1. INTRODUCTION

Although warming from 1979 to 1998 is well supported, major questions exist about long-term trends. Climategate inspired investigations suggest global surface-station data are seriously compromised. The data suffer significant contamination by urbanization and other local factors such as land-use/
land-cover changes and instrument siting that does not meet government standards. There was a major station dropout, which occurred suddenly around 1990 and a significant increase in missing monthly data in the stations that remained. There are also uncertainties in ocean temperatures; no small issue, as oceans cover 71% of the Earth’s surface.

These factors all lead to significant uncertainty and in most cases a tendency for overestimation of century-scale temperature trends. Indeed, numerous peer-reviewed papers cataloged here have estimated that these local issues with the observing networks may account for 30%, 50% or more of the warming shown since 1880. After the data with all its issues are collected, further adjustments are made, each producing more warming,

“[W]hen data conflicts with models, a small coterie of scientists can be counted upon to modify the data” to agree with models’ projections,” says MIT meteorologist Dr. Richard Lindzen.

In this paper, we look at some of the issues in depth and the recommendations made for a reassessment of global temperatures necessary to make sensible policy decisions.

2. THE GLOBAL DATA CENTERS

Five organizations publish global temperature data. Two — Remote Sensing Systems (RSS) and the University of Alabama at Huntsville (UAH) — are satellite data sets. The three terrestrial data sets provided by the institutions — NOAA’s National Climatic Data Center (NCDC), NASA’s Goddard Institute for Space Studies (GISS/ GISTEMP), and the University of East Anglia’s Climatic Research Unit (CRU) — all depend on data supplied by surface stations administered and disseminated by NOAA under the management of the National Climatic Data Center in Asheville, North Carolina. The Global Historical Climatology Network (GHCN) is the most commonly cited measure of global surface temperature for the last 100 years.

Around 1990, NOAA/NCDC’s GHCN data set lost more than three-quarters of the climate measuring stations around the world. A study by Willmott et al. (1991) calculated a +0.2C bias in the global average owing to pre-1990 station closures. Douglas Hoyt had estimated approximately the same value in 2001 due to station closures around 1990. A number of station closures can be attributed to Cold-War era military base closures, such as the DEW Line (The Distant Early Warning Line) in Canada and its counterpart in Russia.

The world’s surface observing network had reached its golden era in the 1960s to 1980s, with more than 6,000 stations providing valuable climate information. Now, there are fewer than 1,500 remaining.

It is a fact that the three data centers each performed some final adjustments to the gathered data before producing their own final analysis. These
adjustments are frequent and often poorly documented. The result was almost always to produce an enhanced warming even for stations which had a cooling trend in the raw data. The metadata, the information about precise location, station moves and equipment changes were not well documented and shown frequently to be in error which complicates the assignment to proper grid boxes and make the efforts of the only organization that attempts to adjust for urbanization, NASA GISS problematic.

As stated here relative to Hansen et al. (2001),1 “The problem [accuracy of the latitude/longitude coordinates in the metadata] is, as they say, “even worse than we thought.” One of the consumers of GHCN metadata is of course GISTEMP, and the implications of imprecise latitude/longitude for GISTEMP are now considerably greater, following the change in January 2010 to use of satellite-observed night light radiance to classify stations as rural or urban throughout the world, rather than just in the contiguous United States as was the case previously. As about a fifth of all GHCN stations changed classification as a result, this is certainly not a minor change.”

Among some major players in the global temperature analyses, there is even disagreement about what the surface air temperature really is. (See “The Elusive Absolute Surface Air Temperature (SAT)” by Dr. James Hansen.2 Essex et al. questioned whether a global temperature existed here.3)

Satellites measurements of the lower troposphere (around 600 mb) are clearly the better alternative. They provide full coverage and are not contaminated by local factors. Even NOAA had assumed satellites would be the future solution for climate monitoring. Some have claimed satellite measurements are subject to error. RSS and UAH in 2005 jointly agreed4 that there was a small net cold bias of 0.03 °C in their satellite-measured temperatures, and corrected the data for this small bias. In contrast, the traditional surface-station data we will show suffer from many warm biases that are orders of magnitude greater in size than the satellite data, yet that fact is often ignored by consumers of the data.

Some argue that satellites measure the lower atmosphere and that this is not the surface. This difference is real but it is irrelevant (CCSP5). The lower troposphere around 600 mb was chosen because it was above the mixing level and so with polar orbiters the issues of the diurnal variations are eliminated. Also there is a high correlation between temperatures in the lower to middle troposphere and the surface.

Anomalies from satellite data and surface-station data have been increasing in the last 3 decades. When the satellites were first launched, their temperature readings were in better agreement with the surface-station data. There has been increasing divergence over time which can be seen below (derived from Klotzbach et al., 2009). In the first plot, we see the temperature anomalies as computed from the satellites and assessed by UAH and RSS and the station-based land surface anomalies from NOAA NCDC (Fig. 1).

The divergence is made clearer when the data are scaled such that the difference in 1979 is zero (Fig. 2).

The Klotzbach paper finds that the divergence between surface and lower-tropospheric trends is consistent with evidence of a warm bias in the surface temperature record but not in the satellite data.

Klotzbach et al. described an ‘amplification’ factor for the lower troposphere as suggested by Santer et al. (2005) and Santer et al. (2008) due to greenhouse gas trapping relative to the warming at the surface. Santer refers to the effect as “tropospheric amplification of surface warming.” This effect is a characteristic of all of the models used in the UNIPCC and the USGCRP “ensemble” of models by Karl et al. (2006) which was the source for Karl et al. (2009) which in turn was relied upon by EPA in its recent Endangerment Finding. (Federal Register/Vol. 74, No. 239/Tuesday, December 15, 2009/Rules and Regulations at 66510.)

As Dr. John Christy, keeper of the UAH satellite data set describes it, “The amplification factor is a direct calculation from model simulations that show over 30-year periods that the upper air warms at a faster rate than the surface — generally 1.2 times faster for global averages. This is the so-called “lapse rate feedback” in which the lapse rate seeks to move toward the moist adiabat as the

![Annual Land Surface vs Equivalent Lower Troposphere Anomalies](image)

**FIGURE 1** Annual land surface anomalies compared to UAH and RSS lower-tropospheric temperature anomalies since 1979 (sources: NOAA and Klotzbach).
surface temperature rises. In models, the convective adjustment is quite rigid, so this vertical response in models is forced to happen. The real world is much less rigid and has ways to allow heat to escape rather than be retained as models show.” This latter effect has been documented by Chou and Lindzen (2005) and Lindzen and Choi (2009).

**FIGURE 2** NOAA annual land temperatures minus annual UAH lower troposphere (blue line) and NOAA annual land temperatures minus annual RSS lower troposphere (green line) over the period from 1979 to 2008.

**FIGURE 3** Model amplification-based forecast lower troposphere (blue line) and actual UAH (green line) and RSS lower troposphere (purple line) over the period from 1979 to 2008.
The amplification factor was calculated from the mean and median of the 19 GCMs that were in the CCSP SAP 1.1 report (Karl et al., 2006). A fuller discussion of how the amplification factor was calculated is available in the Klotzbach paper.6

The ensemble model forecast curve (upper curve) in Fig. 3 was calculated by multiplying the NOAA NCDC surface temperature for each year by the amplification factor, and thus is the model projected tropospheric temperature. The lower curves are the actual UAH and RSS lower-tropospheric satellite temperatures.

This strongly suggests that instead of atmospheric warming from greenhouse effects dominating, surface-based warming very likely due to uncorrected urbanization and land-use contamination is the biggest change. Since these surface changes are not fully adjusted for, trends from the surface networks are not reliable.

3. THE GOLDEN AGE OF SURFACE OBSERVATION

In this era of ever-improving technology and data systems, one would assume that measurements would be constantly improving. This is not the case with the global station observing network. The Golden Age of Observing was several decades ago. It is gone.

The Hadley Centre’s Climate Research Unit (CRU) at East Anglia University is responsible for the CRU global data. NOAA’s NCDC, in Asheville, NC, is the source of the Global Historical Climate Network (GHCN) and of the U.S. Historical Climate Network (USHCN). These two data sets are relied upon by NASA’s GISS in New York City and by Hadley/CRU in England.

All three have experienced degradation in data quality in recent years.

Ian “Harry” Harris, a programmer at the Climate Research Unit, kept extensive notes of the defects he had found in the data and computer programs that the CRU uses in the compilation of its global mean surface temperature anomaly data set. These notes, some 15,000 lines in length, were stored in the text file labeled “Harry_Read_Me.txt”, which was among the data released by the whistle-blower with the Climategate emails. This is just one of his comments:

“[The] hopeless state of their (CRU) database. No uniform data integrity, it’s just a catalogue of issues that continues to grow as they’re found…I am very sorry to report that the rest of the databases seem to be in nearly as poor a state as Australia was. There are hundreds if not thousands of pairs of dummy stations, one with no WMO and one with, usually overlapping and with the same station name and very similar coordinates.

I know it could be old and new stations, but why such large overlaps if that’s the case? Aarrggghhh! There truly is no end in sight.

“This whole project is SUCH A MESS. No wonder I needed therapy!!

“I am seriously close to giving up, again. The history of this is so complex that I can’t get far enough into it before by head hurts and I have to stop. Each parameter has a tortuous history of manual and semi-automated interventions that I simply cannot just go back to early versions and run the updateprog. I could be throwing away all kinds of corrections - to lat/lons, to WMOs (yes!), and more. So what the hell can I do about all these duplicate stations?”

According to Phil Jones, former director of the Climatic Research Unit (CRU), ‘there is some truth’ to the charge that he failed to update and organize the raw data supporting the CRU temperature data set, on which the IPCC relies in its reports to make temperature projections and that at least some of the original raw data were lost. This should raise questions about the quality of global data.

In the following email, CRU’s Director at the time, Dr. Phil Jones, acknowledges that CRU mirrors the NOAA data:

“Almost all the data we have in the CRU archive is exactly the same as in the GHCN archive used by the NOAA National Climatic Data Center.”

In the Russell inquiry into CRU’s role in Climategate, they estimated at least 90% of the data were the same. Steve McIntyre’s analysis showed 95.6% concordance. NASA uses the GHCN as the main data source for the NASA GISS data.

Dr. Roger Pielke Sr. in this post on the three data sets notes:

“The differences between the three global surface temperatures that occur are a result of the analysis methodology as used by each of the three groups. They are not “completely independent.” Each of the three surface temperature analysis suffer from unresolved uncertainties and biases as we documented, for example, in our peer reviewed paper.”

Dr. Richard Anthes, President of the University Corporation for Atmospheric Research, in testimony to Congress in March 2009, noted:

“The present federal agency paradigm with respect to NASA and NOAA is obsolete and nearly dysfunctional, in spite of best efforts by both agencies.”

4. VANISHING STATIONS

More than 6,000 stations were in the NOAA database for the mid-1970s, but just 1,500 or less are used today. NOAA claims the real-time network includes 1,200 stations with 200–300 stations added after several months and included in the annual numbers. NOAA is said to be adding additional U.S. stations now that USHCN v2 is available, which will inflate this number, but make it disproportionately U.S.

There was a major disappearance of recording stations in the late 1980s to the early 1990s. Figure 4 compares the number of global stations in 1900, the 1970s, and 1997, showing the increase and then decrease (Peterson and Vose\(^ {10}\)).

Dr. Kenji Matsuura and Dr. Cort J. Willmott at the University of Delaware have prepared this animation.\(^ {11}\) See the lights go out in 1990, especially in Asia.

---

The following chart of all GHCN stations and the average annual temperature show the drop focused around 1990. In this plot, those stations with multiple locations over time are given separate numbers, which inflates the total number. While a straight average is not meaningful for global temperature calculation (because areas with more stations would have higher weighting), it illustrates that the disappearance of so many stations in an uneven fashion may have introduced a distribution bias (Fig. 5).

As can be seen in the figure, the straight average of all global stations does not fluctuate much until 1990, at which point the average temperature jumps up. This observational bias can influence the calculation of area-weighted averages to some extent. As previously noted, a study by Willmott, Robeson, and Feddema (“Influence of Spatially Variable Instrument Networks on Climatic Averages”, 1991) calculated a +0.2C bias in the global average owing to pre-1990 station closures. Others have attempted experiments (Mosher, Grant, Lilligren) that purport to show this does not necessarily translate into a warm bias given the ‘anomaly method’ (using anomalies or departures from normal base period values instead of the actual temperatures). The effect may not be definitively known until a full data reconstruction can take place.

Global databases all compile data into latitude/longitude-based grid boxes and calculate temperatures inside the boxes using data from the stations within them or use the closest stations (weighted by distance) in nearby boxes. The use of anomalies instead of mean temperatures greatly improve the chances of filling in some of the smaller holes (empty grid boxes) or not producing significant differences in areas where the station density is high, they

FIGURE 5 Plot of the number of total station ID’s in each year since 1950 and the average temperatures of the stations in the given year.

can’t be relied on to accurately estimate anomalies in the many large data sparse areas (Canada, Greenland, Brazil, Africa, parts of Russia). To fill in these areas requires NOAA and NASA to reach out as far as 1200 km.

There are 8,000 grid boxes globally (land and sea). If the Earth is 71% ocean, approximately 2,320 grid boxes would be over land (actual number will vary as some grid boxes will overlap or may just touch the coast).

With 1,200 stations in the real-time GHCN network that would be enough to have 51.7% of the land boxes with a station. However, since stations tend to cluster, that number is smaller. Our calculation is that that number is around 44% or 1,026 land grid boxes without a station.

For data in empty boxes, GHCN will look to surrounding areas as far away as 1,200 km (in other words using Atlanta, GA to estimate a monthly or annual anomaly in Chicago, IL, Birmingham Al to estimate New York City, Los Angeles to estimate Jackson Hole, WY).

Certainly an isolated vacant grid box surrounded by boxes with data in them may be able to yield a reasonably representative extrapolated anomaly value from the surrounding data.

But in data sparse regions, such as is much of the Southern Hemisphere, when you have to extrapolate from more than one grid box away you are increasing the data uncertainty. If you bias it towards having to look towards more urbanized or airport regions or lower elevation coastal locations as E.M. Smith has detected, you are added potential warm bias to uncertainty. This has been the case in the north in the large countries bordering on the arctic (Russia and Canada) where the greatest warming is shown in the data analyses but also in Brazil where fast growing cities are used to estimate anomalies in the Amazon.

To ascertain whether a net bias exists, E.M. Smith has conducted first an analysis of mean temperatures for whatever stations existed by country or continent/sub continent. He then applied a dT method\(^\text{13}\) which is a variation of ‘First Differences’ as a means of examining temperature data anomalies independent of actual temperature. dT/year is the “average of the changes of temperature, month now vs. the same month that last had valid data, for each year”. He then does a running total of those changes, or the total change, the “Delta Temperature” to date. He is doing this for every country (see footnote 14). His next step will be to attempt to splice/blend the data into the grids.

Even then uncertainty will remain that only more complete data set usage would improve. The following graphic powerfully illustrates this was a factor even before the major dropout. Brohan (2006) showed the degree of uncertainty in surface temperature sampling errors for 1969 (here for CRUTEM3). The degree of uncertainty exceeds the total global warming signal (Fig. 6).

\(^\text{13}\) http://chiefio.wordpress.com/2010/02/28/last-delta-t-an-experimental-approach/.

\(^\text{14}\) http://chiefio.wordpress.com/.
5. SEE FOR YOURSELF — THE DATA IS A MESS

Look for yourself following these directions using the window into the NOAA, GHCN data provided by NASA GISS.\(^\text{15}\) Point to any location on the world map (say, central Canada). You will see a list of stations and approximate populations. Locations with less than 10,000 people are assumed to be rural (even though Oke has shown a town of 1,000 can have an urban warming bias of 2.2\(\text{C}\)).

You will see that the stations have a highly variable range of years with data. Try to find a few stations where the data extend to the current year. If you find some, you will likely see gaps in the graphs. To see how incomplete the data set is for that station, click in the bottom left of the graph Download monthly data as text. For many, many stations you will see the data set in a monthly tabular form has many missing data months mostly after 1990 (designated by 999.9) (Fig. 7).

These facts suggest that the golden age of observations was in the 1950s to 1980s. Data sites before then were more scattered and did not take data at standardized times of day. After the 1980s, the network suffered from loss of stations and missing monthly data. To fill in these large holes, data were extrapolated from greater distances away.

Indeed this is more than just Russia. Forty percent of GHCN v2 stations have at least one missing month (Fig. 8).

This is concentrated in the winter months as analyst Verity Jones has shown here.\(^\text{16}\)

---

\(^{15}\) http://data.giss.nasa.gov/gistemp/station_data/

As Verity Jones notes “Much of the warming signal in the global average data can be traced to winter warming (lows are not as low). If we now have a series of cooler years, particularly cooler winter months with lower lows, my concern is that missing months, particularly winter months could lead to a warm bias.”
NOAA tells us that by 2020, we will have as much data for the 1990s and 2000s as we had in the 1960s and 1970s. We are told that other private sources have been able to assemble more complete data sets in near real time (example: WeatherSource). Why can’t our government with a budget far greater than these private sources do the same or better? This question has been asked by others in foreign nations.

6. STATION DROPOUT WAS NOT TOTALLY RANDOM

6.1. Canada

After 1990, just one thermometer remains in the database for everything north of the 65th parallel. That station is Eureka, which has been described as “The Garden Spot of the Arctic” thanks to the flora and fauna abundant around the Eureka area, more so than anywhere else in the High Arctic. Winters are frigid but summers are slightly warmer than at other places in the Canadian Arctic.

NOAA GHCN used only 35 of the 600 Canadian stations in 2009, down from 47 in 2008. A case study by Tim Ball confirmed Environment Canada claims that weather data are available elsewhere from airports across Canada and indeed hourly readings can be found on the internet for many places in Canada (and Russia) not included in the global databases. Environment Canada reported in the National Post,17 that there are 1,400 stations in Canada with 100 north of the Arctic Circle, where NOAA uses just one. See E.M. Smith’s analysis in footnote 18.

Verity Jones plotted the stations from the full network rural, semi-rural and urban for Canada and the northern United States both in 1975 and again in 2009. She also marked with diamonds the stations used in the given year. Notice the good coverage in 1975 and very poor, virtually all in the south in 2009. Notice the lack of station coverage in the higher latitude Canadian region and arctic in 2009 (Fig. 9).

6.2. New Zealand and Australia

Smith found that in New Zealand the only stations remaining had the words “water” or “warm” in the descriptor code. Some 84% of the sites are at airports, with the highest percentage in southern cold latitudes.

In Australia, Torok et al. (2001),19 observed that in European and North American cities urban—rural temperature differences scale linearly with the logarithms of city populations. They also learned that Australian city heat islands are generally smaller than those in European cities of similar size.

which in turn are smaller than those in North American cities. The regression lines for all three continents converge in the vicinity of a population of 1,000 people, where the urban—rural temperature difference is approximately $2.2 \pm 0.2 \, ^\circ C$, essentially the same as what Oke (1973) had reported two decades earlier.

Smith finds the Australian dropout\(^{20}\) was mainly among higher latitude, cooler stations after 1990, with the percentage of city airports increasing to 71%, further enhancing apparent warming. The trend in “island Pacific without Australia and without New Zealand” is dead flat. The Pacific Ocean islands are NOT participating in “global” warming. Changes of thermometers in Australia and New Zealand are the source of any change.

### 6.3. Turkey

Turkey had one of the densest networks of stations of any country. E.M. Smith calculated anomaly process similar to First Differences. Then dT is the running total of those changes, or the total change, the “Delta Temperature” to date. Note the step up after 1990 cumulative change in temperature and the change per year for Turkey.\(^{21}\)

---

His dT method is a variation of ‘First Differences’ as a means of examining temperature data anomalies independent of actual temperature. dT/year is the “average of the changes of temperature, month now vs. the same month that last had valid data, for each year” (Fig. 10).

![Turkey dT(°C) vs dT/year](image)

**FIGURE 10** Smith analysis of Turkey temperatures using ‘First Differences’.

Verity Jones maps showing station temperature trends for (top) all stations active during 1880–2010 and (bottom) for stations active after 1990. The result is that Turkey is shown to be warming when the data shows cooling.

**FIGURE 11**

---


Despite that apparent warming, the Turkish Met Service finds evidence for cooling. This peer-reviewed paper: Murat Turke, Utku M. Sumer, Gonul Kilic, State Meteorological Service, Department of Research, Climate Change Unit, 06120 Kalaba-Ankara, Turkey which concludes:

“Considering the results of the statistical tests applied to the 71 individual stations data, it could be concluded that annual mean temperatures are generally dominated by a cooling tendency in Turkey.” See in Verity Jones website Digging in the Clay here\(^\text{23}\), the dropout of stations from nearly 250 to 39 leaving behind warming stations. 25 of the 39 stations are shown as the other stations did not have complete enough data to determine a reliable trend (less than 10 years without missing months) (Fig. 11).

7. INSTRUMENT CHANGES AND SITING

The World Meteorological Organization (WMO), a specialized agency of the United Nations,\(^\text{24}\) grew out of the International Meteorological Organization (IMO), which was founded in 1873. Established in 1950, the WMO became the specialized agency of the United Nations (in 1951) for meteorology, weather, climate, operational hydrology, and related geophysical sciences.

According to the WMO’s own criteria, followed by the NOAA’s National Weather Service, temperature sensors should be located on the instrument tower at 1.5 m (5 feet) above the surface of the ground. The tower should be on flat, horizontal ground surrounded by a clear surface, over grass or low vegetation kept less than 4 inches high. The tower should be at least 100 m (110 yards) from tall trees, or artificial heating or reflecting surfaces, such as buildings, concrete surfaces, and parking lots.

Very few stations meet these criteria.

8. ALONG COMES ‘MODERNIZATION’

The modernization of weather stations in the United States replaced many human observers with instruments that initially had major errors, or had “warm biases” (HO-83) or were designed for aviation and were not suitable for precise climate trend detection [Automatic Surface Observing Systems (ASOS) and the Automated Weather Observing System (AWOS)]. Also, the new instrumentation was increasingly installed on unsuitable sites that did not meet the WMO’s criteria.

During recent decades there has been a migration away from old instruments read by trained observers. These instruments were generally in shelters that were properly located over grassy surfaces and away from obstacles to ventilation and heat sources.

\(^{23}\) http://diggingintheclay.blogspot.com/2010/03/no-more-cold-turkey.html.

Today we have many more automated sensors (The MMTS) located on poles cabled to the electronic display in the observer’s home or office or at airports near the runway where the primary mission is aviation safety.

The installers of the MMTS instruments were often equipped with nothing more than a shovel. They were on a tight schedule and with little budget. They often encountered paved driveways or roads between the old sites and the buildings. They were in many cases forced to settle for installing the instruments close to the buildings, violating the government specifications in this or other ways (Fig. 12).

Pielke and Davey (2005) found a majority of stations, including climate stations in eastern Colorado, did not meet WMO requirements for proper siting.

They extensively documented poor siting and land-use change issues in numerous peer-reviewed papers, many summarized in the landmark paper Unresolved issues with the assessment of multi-decadal global land surface temperature trends25 (2007).

In a volunteer survey project, Anthony Watts and his more than 650 volunteers www.surfacestations.org found that over 900 of the first 1,067

stations surveyed in the 1,221 station U.S. climate network did not come close to meeting the specifications. Only about 3% met the ideal specification for siting. They found stations located next to the exhaust fans of air conditioning units, surrounded by asphalt parking lots and roads, on blistering-hot rooftops, and near sidewalks and buildings that absorb and radiate heat. They found 68 stations located at wastewater treatment plants, where the process of waste digestion causes temperatures to be higher than in surrounding areas. In fact, they found that 90% of the stations fail to meet the National Weather Service’s own siting requirements that stations must be 30 m (about 100 feet) or more away from an artificial heating or reflecting source.

The average warm bias for inappropriately-sited stations exceeded 1 °C using the National Weather Service’s own criteria, with which the vast majority of stations did not comply.

A report from last spring with some of the earlier findings can be found in footnote 26. Some examples from these sources (Figs. 13, 14):

As of October 25, 2009, 1,067 of the 1,221 stations (87.4%) had been evaluated by the surfacestations.org volunteers and evaluated using the Climate Reference Network (CRN) criteria.27 90% were sited in ways that result in errors exceeding 1 °C according to the CRN handbook.

This siting issue remains true even by the older “100 foot rule” criteria for COOP stations, specified by NOAA28 for the U.S. Cooperative Observer

Numerous sensors are located at waste treatment plants. An infrared image of the scene shows the output of heat from the waste treatment beds right next to the sensor. (Photos by Anthony Watts, surfacestations.org.)

FIGURE 14 One of many waste treatment plants serving as stations in USHCN.

FIGURE 15 Surfacestations.org quality rating by stations for 1,067 U.S. climate stations as of 10/25/2009. Only 10% meet minimal CRN ranking (CRN 1 or 2).
network where they specify “The sensor should be at least 100 feet (~30 m) from any paved or concrete surface (Fig. 15).”

Dr. Vincent Gray, IPPC Reviewer for AR1 through IV published on some issues related to temperature measurements.29

In 2008, Joe D’Aleo asked NCDC’s Tom Karl about the problems with siting and about the plans for a higher quality Climate Reference Network (CRN — at that time called NERON). He said he had presented a case for a more complete CRN network to NOAA but NOAA said it was unnecessary because they had satellite monitoring. The Climate Reference Network was capped at 114 stations and would not provide meaningful trend assessment for about 10 years. NOAA has since reconsidered and now plans to upgrade about 1,000 climate stations, but meaningful results will be even further in the future.

In monthly press releases no satellite measurements are ever mentioned, although NOAA claimed that was the future of observations.

9. ADJUSTMENTS NOT MADE, OR MADE BADLY

The Climategate whistle-blower proved what those of us dealing with data for decades already knew. The data were not merely degrading in quantity and quality: they were also being manipulated. This is done by a variety of post measurement processing methods and algorithms. The IPCC and the scientists supporting it have worked to remove the pesky Medieval Warm Period, the Little Ice Age, and the period emailer Tom Wigley referred to as the “warm 1940s blip”. There are no adjustments in NOAA and Hadley data for urban contamination. The adjustments and non-adjustments instead increased the warmth in the recent warm cycle that ended in 2001 and/or inexplicably cooled many locations in the early record, both of which augmented the apparent trend.

10. HEAT FROM POPULATION GROWTH AND LAND-USE CHANGES

10.1. Urban Heat Island

Weather data from cities as collected by meteorological stations are indisputably contaminated by urban heat-island bias and land-use changes. This contamination has to be removed or adjusted for in order to accurately identify true background climatic changes or trends. In cities, vertical walls, steel and concrete absorb the sun’s heat and are slow to cool at night. More and more of the world is urbanized (population increased from 1.5 B in 1900 to 6.8 B in 2010).

The urban heat-island effect occurs not only for big cities but also for towns. Oke (who won the 2008 American Meteorological Society’s Helmut Landsberg award for his pioneer work on urbanization the effect of urbanization on local microclimates) had a formula for the warming that is tied to population. Oke (1973) found that the urban heat-island (in °C) increases according to the formula:

\[ \text{Urban heat-island warming} = 0.317 \ln P, \quad \text{where } P = \text{population}. \]

Thus a village with a population of 10 has a warm bias of 0.73 °C. A village with 100 has a warm bias of 1.46 °C, and a town with a population of 1,000 people has a warm bias of 2.2 °C. A large city with a million people has a warm bias of 4.4 °C.

Goodrich (1996) showed the importance of urbanization to temperatures in his study of California counties in 1996. He found for counties with a million or more population the warming from 1910 to 1995 was 4 °F, for counties with 100,000 to 1 million it was 1 °F, and for counties with less than 100,000 there was no change (0.1 °F) (Fig. 16).

11. U.S. CLIMATE DATA

Compared to the GHCN global database, the USHCN database is more stable (Fig. 17).
FIGURE 17 Comparison of number of GHCN temperature stations in the U.S. vs. rest of the world (ROW). http://www.appinsys.com/GlobalWarming/ClimateData.htm

FIGURE 18 NOAA NCDC USHCN version 1 annual U.S. temperatures as of 1999.
When first implemented in 1990 as version 1, USHCN employed 1,221 stations across the United States. In 1999, NASA’s James Hansen published this graph of USHCN v.1 annual mean temperature (Fig. 18):

Hansen correctly noted:

“The US has warmed during the past century, but the warming hardly exceeds year-to-year variability. Indeed, in the US the warmest decade was the 1930s and the warmest year was 1934.”

USHCN was generally accepted as the world’s best database of temperatures. The stations were the most continuous and stable and had adjustments made for time of observation, urbanization, known station moves or land-use changes around sites, as well as instrumentation changes.

Note how well the original USHCN agreed with the state record high temperatures.

12. U.S. STATE HEAT RECORDS SUGGEST RECENT DECADES ARE NOT THE WARMEST

The 1930s were, by far, the hottest period for the time-frame. In absolute terms the 1930s had a much higher frequency of maximum temperature extremes than

![U.S. State Maximum and Minimum Monthly Records by Decade](image)

**FIGURE 19** United States all-time monthly record lows and highs by decade. Compiled by Hall from NOAA NCDC data.
the 1990s or 2000s or the combination of the last two decades. This was shown by Bruce Hall and Dr. Richard Keen, also covering Canada (Fig. 19).

NCDC’s Tom Karl (1988) employed an urban adjustment scheme for the first USHCN database (released in 1990). He noted that the national climate network formerly consisted of predominantly rural or small towns with populations below 25,000 (as of 1980 census) and yet that an urban heat-island effect was clearly evident.

Tom Karl et al.’s adjustments were smaller than Oke had found (0.22 °C annually on a town of 10,000 and 1.81 °C on a city of 1 million and 3.73 °C for a city of 5 million). Karl observed that in smaller towns and rural areas the net urban heat-island contamination was relatively small, but that significant anomalies showed up in rapidly growing population centers.

13. MAJOR CHANGES TO USHCN IN 2007

NOAA had to constantly explain why their global data sets which had no such adjustment was showing warming and the U.S., not so much. NOAA began reducing the UHI around 2000 (noticed by state climatologists and seen in this analysis of New York City’s Central Park data here http://icecap.us/index.php/go/new-and-cool/central_park_temperatures_still_a_mystery/) and then in USHCN version 2 released for the U.S. stations in 2009, the urban heat-island adjustment was eliminated which resulted in an increase of 0.3 °F in warming trend since the 1930s. See animating GIF here http://stevengoddard.files.wordpress.com/2010/12/1998uschanges.gif.

In 2007 the NCDC, in its version 2 of USHCN, inexplicably removed the Karl urban heat-island adjustment and substituted a change-point algorithm that looks for sudden shifts (discontinuities). This is best suited for finding site moves or local land-use changes (like paving a road or building next to sensors or shelters), but not the slow ramp up of temperature characteristic of a growing town or city (Fig. 20).

David Easterling, Chief of the Scientific Services Division at NOAA in one of the NASA FOIA emails noted: “One other fly in the ointment, we have a new adjustment scheme for USHCN (V2) that appears to adjust out some, if not most, of the “local” trend that includes land-use change and urban warming.”

The difference between the old and new is shown here. Note the significant post-1995 warming and mid-20th-century cooling owing to de-urbanization of the database (Fig. 21).

The change can be seen clearly in this animation and in ‘blink charts for Wisconsin’ and Illinois. Here are two example stations with USHCN version

---

31. http://climate-skeptic.typepad.com/a/6a00e54eeb9dc18834010535ef5d49970b-pi.
Figure 20  NOAA NCDC USHCN version 2 annual mean temperatures as of 2007.

Figure 21  NOAA NCDC USHCN version 2 minus version 1 annual mean temperatures.
1 and version 2 superimposed (thanks to Mike McMillan). Notice the clear tendency to cool off the early record and leave the current levels near recently reported levels or increase them. The net result is either reduced cooling or enhanced warming not found in the raw data (Fig. 22).

The new algorithms are supposed to correct for urbanization and changes in siting and instrumentation by recognizing sudden shifts in the temperatures (Fig. 23).

It should catch the kind of change shown above in Tahoe City, CA (Fig. 23). It is unlikely to catch the slow warming associated with the growth of cities and towns over many years, as in Sacramento, CA, in figure 24 above.

**FIGURE 22** NOAA USJCN version 1 vs. version 2 for Olney and Lincoln Illinois.

**FIGURE 23** Tahoe City, CA data and photos courtesy of Anthony Watts, surfacestations.org.
In a conversation during Anthony Watts invited presentation about the surface stations projects to NCDC, on 4/24/2008, he was briefed on USHCN2’s algorithms and how they operated by Matt Menne, lead author of the USHCN2 project. While Mr. Watts noted improvements in the algorithm can catch some previously undetected events like undocumented station moves, he also noted that the USHCN2 algorithm had no provision for long-term filtering of signals that can be induced by gradual local urbanization, or by long-term changes in the siting environment, such as weathering/coloring of shelters, or wind blocking due to growth of shrubbery/trees.

When Mr. Menne was asked by Mr. Watts if this lack of detection of such long-term changes was in fact a weakness of the USHCN algorithm, he replied “Yes, that is correct”. Essentially USHCN2 is a short period filter only, and cannot account for long-term changes to the temperature record, such as UHI, making such signals indistinguishable from the climate-change signal that is sought.

See some other examples of urban vs. nearby rural. Doug Hoyt, who worked at NOAA, NCAR, Sacramento Peak Observatory, the World Radiation Center, Research and Data Systems, and Raytheon where he was a Senior

Scientist did this analysis\textsuperscript{35} of the urban heat island. Read beyond the references for interesting further thoughts (Fig. 25).

Even before the version 2 shown above, Balling and Idso (2002)\textsuperscript{36} found that the adjustments being made to the raw USHCN temperature data were “producing a statistically significant, but spurious, warming trend” that “approximates the widely-publicized 0.50 °C increase in global temperatures over the past century”. There was actually a linear trend of progressive cooling of older dates between 1930 and 1995.

“It would thus appear that in this particular case of “adjustments,” the cure was much worse than the disease. In fact, it would appear that the cure may actually be the disease.”

It should be noted even with the changes to the USHCN, the correlations with CO\textsubscript{2} are intermittent with just 44 years warming while CO\textsubscript{2} increased and 62 years cooling even as CO\textsubscript{2} rose, not a convincing story for greenhouse CO\textsubscript{2} climate dominance at least with the U.S. data, even with all its warts, generally accepted as the most complete and stable data sets in the world (Fig. 26).

\textsuperscript{35} http://www.warwickhughes.com/hoyt/uhi.htm.

\textsuperscript{36} http://www.co2science.org/articles/V12/N50/C1.php.
14. HADLEY AND NOAA

No real urbanization adjustment is made for either NOAA’s or CRU’s global data. Jones et al. (1990: Hadley/CRU) concluded that urban heat-island bias in gridded data could be capped at 0.05 °C/century. Jones used data by Wang which Keenan37 has shown was fabricated. Peterson et al. (1998) agreed with the conclusions of Jones, Easterling et al. (1997) that urban effects on 20th century globally and hemispherically-averaged land air temperature time-series do not exceed about 0.05 °C from 1900 to 1990.

Peterson (2003) and Parker (2006) argue urban adjustment is not really necessary. Yet Oke (1973) showed a town of 1,000 could produce a 2.2 °C (3.4 °F warming). The UK Met Office (UKMO) has said38 future heat waves could be especially deadly in urban areas, where the temperatures could be 9 °C or more above today’s, according to the Met Office’s Vicky Pope. NASA summer land surface temperature of cities in the Northeast were an average of 7—9 °C (13—16 °F) warmer than surrounding rural areas over a three year period, new NASA research shows. It appears, the warmers want to have it both ways. They argue that the urban heat-island effect is insignificant, but also argue future heat waves will be most severe in the urban areas. This is especially incongruous given that greenhouse theory has the warming greatest in winters and at night.

The most recent exposition of CRU methodology is Brohan et al. (2006), which included an allowance of 0.1 °C/century for urban heat-island effects in the uncertainty but did not describe any adjustment to the reported average temperature. To make an urbanization assessment for all the stations used in the

HadCRUT data set would require suitable metadata (population, siting, location, instrumentation, etc.) for each station for the whole period since 1850. No such complete metadata are available.

The homepage for the NOAA temperature index\(^\text{39}\) cites Smith and Reynolds (2005) as authority. Smith and Reynolds in turn state that they use the same procedure as CRU: i.e., they make an allowance in the error-bars but do not correct the temperature index itself. The population of the world went from 1.5 to 6.7 billion in the 20th century, yet NOAA and CRU ignore population growth in the database with only a 0.05–0.1 °C uncertainty adjustment.

Steve McIntyre challenged Peterson (2003), who had said, “Contrary to generally accepted wisdom, no statistically significant impact of urbanization could be found in annual temperatures”\(^\text{40}\), by showing that the difference between urban and rural temperatures for Peterson’s station set was 0.7 °C and between temperatures in large cities and rural areas 2 °C. He has done the same for Parker (2006) (Fig. 27).\(^\text{41}\)

Runnalls and Oke (2006) concluded that:

“Gradual changes in the immediate environment over time, such as vegetation growth or encroachment by built features such as paths, roads, runways, fences, parking lots, and buildings into the vicinity of the instrument site, typically lead to trends in the series.”

\(^\text{41}\) http://climateaudit.org/2007/06/14/parker-2006-an-urban-myth/.
“Distinct régime transitions can be caused by seemingly minor instrument relocations (such as from one side of the airport to another or even within the same instrument enclosure) or due to vegetation clearance. This contradicts the view that only substantial station moves involving significant changes in elevation and/or exposure are detectable in temperature data.”

Numerous other peer-reviewed papers and other studies have found that the lack of adequate urban heat-island and local land-use change adjustments could account for up to half of all apparent warming in the terrestrial temperature record since 1900.

Siberia is one of the areas of greatest apparent warming in the record. Besides station dropout and a 10-fold increase in missing monthly data, numerous problems exist with prior temperatures in the Soviet era. City and town temperatures determined allocations for funds and fuel from the Supreme Soviet, so it is believed that cold temperatures were exaggerated in the past. This exaggeration in turn led to an apparent warming when more honest measurements began to be made. Anthony Watts has found that in many Russian towns and cities uninsulated heating pipes are in the open. Any sensors near these pipes would be affected. The pipes also contribute more waste heat to the city over a wide area.

The physical discomfort and danger to observers in extreme environments led to some estimations or fabrications being made in place of real observations, especially in the brutal Siberian winter. See this report. This was said to be true also in Canada along the DEW Line where radars were set up to detect incoming Soviet bombers during the Cold War.

McKitrick and Michaels (2004) gathered weather station records from 93 countries and regressed the spatial pattern of trends on a matrix of local climatic variables and socioeconomic indicators such as income, education, and energy use. Some of the non-climatic variables yielded significant coefficients, indicating a significant contamination of the temperature record by non-climatic influences, including poor data quality.

The two authors repeated the analysis on the IPCC gridded data covering the same locations. They found that approximately the same coefficients emerged. Though the discrepancies were smaller, many individual indicators remained significant. On this basis they were able to rule out the hypothesis that there are no significant non-climatic biases in the data. Both de Laat and Maurellis and McKitrick and Michaels concluded that non-climatic influences add up to a substantial warming bias in measured mean global surface temperature trends.

Ren et al. (2007), in the abstract of a paper on the urban heat-island effect in China, published in *Geophysical Research Letters*, noted that “annual and


seasonal urbanization-induced warming for the two periods at Beijing and Wuhan stations is also generally significant, with the annual urban warming accounting for about 65—80% of the overall warming in 1961—2000 and about 40—61% of the overall warming in 1981—2000.”

This result, along with the previous mentioned research results, indicates a need to pay more attention to the urbanization-induced bias that appears to exist in the current surface air temperature records.

Numerous recent studies show the effects of urban anthropogenic warming on local and regional temperatures in many diverse, even remote, locations. Jáuregui et al. (2005) discussed the UHI in Mexico, Torok et al. (2001) in southeast Australian cities. Block et al. (2004) showed effects across central Europe. Zhou et al. (2004) and He et al. (2005) across China, Velázquez-Lozada et al. (2006) across San Juan, Puerto Rico, and Hinkel et al. (2003) even in the village of Barrow, Alaska. In all cases, the warming was greatest at night and in higher latitudes, chiefly in winter.

Kalnay and Cai (2003) found regional differences in U.S. data but overall very little change and if anything a slight decrease in daily maximum temperatures for two separate 20-year periods (1980—1999 and 1960—1979), and a slight increase in night-time readings. They found these changes consistent with both urbanization and land-use changes from irrigation and agriculture.

Christy et al. (2006) showed that temperature trends in California’s Central Valley had significant nocturnal warming and daytime cooling over the period of record. The conclusion is that, as a result of increases in irrigated land, daytime temperatures are suppressed owing to evaporative cooling and night-time temperatures are warmed in part owing to increased heat capacity from water in soils and vegetation. Mahmood et al. (2006b) also found similar results for irrigated and non-irrigated areas of the Northern Great Plains.

Two Dutch meteorologists, Jos de Laat and Ahilleas Maurellis, showed in 2006 that climate models predict there should be no correlation between the spatial pattern of warming in climate data and the spatial pattern of industrial development. But they found that this correlation does exist and is statistically significant. They also concluded it adds a large upward bias to the measured global warming trend.

Ross McKitrick and Patrick Michaels in 2007 showed a strong correlation between urbanization indicators and the “urban adjusted” temperatures and that the adjustments are inadequate. Their conclusion: “Fully correcting the surface temperature data for non-climatic effects reduce the estimated 1980—2002 global average temperature trend over land by about half.”

As Pielke (2007) also notes:

“Changnon and Kunkel (2006) examined discontinuities in the weather records for Urbana, Illinois, a site with exceptional metadata and concurrent records when
important changes occurred. They identified a cooling of 0.17°C caused by a non-standard height shelter of 3 m from 1898 to 1948. After that there was a gradual warming of 0.9 °C as the University of Illinois campus grew around the site from 1900 to 1983. This was followed by an immediate 0.8 °C cooling when the site moved 2.2 km to a more rural setting in 1984. A 0.3 °C cooling took place with a shift in 1988 to Maximum-Minimum Temperature systems, which now represent over 60% of all USHCN stations. The experience at the Urbana site reflects the kind of subtle changes described by Runnalls and Oke (2006) and underscores the challenge of making adjustments to a gradually changing site.”

A 2008 paper by Hadley’s Jones et al., has shown a considerable contamination in China, amounting to 1 °C/century. This is an order of magnitude greater than the amount previously assumed (0.05–0.1 °C/century uncertainty).

In a 2009 article, Brian Stone of Georgia Tech wrote:

“Across the US as a whole, approximately 50 percent of the warming that has occurred since 1950 is due to land use changes (usually in the form of clearing forest for crops or cities) rather than to the emission of greenhouse gases. Most large US cities, including Atlanta, are warming at more than twice the rate of the planet as a whole. This is a rate that is mostly attributable to land use change.”

In a paper posted on SPPI, Dr. Edward Long summarized his findings as follows: both raw and adjusted data from the NCDC has been examined for a selected Contiguous U.S. set of rural and urban stations, 48 each or one per State. The raw data provides 0.13 and 0.79 °C/century temperature increase for the rural and urban environments (Figs. 28, 29).

One would expect the urban would be adjusted to match the uncontaminated rural data. Instead the rural is adjusted to look more like the urban with the warming since 1895 increased over half a degree from just 0.13 °C to 0.64 °C while the urban trend decreased an insignificant 0.02 °C (Fig. 30).

The adjusted data provide 0.64 and 0.77 °C/century respectively. Comparison of the adjusted data for the rural set to that of the raw data shows a systematic treatment that causes the rural adjusted set’s temperature rate of increase to be five-fold more than that of the raw data. This suggests the consequence of the NCDC’s protocol for adjusting the data is to cause historical data to take on the time-line characteristics of urban data. The consequence intended or not, is to report a false rate of temperature increase for the Contiguous U.S.

FIGURE 28  Edward Long analysis of rural raw stations for the lower 48 states, USHCN version 2. Note the very small trend 0.12 °C/century in this data set and at the significant peak in the 1930s.

FIGURE 29  Edward Long urban annual temperatures and trend from USHCN version 2 annual temperatures for the lower 48 states. Note the trend of 0.79 °C for this data set with the 1930 peak but with the second recent peak higher.
15. FINAL ADJUSTMENTS – HOMOGENIZATION

Dr. William Briggs in a 5 part series on the NOAA/NASA process of homogenization on his blog\(^{47}\) noted the following:

“At a loosely determined geographical spot over time, the data instrumentation might have changed, the locations of instruments could be different, there could be more than one source of data, or there could be other changes. The main point is that there are lots of pieces of data that some desire to stitch together to make one whole.

Why?

I mean that seriously. Why stitch the data together when it is perfectly useful if it is kept separate? By stitching, you introduce error, and if you aren’t careful to carry that error forward, the end result will be that you are far too certain of yourself. And that condition - unwarranted certainty - is where we find ourselves today.”

It has been said by NCDC in Menne et al. “On the reliability of the U.S. surface temperature record” (in press) and in the June 2009\(^{48}\) “Talking Points: related to Is the U.S. Surface Temperature Record Reliable?” that

\(^{47}\) http://wmbriggs.com/blog/?p=1459.

\(^{48}\) www.ncdc.noaa.gov/oa/about/response-v2.pdf.
station siting errors do not matter. However, the way NCDC conducted the analysis gives a false impression because of the homogenization process used.

Here’s a way to visualize the homogenization process. Think of it like measuring water pollution. Here’s a simple visual table of CRN station quality ratings and what they might look like as water pollution turbidity levels, rated as 1–5 from best to worst turbidity (Fig. 31):

In homogenization the data is weighted against the nearby neighbors within a radius. And so a station might start out as a “1” data wise, might end up getting polluted with the data of nearby stations and end up as a new value, say weighted at “2.5”. Even single stations can affect many other stations in the GISS and NOAA data homogenization methods carried out on U.S. surface temperature data (Fig. 32).49,50

In the map above, applying a homogenization smoothing, weighting stations by distance nearby the stations with question marks, what would you imagine the values (of turbidity) of them would be? And, how close would these two values be for the east coast station in question and the west coast station in question? Each would be closer to a smoothed center average value based on the neighboring stations.

Essentially, NCDC is comparing *homogenized data* to *homogenized data*, and thus there would not likely be any large difference between “good” and “bad” stations in that data. All the differences have been smoothed out by homogenization (pollution) from neighboring stations!

The best way to compare the effect of siting between groups of stations is to use the “raw” data, before it has passed through the multitude of adjustments that NCDC performs. However NCDC is apparently using homogenized data. So instead of comparing apples and oranges (poor sited vs. well sited stations) they essentially just compare apples (Granny Smith vs. Golden Delicious) of which there is little visual difference beyond a slight color change.

They cite 60 years of data in the graph they present, ignoring the warmer 1930s. They also use an early and incomplete surfacesations.org data set that was never intended for analysis in their rush to rebut the issues raised. However, our survey most certainly cannot account for changes to the station locations or station siting quality any further back than about 30 years. By NCDC’s own admission, (see Quality Control of pre-1948 Cooperative Observer Network Data\(^51\)) they have little or no metadata posted on station siting much further back than about 1948 on their MMS meta-database. Clearly, siting quality is dynamic over time.

The other issue about siting that NCDC does not address is that it is a significant contributor to extreme temperature records. By NOAA’s own admission in PCU6 — Unit No. 2 Factors Affecting the Accuracy and

Continuity of Climate Observations such siting issues as the rooftop weather station in Baltimore contributed many erroneous high temperature records, so many in fact that the station had to be closed.

NOAA wrote about the Baltimore station:

“A combination of the rooftop and downtown urban siting explain the regular occurrence of extremely warm temperatures. Compared to nearby ground-level instruments and nearby airports and surrounding COOPs, it is clear that a strong warm bias exists, partially because of the rooftop location.

Maximum and minimum temperatures are elevated, especially in the summer. The number of 80 plus minimum temperatures during the one-year of data overlap was 13 on the roof and zero at three surrounding LCD airports, the close by ground-based inner Baltimore harbor site, and all 10 COOPs in the same NCDC climate zone. Eighty-degree minimum are luckily, an extremely rare occurrence in the mid-Atlantic region at standard ground-based stations, urban or otherwise.”

Clearly, siting does matter, and siting errors have contributed to the temperature records of the United States, and likely the world GHCN network. Catching such issues isn’t always as easy as NOAA demonstrated in Baltimore (Fig. 33).

There is even some evidence that the change-point algorithm does not catch some site changes it should catch and that homogenization doesn’t help. Take, for example, Lampasas, Texas, as identified by Anthony Watts (Fig. 34).

The site at Lampasas, TX, moved close to a building and a street from a more appropriate grassy site after 2001. Note even with the GISS “homogeneity” adjustment (red) applied to the NOAA adjusted data, this artificial

warming remains although the old data (blue) is cooled to accentuate warming even further (Fig. 35).

The net result is to make the recent warm cycle maximum more important relative to the earlier maximum in the 1930s, and note the sudden warm blip after the station move remains.

Other examples (and there are many, many such examples) include (Fig. 36): Adjustments to the raw data are responsible for the New Zealand warming trend shown by NIWA, the National Institute of Water and Atmospheric

![Image of Lampasas, Texas relocated station](https://surfacestations.org)

**FIGURE 34** Lampasas, Texas relocated station (Photograph by Julie K. Stacy surfacestations.org).

**FIGURE 35** Lampasas, Texas relocated station before (blue) and after (red) homogenization. Note the cooling of the old data but no correction for the station move in 2001.
Research (NIWA). New Zealand Climate Science Coalition (NZCSC) publicly called on NIWA [link] to admit no valid statistical justification for its claims of a 0.91 °C rise in New Zealand’s average temperature last century for the Seven Station Series (7SS) (Fig. 37).

For the globe the final adjusted data set is then used to populate a global grid, interpolating up to 1200 km (745 miles) to grid boxes that had become now vacant by the elimination of stations.

The data are then used for estimating the global average temperature and anomaly and for initializing or validating climate models.

After the Menne et al. (2009) paper, NCDC recognized their position on station siting was untenable and requested $100 million to upgrade the siting of 1,000 climate stations in the 1,220 station network.

NASA/NOAA homogenization process has been shown to significantly alter the trends in many stations where the siting and rural nature suggest the

![Figure 36: GHCN raw vs. adjusted for Davis, CA and Auckland, NZ.](image)

![Figure 37: NIWA raw vs. adjusted for Seven Sisters Stations (7SS). Adjusted NIWA becomes GHCN raw.](image)
data are reliable. In fact, adjustments account for virtually all the trend in the data (multi-author paper accepted 2011).

**16. PROBLEMS WITH SEA SURFACE TEMPERATURE MEASUREMENTS**

The world is 71% ocean. The Hadley Centre only trusts data from British merchant ships, mainly plying northern hemisphere routes. Hadley has virtually no data from the southern hemisphere’s oceans, which cover four-fifths of the hemisphere’s surface. NOAA and NASA use ship data reconstructions. The gradual change from taking water in canvas buckets to taking it from engine intakes introduces uncertainties in temperature measurement. Different sampling levels will make results slightly different. How to adjust for this introduced difference and get a reliable data set has yet to be resolved adequately, especially since the transition occurred over many decades. The chart, taken from Kent (2007), shows how methods of ocean-temperature sampling have changed over the past 40 years (Fig. 38).

We have reanalysis data based on reconstructions from ships, from buoys (which also have problems with changing methodology) and, in recent decades, from satellites. The oceans offer some opportunity for mischief, as the emails released by the Climategate whistle-blower showed clearly.

This report53 analyzed climate model (Barnet et al., 2001) forecasts of ocean temperatures from 1955 to 2000 vs. actual changes. It found models greatly overstated the warming especially at the surface where the actual change was just about 0.1 °C over that period.

![FIGURE 38](http://www.worldclimatereport.com/archive/previous_issues/vol6/v6n16/feature1.htm)  

---

There is another data set that may better resolve this discrepancy with time, the ARGO buoys. The Argo network\(^{54}\) which may eventually overcome many of the prior problems, became operational in mid-2003.

Before Argo, starting in the early 1960s, ocean temperatures were measured with bathythermographs (XBTs). They are expendable probes fired into the water by a gun, that transmit data back along a thin wire. They were nearly all launched from ships along the main commercial shipping lanes, so geographical coverage of the world’s oceans was poor—for example the huge southern oceans were not monitored. XBTs do not go as deep as Argo floats, and their data are much less accurate (Met Office,\(^ {55}\) Argo\(^ {56}\)) (Fig. 39).

Early results showed a cooling, but some issues may exist with the quality control of the early measurements and the strong El Nino in 2009/10 produced a brief pop up now reversing. We believe in the future this data set may give us the best indication of ocean heat content which could be the most robust and reliable indication of climate trends (Pielke, 2008\(^ {57}\)).

17. LONG-TERM TRENDS

Just as the Medieval Warm Period was an obstacle to those trying to suggest that today’s temperature is exceptional, and the UN and its supporters tried to abolish it with the “hockey-stick” graph, the warmer temperatures in the 1930s and 1940s were another inconvenient fact that needed to be “fixed”.

In each of the databases, the land temperatures from that period were simply adjusted downward, making it look as though the rate of warming in the 20\(^{th}\)

---

century was higher than it was, and making it look as though today’s temperatures were unprecedented in at least 150 years (Figs. 40—44).

Wigley\textsuperscript{58} even went so far as to suggest that sea surface temperatures for the period should likewise be “corrected” downward by 0.15°C, making the 20\textsuperscript{th}-century warming trend look greater but still plausible. This is obvious data doctoring.

In the Climategate emails, Wigley also noted\textsuperscript{59}:

“Land warming since 1980 has been twice the ocean warming — and skeptics might claim that this proves that urban warming is real and important.”

NOAA, then, is squarely in the frame. First, the unexplained major station dropout with a bias towards warmth in remaining stations. Next, the removal of the urbanization adjustment and lack of oversight and quality control in the siting of new instrumentation in the United States database degrades what once was the world’s best data set, USHCNv1. Then, ignoring a large body of peer review research demonstrating the importance of urbanization and land-use changes NOAA chooses not to include any urban adjustment for the global data set, GHCN.


FIGURE 41  Comparing NASA GISS values for recent years as reported in the year shown. Note the shift down in 2007 after correction for the millennium bug identified by McIntyre and then the shift up again in 2009.

FIGURE 42  Comparing NASA GISS global values from 1980 to 2010. Note the string cooling prior to 1980. Warming post 1980 was due to many issues unaccounted for. Compare to UAH value for 2009.
As shown, these and other changes that have been made alter the historical record and mask cyclical changes that could be readily explained by natural factors like multi-decadal ocean and solar changes.60

FIGURE 43 Comparing NOAA GHCN version 2 raw vs. adjusted. A similar cooling in the early record and recent warming is clearly shown. Raw data show little warming from peak in 1930s to 1940s to 1990s and 2000s.

FIGURE 44 Central England Temperatures (CETs) for December from 1659 to 2010 second only 1890.

The CRU data have seen changes even in the last decade with a cooling of the early and middle parts of the 20th century and dramatic post 1990 warming when most of the issues discussed emerged (Fig. 40).

The green is the 2001 global temperature plot and the red that in 2010 (with data through 2009).

Is NASA in the clear? No. It works with the same GHCN/USHCN base data (plus the SCAR data from Antarctica). To its credit, as we have shown its U.S. database includes an urban adjustment that is reasonable, but as Steve McIntyre showed uses NASA used population data and adjusted GHCN temperature records for cities in a warming direction as often as they do in a cooling direction. This we have seen is due to very poor metadata from GHCN which GISS uses to match with satellite night light to define a station as urban, suburban or rural.

And their homogenization process and other non-documented final adjustments result in an increase in apparent warming, often by cooling the early record as can be seen in several case studies that follow.

NASA also seems to constantly rehash the surface data. John Goetz showed that 20% of the historical record was modified 16 times in the 2½ years ending in 2007. 1998 and 1934 ping pong regularly between first and second warmest year as the fiddling with old data continues.

In 2007, NASA adjusted post-2000 data when Steve McIntyre found a bug in the USHCN data down by 0.12 to 0.15C. Note how the data were adjusted up again in 2009 (USHCN V2.) (see Fig. 41)

Earlier version of NASA data was extracted from an earlier paper by Hansen in 1980 and is compared in the graph below. In 1987 (green on the graph in figure 42), the GISS temperatures were modified down in the middle part of the century from the 1980 version (blue) which enhanced the apparent warming in time for Dr. Hansen’s testimony in front of congress.

Cooling before 1980 is dramatic. Warming after 1990 was due to the myriad of issues with the data in this period as we have identified above.

E-mail messages obtained by a Freedom of Information Act request reveal that NASA concluded that its own climate findings were inferior to those maintained by both the University of East Anglia’s Climatic Research Unit (CRU) — the scandalized source of the leaked Climategate e-mails — and the National Oceanic and Atmospheric Administration’s National Climatic Data Center.

The e-mails from 2007 reveal that when a USA Today reporter asked if NASA’s data “was more accurate” than other climate-change data sets, NASA’s Dr. Reto A. Ruedy replied with an unequivocal no. He said “the National

---

Climatic Data Center’s procedure of only using the best stations is more accurate,” admitting that some of his own procedures led to less accurate readings.

“My recommendation to you is to continue using NCDC’s data for the U.S. means and [East Anglia] data for the global means,” Ruedy told the reporter.

A similar tale is seen with NOAA GHCN version 2 before and after adjustment (Fig. 43).

The longest history of unaltered data is the Central England Temperature set established during the Little Ice Age in 1659. Note how this past December 2010 was the second coldest December in the record (just 0.1C above 1890). Long-term warming is seen coming out of the LIA but no acceleration upwards can be detected. (Fig. 44)

18. SUMMARY

Climategate has sparked a flurry of examinations of the global data sets not only at CRU, NASA, and NOAA, but in various countries throughout the world. Though the Hadley Centre implied their data were in agreement with other data sets and was thus trustworthy, the truth is that other data centers and the individual countries involved were forced to work with degraded data and appear to be each involved in data manipulation.

Kevin Trenberth, IPCC Lead Author, NCAR and CRU associate said:

“It’s very clear we do not have a climate observing system...This may be a shock to many people who assume that we do know adequately what’s going on with the climate, but we don’t.”

Climate change is real. There has been localized warming due to population growth and land-use changes. There are cooling and warming periods that can be shown to correlate well with solar and ocean cycles. You can trust in the data that shows that there has been warming from 1979 to 1998, just as there was warming around 1920 to 1940. But there has been cooling from 1940 to the late 1970s and since 2001. The long-term trend on which this cyclical pattern is superimposed that is exaggerated.

As shown, record highs and rural temperatures in North America show the cyclical pattern but suggest the 1930s to 1940 peak was higher than the recent peak around 1998. Recent ranking was very likely exaggerated by the numerous data issues discussed. Given these data issues and the inconvenient truths in the Climategate emails, the claim that the 2010 was the warmest year and the 2000s was the warmest decade in the record or as some claim in a millennium or two is not credible.

These factors all lead to significant uncertainty and a tendency for over-estimation of century-scale temperature trends. An obvious conclusion from all findings above is that the global databases are seriously flawed and can no
longer be trusted to assess climate trends. And, consequently, such surface data should not be used for decision making.

We enthusiastically support Roger Pielke Sr. who, after exchanges with Phil Jones over data sets, called for\(^\text{65}\):

> “an inclusive assessment of the surface temperature record of CRU, GISS and NCDC. We need to focus on the science issues. This necessarily should involve all research investigators who are working on this topic, with formal assessments chaired and paneled by mutually agreed to climate scientists who do not have a vested interest in the outcome of the evaluations.”

Georgia Tech’s Dr. Judith Curry’s comments on Roger Pielke Jr.’s blog also support such an effort:

> “In my opinion, there needs to be a new independent effort to produce a global historical surface temperature dataset that is transparent and that includes expertise in statistics and computational science...The public has lost confidence in the data sets...Some efforts are underway in the blogosphere to examine the historical land surface data (e.g. such as GHCN), but even the GHCN data base has numerous inadequacies.”

Judith is part of the newly announced Berkeley Earth Surface Temperature (BEST) Project which aims to develop an independent analysis of the data from land stations, which would include many more stations than had been considered by the Global Historic Climatology Network. We trust they will include scientists who understand the issues we have raised and will make the reconstructed data sets available for independent review and analysis.

It should be noted that replication is required by the data quality act (DQA) according to the government’s own Office of Management and Budget (OMB). Though such an effort can be done locally through tedious research and analysis in the United States, the status of the publicly available global databases (GHCN, GISS, CRU) makes that extremely difficult or impossible currently. Until then, satellite data is the only trustworthy data set.

ACKNOWLEDGMENTS

I wish to thank Anthony Watts who provided invaluable analysis, and considerable constructive feedback and suggestions for this analysis. I wish to also thank Roger Pielke Sr., Steve McIntyre, E.M. Smith and Verity Jones and many others cited in this compilation study for their tireless efforts with regards to issues with temperature measurements.

REFERENCES


Pielke Sr., R.A., Matsui, T., 2005. Should light wind and windy nights have the same temperature trends at individual levels even if the boundary layer averaged heat content change is the same? Geophysical Research Letters 32 (No. 21), L21813. 10.1029/2005GL024407.


