and not one station changed the sign of its trend. Thus trends identified are relatively insensitive to the slight change in data period in this case.
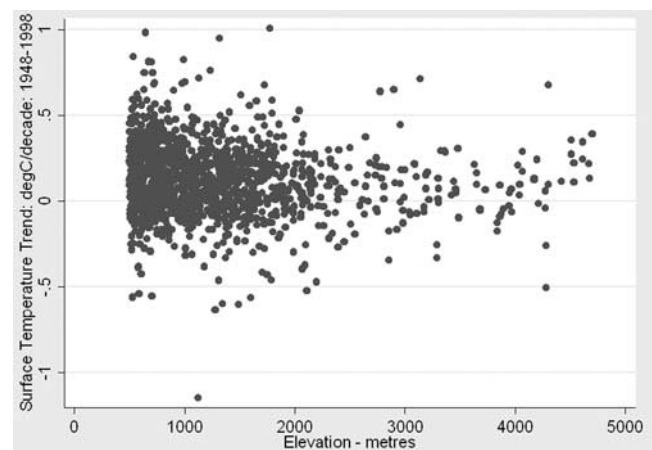
Figure 7. Surface temperature trend magnitude (°C/decade) versus elevation.
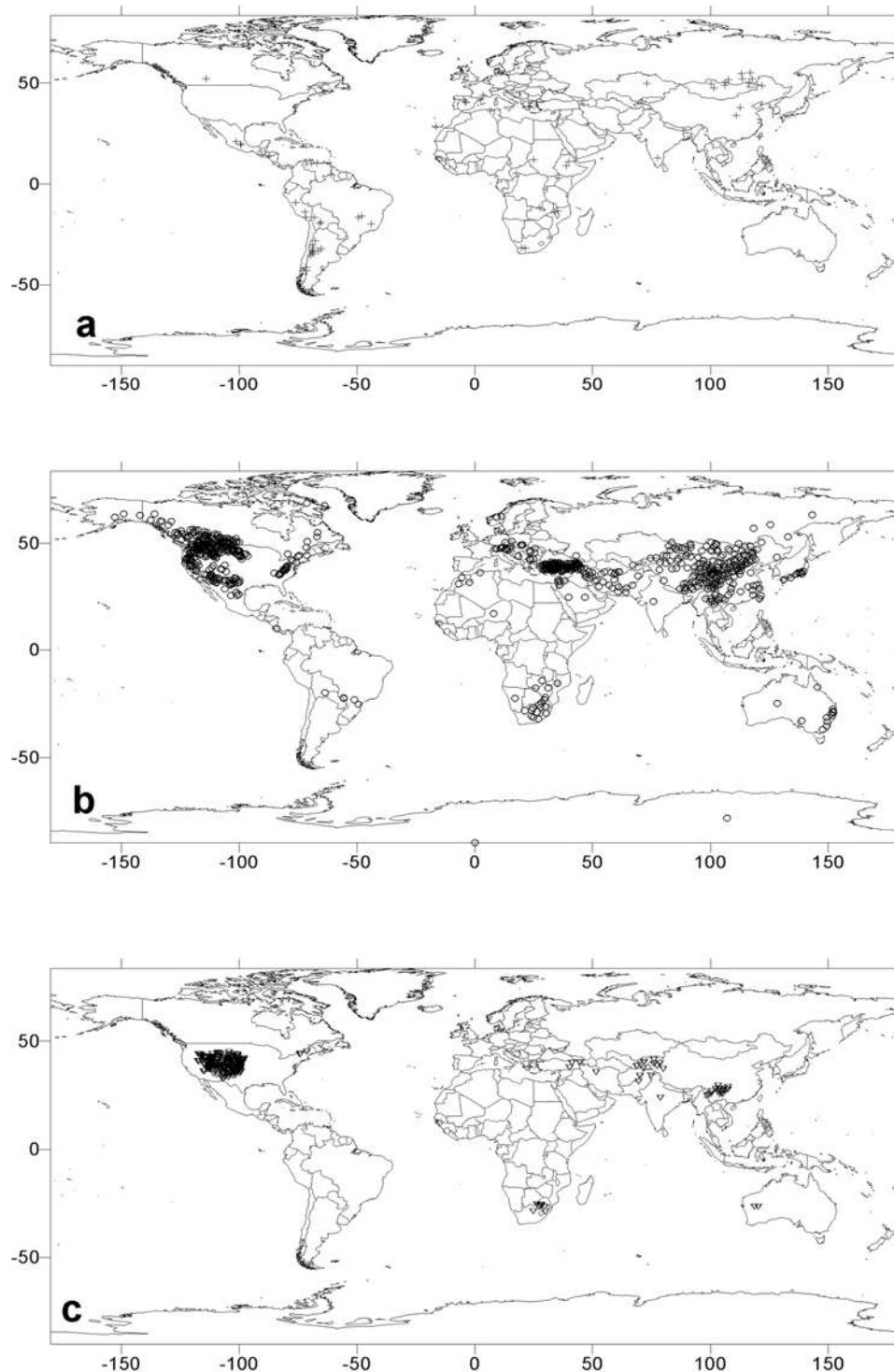
Figure 8. Maps of temporal trend in mean monthly free-air temperature anomalies (1948–1998) showing stations with (a) significant warming, (b) no trend, and (c) significant cooling.

6.1.2. Influence of Pre-1959 Data

[39] Similar trend comparisons were performed for 1959–1998 versus 1948–1998 to assess the impact of the radiosonde time change. Figures 12a–12c show the relationships between surface, free air and ΔT trends, respectively. In all cases the correlation between trend magnitudes for the two time periods are strong. Predictably, the free-air trends are the most unstable, but this does

not manifest itself in ΔT because most of the variability in ΔT trends is accounted for by surface trend variability. Over 90% of stations retain the same sign of ΔT trend for the two periods.

6.1.3. Influence of 1979 Satellite Introduction

[40] It is more difficult to assess the impact of the introduction of satellite data than that of the radiosonde time change above, since this occurred toward the middle of
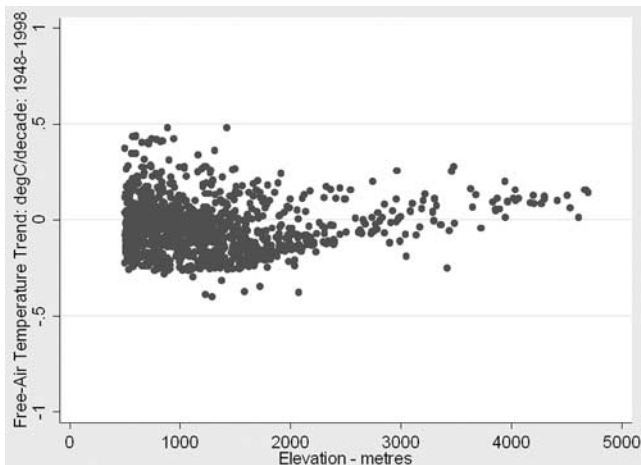
Figure 9. Free-air temperature trend magnitude (°C/decade) versus elevation.

the record. Thus trends for the satellite era might be expected to be substantially different from those of the whole period anyway. Figures 12d–12f (for surface, free air and ΔT) plot trends for 1959–1978 (presatellite) versus 1959–1998 (longer period). The homogenized surface data does show some difference in trends between the two periods (r = 0.540). The reanalysis shows much more coherence (r = 0.711). As a result the ΔT trends show reasonable consistency, again less than 20% changing sign. Finally, trends for 1979–1998 (satellite era) were also compared with those for 1959–1998. Again, surface trend magnitudes showed reasonable correlation (r = 0.519), but free-air trends less so (r = 0.236). Thus the change in the reanalysis that dominates the long-period trend is concentrated in the presatellite era. ΔT trends were somewhat more sensitive to time period in this case, with the trend for 41% of stations changing sign. In all three periods, 1959–1998, 1959–1978 and 1979–1998, the median ΔT trend is positive (the surface is warming more rapidly than free air), but the value is only 0.04°C/decade in the later period as opposed to 0.15°C/decade for the whole period.

[41] With the possible exception of 1979–1998, the sign of the ΔT trend is not very sensitive to time period. This may be because ΔT trend variability is presently dominated by the surface trend variability, free-air trends being both weak and less spatially variable. However, if the reanalysis trends are unrealistically weak, any systematic error in free-air trends could have influenced this result.

6.2. Spatial Sampling

[42] The decision to include as many stations as possible in the initial study by using a conservative elevation threshold of 500 m was driven by a concern to be spatially extensive. However, this did include using relatively low-elevation sites, sites in mountain valleys, and sites below the elevation of the surrounding reanalysis grid points. The consequences of this on the surface versus free-air anomaly correlations has been illustrated. There is no reason to imagine that boundary layer effects would systematically bias the trend results. However, mean trends (1948–1998) were calculated for subsets of stations to assess this possi-

bility. Subsets of stations were defined using four methods: data source, anomaly correlation, topography and relationship to surrounding reanalysis grid elevations.

[43] The influence of data source on trends was examined by comparing mean trends for (1) GHCN sites outside the United States, (2) GHCN sites within the United States, which were originally part of the HCN, and (3) CRU sites. There were significant differences in mean surface, free-air and ΔT trends, although absolute differences were small (not shown). The non-U.S. GHCN and CRU stations were similar, whereas the U.S. stations stood out as unusual, particularly for ΔT trends which were much more strongly positive on average than the global mean. However, since there is also inevitable spatial clustering of the three types of station, some of this difference could be spatially induced rather than due to the different methods used to develop homogeneity [*Peterson and Vose*, 1997].

[44] It could be asserted that trends based on stations that show a high anomaly correlation are more globally representative because of reduced boundary layer effects. The correlations between trend magnitudes and the surface/free-air anomaly correlation from section 4 (a surrogate measure of confidence in the spatial representativeness of the station) were examined. In all cases there is no significant relationship. For sites with poor anomaly correlations there are often weak surface and free-air trends (not shown). There is certainly no tendency for trends to be inflated at the stations where the anomaly correlation is weak. Table 5 compares mean and median trends for all sites versus trends for sites with anomaly correlations above and below 0.9. In all cases there is no significant difference in mean or median trend magnitudes.

[45] However, there are significant differences in ΔT trends between sites of differing topographic type. Although the mean ΔT trend is positive in all four topographic categories, flat sites (FL) show the largest trend, and mountaintop sites (MT) show the weakest (meaning that they are behaving more similarly to the free air).

[46] Finally, ΔT trends were examined according to how far the surface site was above or below the surrounding reanalysis grid points. There is some substantial overlap with topography here since mountaintop sites tend to be well above the surrounding grid, and valley sites below. For the 100 sites more than 500 m above the surrounding grid the mean ΔT trend is still positive (0.06°C/decade) but lower than the trend for all sites below the reanalysis topography (0.23°C/decade). This difference is statistically significant, and is consistent with the analysis by topography. Thus the difference in observed trends between surface and free-air temperatures is larger at sites where boundary layer effects are expected to be more influential.

7. Discussion and Conclusions

[47] An analysis of trends in surface temperatures, free-air temperatures at the same elevation, and the difference between the two (ΔT) for 1084 stations shows that at a majority of stations the surface is warming more rapidly than the free air. This finding is relatively insensitive to data period, since most of the variability in ΔT trends is controlled by surface trend variability. Concerns about reanalysis homogeneity are still important since unrealisti-
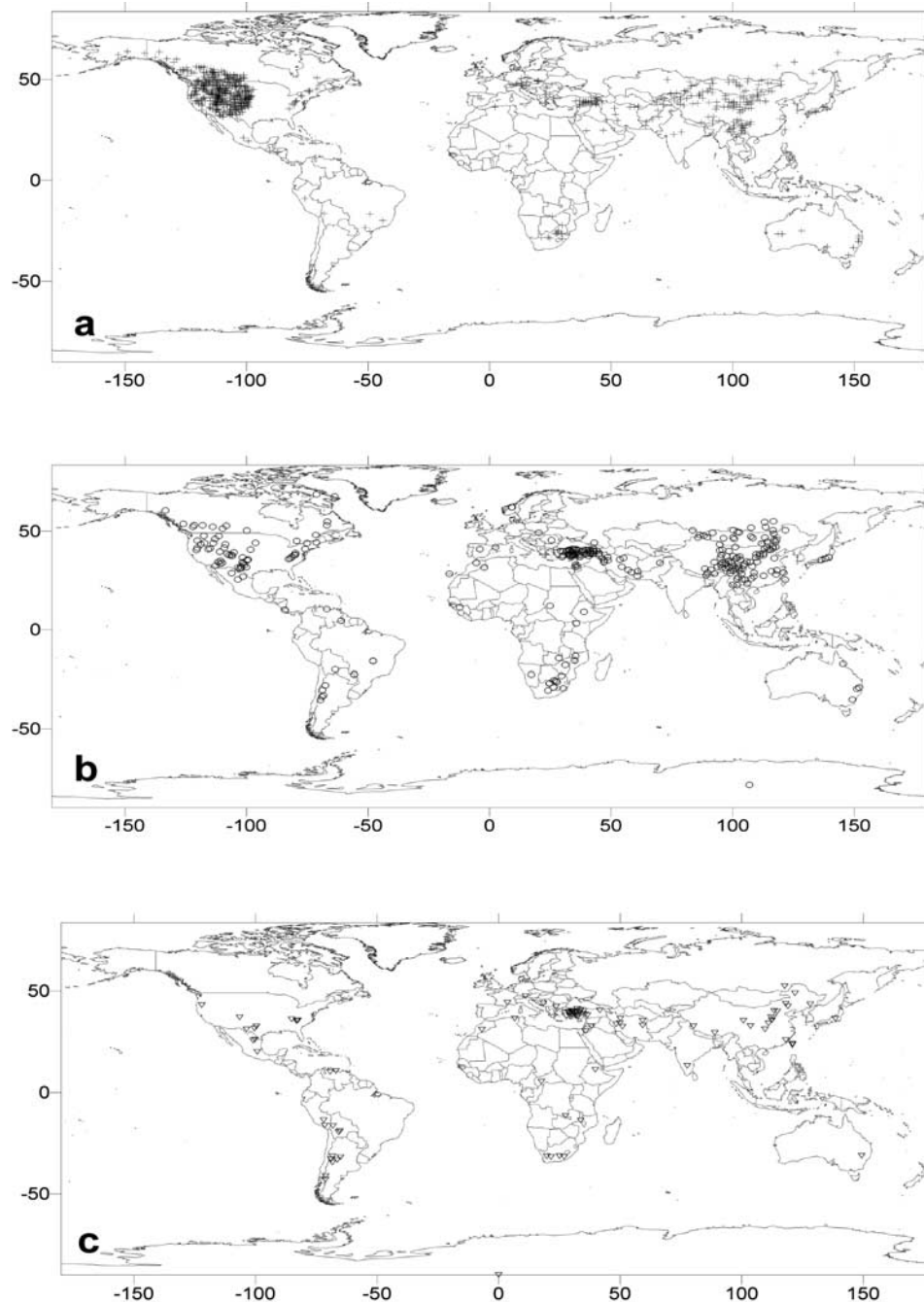
Figure 10. Maps of temporal trend in mean monthly ΔT anomalies (1948–1998) showing stations with (a) significant increase, (b) no trend, and (c) significant decrease.

cally weak free-air trends could lead to this statistical situation.

[48] The findings are sensitive to topography but not absolute elevation. True mountain summit sites show weaker increases in ΔT than valley sites (although both are significant). Thus the discrepancy between surface and free-air warming diminishes somewhat at mountain peaks. It is important not to confuse the topographical effect with elevation since many of the highest-elevation sites in the GHCN/CRU data sets are in valleys. This is an issue in regions such as the Tibetan plateau where observing sites are skewed toward high-elevation valleys with distinct microclimates.

[49] Past observational studies have shown that the main control of ΔT is expected to be the energy balance of the surface [Samson, 1965; Tabony, 1985; Barry, 1992]. The net radiation budget, cloud cover, the presence or absence of snow cover, and the strength of the airflow at the surface are physical controls. A more detailed examination of how these factors vary in their influence, over time and space, is beyond the scope of this paper. It may help explain the patterns of change identified here, in particular as to
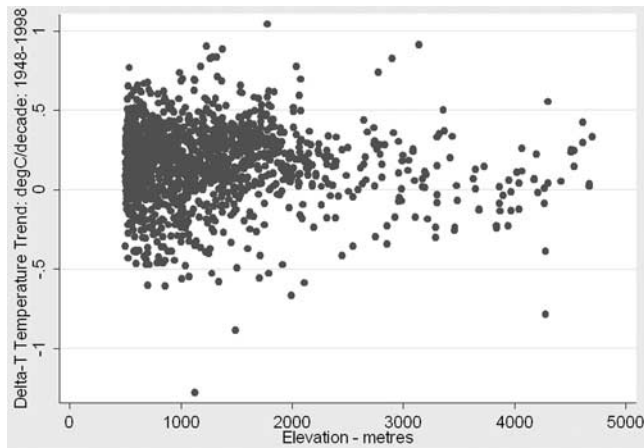
Figure 11. ΔT trend magnitude (°C/decade) versus elevation.

whether the ΔT changes identified are substantiated by attendant changes in snow cover, cloud cover etc. Unfortunately, the reanalysis is unreliable in its simulation of energy balance (in particular clouds) and so independent cloud and snow data will be preferred in future research.

[50] Because ΔT variability in this case is dominated by surface temperature variability, the quality of the surface data is important. As for the "free air," although the NCEP/NCAR temperatures are highly dependent on the combination of the raw radiosonde and satellite data upon which they are based, the NCEP/NCAR output is obscure. The relatively unstable nature of trends derived from the reanalysis is also an important lesson (Figure 12), and it would be beneficial to separate the free-air data sources by using radiosonde and satellite data separately in future analyses, although this would require interpolation from an irregular and sparse station network in the case of radiosondes, and vertically from satellite data. It may be more realistic to concentrate such effort on making comparisons where reliable radiosonde sites and surface high-elevation sites exist in close proximity, rather than attempting an extensive global comparison.
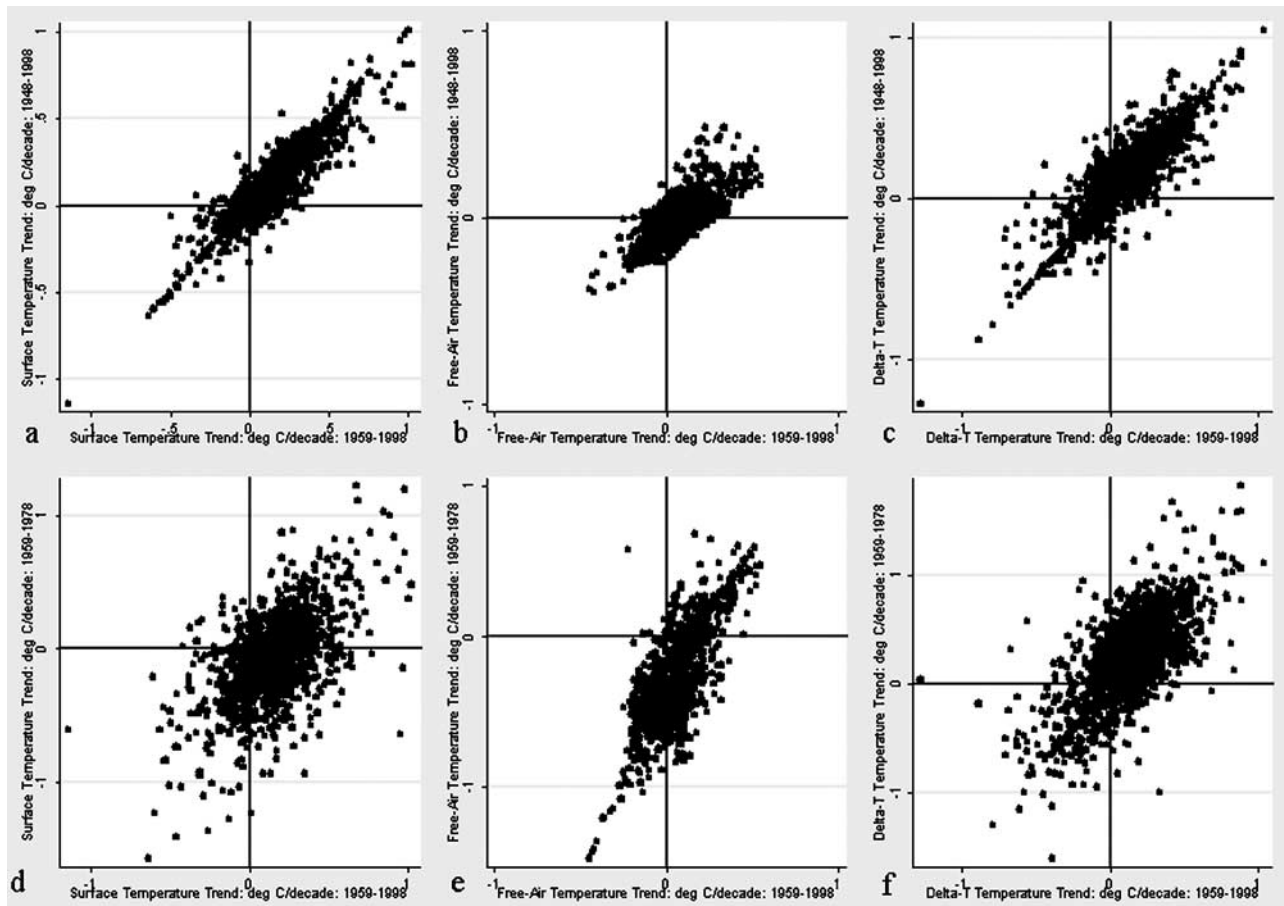
Figure 12. Trend magnitudes (°C/decade) for 1959–1998 (postradiosonde time change) versus 1948–1998 (whole period): (a) mean monthly surface temperature anomalies, (b) mean monthly free-air temperature anomalies, and (c) mean monthly ΔT anomalies. Trend magnitudes (°C/decade) for 1959–1978 (presatellite) versus 1959–1998 (longer period): (d) mean monthly surface temperature anomalies (e) mean monthly free-air temperature anomalies, and (f) mean monthly ΔT anomalies.

Table 5. Mean and Median Trends (1948–1998) for Stations With Differing Surface/Free-Air Anomaly Correlations[a]

Temperature	All Sites		$r \geq 0.9$		$r < 0.9$	
	Mean	Median	Mean	Median	Mean	Median
Surface trend	0.13	0.13	0.14	0.11	0.12	0.13
Free-air trend	−0.04	−0.05	−0.03	−0.03	−0.05	−0.06
ΔT trend	0.17	0.19	0.16	0.16	0.17	0.21

[a]Trends are given in °C/decade. There are no significant differences between columns.

[51] The significant changes in ΔT outlined in this paper illustrate a potential decoupling of the Earth's surface at high elevations from the free air in terms of response to radiative forcing. This change is dependent on location and time period. Mountain sites, although they show a high degree of affinity with free-air climate, may therefore not respond to global warming in ways we expect.

References

Barry, R. G. (1990), Changes in mountain climate and glacio-hydrological responses, *Mt. Res. Dev.*, *10*, 161–170.

Barry, R. G. (1992), *Mountain Weather and Climate*, 2nd ed., 402 pp., Routledge, Boca Raton, Fla.

Beniston, M., H. F. Diaz, and R. S. Bradley (1997), Climatic change at high elevation sites: An overview, *Clim. Change*, *36*, 233–252.

Cayan, D. R., and A. V. Douglas (1984), Urban influences on surface temperatures in southwestern United States during recent decades, *J. Clim. Appl. Meteorol.*, *23*, 1520–1530.

Christy, J. R., R. Spencer, and W. Braswell (2000), MSU tropospheric temperatures: Dataset construction and radiosonde comparisons, *J. Atmos. Oceanic Technol.*, *17*, 1153–1170.

Diaz, H. F., and R. S. Bradley (1997), Temperature variations during the last century at high elevation sites, *Clim. Change*, *36*, 253–279.

Diaz, H. F., and N. C. Graham (1996), Recent changes in tropical freezing heights and the role of sea surface temperature, *Nature*, *383*, 152–155.

Diaz, H. F., J. K. Eischeid, C. Duncan, and R. Bradley (2003), Variability of freezing levels, melting season indicators, and snow cover for selected high-elevation and continental regions in the last 50 years, *Clim. Change*, *59*, 33–52.

Everitt, B., S. Landau, and M. Leese (2001), *Cluster Analysis*, 4th ed., 248 pp., Edward Arnold, London.

Folland, C. K., et al. (2001), Observed climate variability and change, in *Climate Change 2001: The Scientific Basis*, edited by J. T. Houghton et al., 881 pp., Cambridge Univ. Press, New York.

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

Giorgi, F., J. W. Hurrell, M. R. Marinucci, and M. Beniston (1997), Elevation dependency of the surface climate signal: A model study, *J. Clim.*, *10*, 288–296.

Jones, P. D., and A. Moberg (2003), Hemispheric and large-scale surface air temperature variations: An extensive revision and update to 2001, *J. Clim.*, *16*, 206–223.

Jones, P. D., P. Y. Groisman, M. Coughlan, N. Plummer, W. C. Wang, and T. R. Karl (1990), Assessment of urbanisation effects in time series of surface air temperature over land, *Nature*, *347*, 169–172.

Jones, P. D., M. New, D. E. Parker, S. Martin, and I. G. Rigor (1999), Surface air temperature and its changes over the past 150 years, *Rev. Geophys.*, *37*, 173–199.

Kalnay, E., and M. Cai (2003), Impact of urbanization and land-use change on climate, *Nature*, *423*, 528–531.

Kalnay, E., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*, 437–471.

Kanamitsu, M., et al. (2002), NCEP-DOE AMIP-II reanalysis (R-2), *Bull. Am. Meteorol. Soc.*, *83*, 1631–1643.

Kaser, G., D. R. Hardy, T. Mölg, R. S. Bradley, and T. M. Hyera (2004), Modern glacier retreat on Kilimanjaro as evidence of climate change, observations and facts, *Int. J. Climatol.*, *24*, 329–339.

Kistler, R., et al. (2001), The NCEP/NCAR 50-year reanalysis, *Bull. Am. Meteorol. Soc.*, *82*, 247–268.

Lanzante, J. R., S. A. Klein, and D. J. Seidel (2003a), Temporal homogenisation of monthly radiosonde temperature data: Part 1: Methodology, *J. Clim.*, *16*, 224–240.

Lanzante, J. R., S. A. Klein, and D. J. Seidel (2003b), Temporal homogenisation of monthly radiosonde temperature data: Part 2: Trends, sensitivities and MSU comparison, *J. Clim.*, *16*, 241–262.

Mears, C. A., M. C. Schabel, and F. J. Wentz (2003), A reanalysis of the MSU channel 2 tropospheric temperature record, *J. Clim.*, *16*, 3650–3664.

National Research Council (NRC) (2000), *Reconciling Observations of Global Temperature Change*, 85 pp., Natl. Acad. Press, Washington, D. C.

Pepin, N. C. (2000), Twentieth century change in the climate record for the Front Range, Colorado, U.S.A., *Arct. Antarct. Alp. Res.*, *32*(2), 135–146.

Pepin, N. C., and M. L. Losleben (2001), Long term changes in surface screen and free-air equivalent temperatures in continental and maritime mid-latitude mountain ranges, paper presented at 2nd International Workshop on Climate Change at High Elevation Sites: Emerging Impacts, Past Global Changes, Davos, Switzerland, 25 –28 June.

Pepin, N. C., and M. L. Losleben (2002), Climate change in the Colorado Rocky Mountains: Free-air versus surface temperature trends, *Int. J. Climatol.*, *22*, 311–329.

Peterson, T. C. (2003), Assessment of urban vs rural in situ surface temperatures in the contiguous United States: No difference found, *J. Clim.*, *16*, 2941–2959.

Peterson, T. C., and R. S. Vose (1997), An overview of the Global Historical Climatology Network temperature database, *Bull. Am. Meteorol. Soc.*, *78*, 2837–2848.

Peterson, T. C., et al. (1998), Homogeneity adjustments of in situ atmospheric climate data: A review, *Int. J. Climatol.*, *18*, 1493–1517.

Peterson, T. C., K. P. Gallo, J. Lawrimore, T. W. Owen, A. Huang, and D. A. McKittrick (1999), Global rural temperature trends, *Geophys. Res. Lett.*, *26*, 329–332.

Richner, H., and P. D. Phillips (1984), A comparison of temperatures from mountaintops and the free atmosphere: Their diurnal variation and mean difference, *Mon. Weather Rev.*, *112*, 1328–1340.

Samson, C. A. (1965), A comparison of mountain slope and radiosonde observations, *Mon. Weather Rev.*, *93*, 327–330.

Santer, B. D., T. M. L. Wigley, J. S. Boyle, D. J. Gaffen, J. J. Hnilo, D. Nychka, D. E. Parker, and K. E. Taylor (2000), Statistical significance of trends and trend differences in layer-average atmospheric time series, *J. Geophys. Res.*, *105*, 7337–7356.

Seidel, D. J., and M. Free (2003), Comparison of lower-tropospheric temperature climatologies and trends at low and high elevation radiosonde sites, *Clim. Change*, *59*, 53–74.

Tabony, R. C. (1985), The variation of surface temperature with altitude, *Meteorol. Mag.*, *114*, 37–48.

Thompson, L. G. (2000), Ice core evidence for climate change in the tropics: Implications for our future, *Quat. Sci. Rev.*, *19*, 19–35.

Trenberth, K. (2004), Rural land-use change and climate, *Nature*, *427*, 213–214.

Vose, R. S., T. R. Karl, D. R. Easterling, C. N. Williams, and M. J. Menne (2004), Impact of land-use change on climate, *Nature*, *427*, 213–214.

Vuille, M., and R. S. Bradley (2000), Mean annual temperature trends and their vertical structure in the tropical Andes, *Geophys. Res. Lett.*, *27*, 3885–3888.

N. C. Pepin, Department of Geography, Buckingham Building, Lion Terrace, University of Portsmouth, Hants PO1 3HE, UK. (nicholas.pepin@port.ac.uk)

D. J. Seidel, Air Resources Laboratory, NOAA, Silver Spring, MD 20910, USA.