

Global Surface Temperature Change

J. Hansen, R. Ruedy, M. Sato, and K. Lo

NASA Goddard Institute for Space Studies, New York, New York, USA

Abstract. We update the Goddard Institute for Space Studies (GISS) analysis of global surface temperature change, compare alternative analyses, and address questions about perception and reality of global warming. Satellite-observed nightlights are used to identify measurement stations located in extreme darkness and adjust temperature trends of urban and peri-urban stations for non-climatic factors, verifying that urban effects on analyzed global change are small. Because the GISS analysis combines available sea surface temperature records with meteorological station measurements, we test alternative choices for the ocean data, showing that global temperature change is sensitive to estimated temperature change in polar regions where observations are limited. We use simple 12-month (and $n \times 12$) running means to improve the information content in our temperature graphs. Contrary to a popular misconception, the rate of warming has not declined. Global temperature is rising as fast in the past decade as in the prior two decades, despite year-to-year fluctuations associated with the El Nino-La Nina cycle of tropical ocean temperature. Record high global 12-month running-mean temperature for the period with instrumental data was reached in 2010.

1. Introduction

Analyses of global surface temperature change are routinely carried out by several groups, including the NASA Goddard Institute for Space Studies, the NOAA National Climatic Data Center (NCDC), and a joint effort of the UK Met Office Hadley Centre and the University of East Anglia Climatic Research Unit (HadCRUT). These analyses are not independent, as they must use much the same input observations. However, the multiple analyses provide useful checks because they employ different ways of handling data problems such as incomplete spatial and temporal coverage and non-climatic influences on measurement station environment.

Here we describe the current GISS analysis of global surface temperature change. We first provide background on why and how the GISS method was developed and then describe the input data that go into our analysis. We discuss sources of uncertainty in the temperature records and provide some insight about the magnitude of the problems via alternative choices for input data and adjustments to the data. We discuss a few of the salient features in the resulting temperature reconstruction and compare our global mean temperature change with that obtained in the NCDC and HadCRUT analyses. Given our conclusion that global warming is continuing unabated, and that this conclusion differs from some popular perceptions, we discuss reasons for such misperceptions including the influence of short-term weather and climate fluctuations.

2. Background of GISS Analysis Method

GISS analyses of global surface temperature change were initiated by one of us (JH) in the late 1970s and first published in 1981 [*Hansen et al.*, 1981]. The objective was an estimate of global temperature change that could be compared with expected global climate change in response to known or suspected climate forcing mechanisms such as atmospheric carbon dioxide, volcanic aerosols, and solar irradiance changes. There was a history of prior analyses of temperature change, as discussed by *Jones et al.* [1982] and summarized in Figure 1.3 of the *IPCC* [2007] report, with most of the studies covering large but less than global regions.

A principal question at the time of the first GISS analysis was whether there were sufficient stations in the Southern Hemisphere to allow a meaningful evaluation of global temperature change. The supposition in the GISS analysis was that an estimate of global temperature change with useful accuracy should be possible because seasonal and annual temperature anomalies, relative to a long-term average (climatology), present a much smoother geographical field than temperature itself. For example, when New York City has an unusually cold winter, it is likely that Philadelphia is also colder than normal.

One consequence of working only with temperature change is that our analysis does not produce estimates of absolute temperature. For the sake of users who require an absolute global mean temperature, we have estimated the 1951-1980 global mean surface air temperature as 14°C with uncertainty several tenths of a degree Celsius. That value was obtained by using a global climate model [*Hansen et al.*, 2007] to fill in temperatures at grid points without observations, but it is consistent with results of *Jones et al.* [1999] based on observational data. The review paper of *Jones et al.* [1999] includes maps of absolute temperature as well as extensive background information on studies of both absolute temperature and surface temperature change.

The correlation of temperature anomaly time series for neighboring stations was illustrated by *Hansen and Lebedeff* [1987] as a function of station separation for different latitude bands. The average correlation coefficient was shown to remain above 50 percent to distances of about 1200 km at most latitudes, but in the tropics the correlation falls to about 35 percent at station separation of 1200 km. The GISS analysis specifies the temperature anomaly at a given location as the weighted average of the anomalies for all stations located within 1200 km of that point, with the weight decreasing linearly from unity for a station located at that point to zero for stations located 1200 km or further from the point in question.

Hansen and Lebedeff [1987] found that the correlation of neighboring station temperature records had no significant dependence on direction between the stations. They also examined the sensitivity of analyzed global temperature to the chosen limit for station radius of influence (1200 km). The global mean temperature anomaly was insensitive to this choice for the range from 250 to 2000 km. The main effect is to make the global temperature anomaly map smoother as the radius of influence increases. However, global maps of temperature anomalies using a small radius of influence are useful for detecting stations with a temperature record that is inconsistent with stations in neighboring regions. Thus the GISS web page provides results for 250 km as well as 1200 km.

The standard GISS analysis thus interpolates among station measurements and extrapolates anomalies as far as 1200 km into regions without measurement stations. Resulting regions with defined temperature anomalies are used to calculate a temperature anomaly history for large latitude zones. Early versions of the GISS analysis, including *Hansen and Lebedeff* [1987], calculated the global temperature anomaly time series as the average for these several latitude zones, with each zone weighted by the area with defined temperature anomaly. That definition can result in the global temperature anomaly differing from the average anomaly for the two hemispheres by as much as several hundredths of a degree during the early decades (1880-1920) when spatial coverage was especially poor. This characteristic caused confusion and questions from users of the GISS data set. Our present analysis weights the temperature anomalies of zones by the zone's full area, thus obtaining consistent zonal and global results. Differences among these arbitrary choices for definition of the global average anomaly are within the uncertainty in the global result discussed below.

Hansen and Lebedeff [1987] calculated an error estimate due to incomplete spatial coverage of stations using a global climate model that was shown to have realistic spatial and temporal variations of temperature anomalies. The average error was found by comparing global temperature variations from the spatially and temporally complete model fields with the results when the model was sampled only at locations and times with measurements. Calculated errors increased toward earlier times as the area covered by stations diminished, with the errors becoming comparable in magnitude to estimated global temperature changes at about 1880. Thus the GISS temperature anomaly estimates are restricted to post-1880. Independent analyses [*Karl et al.*, 1994; *Brohan et al.*, 2006; *Smith et al.*, 2008] of hemispheric temperature errors using observational data yield similar results, supporting the date at which meaningful global analyses become possible.

The GISS analysis uses 1951-1980 as the base period. The United States National Weather Service uses a three-decade period to define "normal" or average temperature. When we began our global temperature analyses and comparisons with climate models that climatology period was 1951-1980. There is considerable merit in keeping the base period fixed, including the fact that many graphs have been published with that choice for climatology. Besides, a different base period only alters the zero point for anomalies, without changing the magnitude of the temperature change over any given period. Note also that many of today's adults grew up during that period, so they can remember what climate was like then. Finally, the data for a base period must have good global coverage, which eliminates periods prior to the 1950s.

GISS analyses beginning with *Hansen et al.* [1999] include a homogeneity adjustment to minimize local (non-climatic) anthropogenic effects on measured temperature change. Such effects are usually largest in urban locations where buildings and energy use often cause a warming bias. Local anthropogenic cooling can also occur, for example from irrigation and planting of vegetation [*Oke*, 1989], but on average these effects are probably outweighed by urban warming. The homogeneity adjustment procedure [Figure 3 of *Hansen et al.*, 1999] changes the long-term temperature trend of an urban station to make it agree with the mean trend of nearby rural stations. The effect of this adjustment on global temperature change was found to be small, less than 0.1°C for the past century. Discrimination between urban and rural areas was based on the population of the city associated with the meteorological station. Location of stations relative to population centers varies, however, so in the present paper we use the intensity of high resolution satellite nighttime measurements to specify which stations are in population centers and which stations should be relatively free of urban influence.

The GISS temperature analysis has been available for many years on the GISS web site (www.nasa.giss.gov), including maps, graphs and tables of the results. The analysis is updated monthly using several data sets compiled by other groups from measurements at meteorological stations and satellite measurements of ocean surface temperature. Ocean data in the pre-satellite era is based on measurements by ships and buoys. The computer program that integrates these data sets into a global analysis is freely available on the GISS web site.

Here we describe the current GISS analysis and present several updated graphs and maps of global surface temperature change. We compare our results with those of HadCRUT and NCDC, the main purpose being to investigate differences in recent global temperature trends and the ranking of annual temperatures among different years.

3. Input Data

The current GISS analysis employs several independent input data streams that are publicly available on the internet and updated monthly. In addition the analysis requires a data

set for ocean surface temperature measurements in the pre-satellite era. In this paper and in the monthly updates on our web site, we now show results using alternative choices for pre-satellite ocean data and alternative procedures for concatenating satellite and pre-satellite data.

3.1. Meteorological Station Measurements

The source of monthly mean station measurements for our current analysis is the Global Historical Climatology Network (GHCN) version 2 of *Peterson and Vose* [1997], which is available monthly from NCDC. GHCN includes data from about 7000 stations. We use only those stations that have a period of overlap with neighboring stations (within 1200 km) of at least 20 years (see Figure 2 of *Hansen et al.*, 1999), which reduces the number of stations used in our analysis to about 6300. When GHCN version 3 becomes available, expected in late 2010, we will make results of our analysis available on our web site for both versions 2 and 3 for a period that is at least long enough to assess the effect of differences between the two versions.

The GISS, NCDC and HadCRUT global temperature results all make use of the GHCN data collection in their analyses. This is a matter of practicality, given the magnitude of the task of assembling and checking observations made in many different nations. However, a separate data collection effort that originated in the old Soviet Union led to another global temperature analysis, which does not use GHCN. This alternative analysis leads to a global temperature record in good agreement with that found by the groups using GHCN data [*Lugina et al.*, 2006]

We use the unadjusted version of GHCN. However, note that a subset of GHCN, the United States Historical Climatology Network (USHCN), has been adjusted via a homogenization intended to remove urban warming and other artifacts [*Karl et al.*, 1990; *Peterson and Vose*, 1997]. Also bad data in GHCN were minimized at NCDC [*Peterson and Vose*, 1997; *Peterson et al.*, 1998b] via checks of all monthly mean outliers that differed from their climatology by more than 2.5 standard deviations. About 15 percent of these outliers were eliminated for being incompatible with neighboring stations, with the remaining 85 percent being retained.

The current GISS analysis adjusts the long-term temperature trends of urban stations based on neighboring rural stations, and we correct discontinuities in the records of two specific stations as described below. Our standard urban adjustment now utilizes satellite observations of nightlights to identify whether stations are located in rural or urban areas. The urban adjustment, described below, is carried out via our published computer program and the publicly available nightlight data set.

Our analysis also continues to include specific adjustments for two stations, as described by *Hansen et al.* [1999]. The two stations have discontinuities that would cause artificial global warming without homogeneity adjustment. The stations, St. Helena in the tropical Atlantic Ocean and Lihue, Kauai, in Hawaii, are on islands with few if any neighboring stations, so their records have noticeable impact on analyzed regional temperature change. The St. Helena station, based on metadata in MCDW records, was moved from elevation 604 m to 436 m between August 1976 and September 1976. Thus, assuming a temperature lapse rate of 6°C/km, we added 1°C to St. Helena temperatures before September 1976. Lihue had an obvious discontinuity in its temperature record around 1950. On the basis of minimizing the discrepancy with its few neighboring stations, we added 0.8°C to Lihue temperatures prior to 1950.

Our procedure in the current monthly updates of the GISS analysis when we find what seems to be a likely error in a station record is to report the problem to NCDC for their consideration and possible correction of the GHCN record. Our rationale is that verification of correct data entry from the original meteorological source is a person-intensive activity that is

best handled by NCDC with its existing communications network. Also it seems better not to have multiple versions of the GHCN data set in the scientific community.

The contiguous United States presents special homogeneity problems. One problem is the bias introduced by change in the time of daily temperature recording [Karl *et al.*, 1986], a problem that does not exist in the temperature records from most other nations. High energy use, built up local environments, and land use changes also cause homogeneity problems [Karl *et al.*, 1990]. The adjustments included by NCDC in the current USHCN data set [version 2, Menne *et al.*, 2009] should reduce these problems. As a test of how well urban influences have been minimized, we illustrate below the effect of our nightlight-based urban adjustment on the current USHCN data set.

3.2. Antarctic Research Station Measurements

Measurements at Antarctic research stations help fill in what would otherwise be a large hole in the GHCN land-based temperature record. Substantial continuous data coverage in Antarctica did not begin until the International Geophysical Year (1957). However, the period since 1957 includes the time of rapid global temperature change that began in about 1980.

The GISS analysis uses Scientific Committee on Antarctic Research (SCAR) monthly data [Turner *et al.*, 2004], which are publicly available. Specifically, the data are from the SCAR READER (Reference Antarctic Data for Environmental Research) project (<http://www.antarctica.ac.uk/met/READER/>).

3.3. Ocean Surface Temperature Measurements

Our standard global land-ocean temperature index uses a concatenation of the Met Office Hadley Centre analysis of sea surface temperatures (HadISST1) [Rayner *et al.*, 2003] for 1880-1981, which is ship-based during that interval, and satellite measurements of sea surface temperature for 1982-present (OISST.v2) [Reynolds *et al.*, 2002]. The satellite measurements are calibrated with the help of ship and buoy data [Reynolds *et al.*, 2002].

Ocean surface temperatures have their own homogeneity issues. Measurement methods changed over time as ships changed, most notably with a change from measurements of bucket water to engine intake water. Homogeneity adjustments have been made to the ship-based record [Parker *et al.*, 1995; Rayner *et al.*, 2003], but these are necessarily imperfect. The spatial coverage of ship data is poor in the early 20th century and before. Also, although the land-based data confirm the reality of an early 1940s peak in global temperature, it has been suggested that changes in ship measurements during and after World War II cause an instrumental artifact that contributes to the magnitude of the peak of ocean temperature in the HadISST1 data set in the early 1940s [Thompson *et al.*, 2008]. Ocean coverage and the quality of sea surface temperature data have been better since 1950 and especially during the era of satellite ocean data, i.e., since 1982. Satellite data, however, also have their own sources of uncertainty, despite their high spatial resolution and broad geographical coverage.

Thus we compare the global temperature change obtained in our standard analysis, which concatenates HadISST1 and OISST.v2, with results from our analysis program using alternative ocean data sets. Specifically, we compare our standard case with results when the ocean data are replaced by the Extended Reconstructed SST (ERSST.v3b) [Smith *et al.*, 2008] for the full period 1880-present, and also when HadISST1 is used as the only ocean data for the full period 1880-present. In the Appendix we compare the alternative ocean data sets themselves over

regions of common data coverage to help isolate differences among input SSTs as opposed to differences in the area covered by ocean data.

We use ocean temperature change only in regions that are ice-free all year (a map of this area is included in the Appendix), because our data set is intended to be temperature change of surface air. Surface air temperature (SAT), measured at heights of 1.25-2 meters at meteorological stations, is of most practical significance to humans and it is usually SAT change that is reported in climate model studies. Change of sea surface temperature (SST) should be a good approximation to change of SAT in ice-free ocean areas; climate model simulations [Hansen *et al.*, 2007] suggest that long-term SAT change over ice-free ocean is only slightly larger than SST change. However, ocean water temperature does not go below the freezing point of water, while surface air temperature over sea ice can be much colder. As a result SST change underestimates SAT change when sea ice cover changes. Indeed, most climate models find that the largest SAT changes with global warming occur in regions of sea ice [IPCC, 2007]. Thus we estimate SAT changes in sea ice regions by extrapolating actual SAT measurements on nearby land or islands; if there are no stations within 1200 km we leave the temperature change undefined.

4. Urban Adjustments

A major concern about the accuracy of analyses of global temperature change has long been the fact that many of the station locations are located in or near urban areas. Human-made structures and energy sources can cause a substantial local warming that affects measurements in the urban environment. This local warming must be eliminated to obtain a valid measure of global climate change. Global temperature analyses now routinely either omit urban stations or adjust their long-term trends to try to eliminate or minimize the urban effect. A comprehensive review of the topic is provided by Parker [2010].

The urban influence on long-term global temperature change is generally found to be small. It is possible that the overall small urban effect is in part a consequence of partial cancellation of urban warming and urban cooling effects. A significant urban cooling can occur, for example, if a station is moved from central city to an airport and if the new station continues to be reported with the same station number and is not treated properly as a separate station in the global analysis.

Global satellite measurements of nightlights allow the possibility for an additional check on the magnitude of the urban influence on global temperature analyses. We describe in this section a procedure in which all stations located in areas with nightlight brightness exceeding a value ($32 \mu\text{W}/\text{m}^2/\text{sr}/\mu\text{m}$) that approximately divides the stations into two categories: rural and urban or peri-urban [Imhoff *et al.*, 1997]. The standard GISS global temperature analysis now adjusts the long-term trends of stations located in regions with nightlight brightness exceeding this limit to agree with the long-term trend of nearby rural stations. If there are not a sufficient number of nearby rural stations, the "bright" station is excluded from the analysis.

We present evidence here that the urban warming has little effect on our standard global temperature analysis. However, in the Appendix we carry out an even more rigorous test. We show there that there are a sufficient number of stations located in "pitch black" regions, i.e., regions with brightness below the satellite's detectability limit ($\sim 1 \mu\text{W}/\text{m}^2/\text{sr}/\mu\text{m}$), to allow global analysis with only the stations in pitch black regions defining long-term trends. The effect of this more stringent definition of rural areas on analyzed global temperature change is immeasurably small ($<0.01^\circ\text{C}$ per century). The finding of a negligible effect in this test (using only stations in pitch black areas) also addresses, to a substantial degree, the question of whether

movement of weather stations to airports has an important effect on analyzed global temperature change. The pitch black requirement eliminates not only urban and peri-urban stations but also three-quarters of the stations in the more than 500 GHCN records that are identified as airports in the station name. (The fact that one-quarter of the airports are pitch black suggests that they are in extreme rural areas and are shut down during the night.)

Station location in the meteorological data records is provided with a resolution of 0.01 degrees of latitude and longitude, corresponding to a distance of about 1 km. This resolution is useful for investigating urban effects on regional atmospheric temperature. Much higher resolution would be needed to check for local problems with the placement of thermometers relative to possible building obstructions, for example. In many cases such local problems are handled via site inspections and reported in the "metadata" that accompanies station records, as discussed by *Karl and Williams* [1987], *Karl et al.* [1989], and *Peterson et al.* [1998a].

Of course the thousands of meteorological station records include many with uncorrected problems. The effect of these problems tends to be reduced by the fact that they include errors of both signs. Also problems are usually greater in urban environments or associated with movement of stations from urban areas. An urban adjustment based on nightlights should tend to reduce the effect of otherwise unnoticed non-climatic effects, and our check that the use of only pitch black stations to set long-term temperature trends (see Appendix) yields a practically identical result with our standard analysis adds further confidence in the result.

We use a nightlight radiance data set [*Imhoff et al.*, 1997] that is publicly available [http://www.ngdc.noaa.gov/dmsp/download_rad_cal_96-97.html] (measurements made between March 1996 and February 1997)] at a resolution of $30'' \times 30''$ ($0.0083^\circ \times 0.0083^\circ$), which is a linear scale of about 1 km. In Figure 1(a) the global radiances are shown after averaging to $0.5^\circ \times 0.5^\circ$ resolution, i.e., a pixel (resolution element) in Figure 1(a) is an average over 3600 high resolution pixels. This averaging reduces the maximum radiance from about $3000 \mu\text{W}/\text{m}^2/\text{sr}/\mu\text{m}$ to about $670 \mu\text{W}/\text{m}^2/\text{sr}/\mu\text{m}$.

Imhoff et al. [1997] investigated the relation between nightlight radiances and population density in the United States. We find that radiances $< 32 \mu\text{W}/\text{m}^2/\text{sr}/\mu\text{m}$ correspond well with the "unlit" or "dark" category of *Imhoff et al.* [1997] within the United States, as can be seen by comparing Figure 1(b) here with Plate 1 of *Hansen et al.* [2001]. The "unlit" regions, according to *Imhoff et al.* [1997], correspond to populations densities of about 0.1 persons/ha or less in the United States.

The relation between population and nightlight radiance in the United States is not valid in the rest of the world, as energy use per capita is higher in the United States than in most countries. However, energy use is probably a better metric than population for estimating urban influence, so we employ nightlight radiance of $32 \mu\text{W}/\text{m}^2/\text{sr}/\mu\text{m}$ as the dividing point between rural and urban areas in our global nightlight test of urban effects. Below we show, using data for regions with very dense station coverage, that use of a much more stringent criterion for darkness does not significantly alter the results.

We first compare two alternatives for the urban correction. One case uses the definition of rural stations used by *Hansen et al.* [1999], i.e., stations associated with towns of population less than 10,000 (population data available at <ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/v2>). The second case defines rural stations as those located in a region with nightlight radiance less than $32 \mu\text{W}/\text{m}^2/\text{sr}/\mu\text{m}$. This nightlight criterion is stricter than the population criterion in the United States, i.e., many sites classified as rural based on population below 10,000 are classified as urban based on nightlight brightness, as shown by the fact that some of the blue dots (towns of population below 10,000) in Figure 1c fall within the urban areas defined by nightlights (yellow

area). However, as we will see, the opposite is true in places such as Africa, i.e., a population criterion of less than 10,000 results in fewer rural stations than the nightlight criterion.

Our adjustment of urban station records uses nearby rural stations to define the long-term trends while allowing the local urban station to define high frequency variations, nominally as described by *Hansen et al.* [1999], but with details as follows. The reason to retain the urban record is to increase the sample for high frequency data and because urban stations often extend a bit further toward earlier times. The adjustment to the urban record is a two-segment broken line that approximates the difference between the averaged regional rural record and the urban record. It is computed as in *Hansen et al.* [1999], but the knee of the broken line is variable (rather than being fixed at 1950), chosen so as to optimize the approximation. If there are not at least three rural stations to define the regional trend, the urban station is dropped.

Figure 2 shows the resulting temperature change over the period 1900-2009 for urban adjustment based on nightlights, urban adjustment based on population, and no urban adjustment. The effect of urban adjustment on global temperature change is only of the order of 0.01°C for either nightlight or population adjustment. The small magnitude of the urban effect is consistent with results found by others [*Karl et al.*, 1988; *Jones et al.*, 1990; *Peterson et al.*, 1999; *Peterson*, 2003; *Parker*, 2004]. Our additional check is useful, however, because of the simple reproducible way that nightlights define rural areas. We previously used this nightlight method [*Hansen et al.*, 2001], but only for the contiguous United States.

The most noticeable effect of the urban adjustment in Figure 2 is in Africa and includes changes of both signs. The large local changes are not due to addition of an urban correction to specific stations, but rather to the deletion of urban stations because of the absence of three rural neighbors. African station records are especially sparse and unreliable [*Peterson et al.*, 1998b; *Christy et al.*, 2009]. Thus the large local temperature changes between one adjustment and another may have more to do with variations in station reliability rather than urban warming. Given the small number of long station records in Africa and South America [*Peterson et al.*, 1997], a single bad station record can affect a noticeable area in a map of temperature change. However, even though the paucity of station records in Africa and South America makes the results less reliable there, the general consistency between land and nearby ocean temperature patterns suggests that the principal trends in Africa and South America are real.

Figure 3 compares the global mean temperature versus time for the two alternative urban adjustments and no urban adjustment. The main conclusion to be drawn is that the differences among the three curves are small. Nevertheless, we know that the adjustment is substantial for some urban stations, so it is appropriate to include an urban adjustment.

How can we judge whether the nightlight or population adjustment is better in the sense of yielding the most realistic result? One criterion might be based on which one yields more realistic continuous meteorological patterns for temperature anomaly patterns. Nightlights arguably do very slightly better based on that criterion (Figure 2). Independently, we expect nightlight intensity to be a better indication than population of urban heat generation. Nightlights also preserve a greater area with defined temperature anomalies (Figure 2). Finally, stations can be more accurately associated with nightlight intensity (within about 1km) than with population, and the nightlight data are easily accessible, so anyone can check our analysis.

For these reasons, beginning in January 2010 the standard GISS analysis employs global nightlights in choosing stations to be adjusted for urban effects. Use of nightlights is a well-defined objective approach for urban adjustment, and the nightlight data set that we use [http://www.ngdc.noaa.gov/dmsp/download_rad_cal_96-97.html] is readily available. In the future, as additional urban areas develop, it will be useful to employ newer satellite measurements.

The small urban correction is somewhat surprising, even though it is consistent with prior studies. We know, for example, that urban effects of several degrees exist in some cities such as Tokyo, Japan and Phoenix, Arizona, as illustrated in Figure 3 of *Hansen et al.* [1999]. Although such stations are adjusted in the GISS analysis, is it possible that our "rural" stations themselves contain substantial human-made warming? There is at least one region, the United States, where we can do a stricter test of urban warming, because of the high density of meteorological stations. The United States is a good place to search for greater urban effects, because of its high energy use and a consequent expectation of large urban effects.

In Figures 4 and 5 we compare no adjustment, population adjustment, and two nightlight adjustments. The standard nightlight adjustment defines rural stations as those with nightlight radiance $<32 \mu\text{W}/\text{m}^2/\text{sr}/\mu\text{m}$. The strict nightlight adjustment defines rural stations as those that are pitch black (radiance $<1 \mu\text{W}/\text{m}^2/\text{sr}/\mu\text{m}$). There are about 300 pitch black stations in the contiguous United States (Figure 1b), sufficient to yield a filled-in United States temperature anomaly map even with the station radius of influence set at 250 km. We use the 250 km radius of influence to provide higher resolution in Figure 4, compared with 1200 km radius of influence, and we exclude smoothing from the plotting package so that results are shown at the $2^\circ \times 2^\circ$ resolution of the calculation, thus allowing more quantitative inspection.

The largest urban adjustment is in the Southwest United States, where a warming bias is removed. In a few locations the adjustment yields greater warming, which can result from either the spatial smoothing inherent in adjusting local trends to match several neighboring stations or neighboring rural stations that have greater warming than the urban station. The standard nightlight adjustment removes slightly more warming than does the population adjustment. The important conclusion is that the strict nightlight adjustment has no significant additional effect, compared with the standard nightlight adjustment. The 1900-2009 temperature change over the contiguous United States, based on linear fit to the data in Figure 5, is 0.70°C (no adjustment), 0.64°C (population adjustment), 0.63°C (standard nightlight adjustment), and 0.64°C (strict nightlight adjustment). If only USHCN stations are employed (as in NCDC analyses) we find a 1900-2009 temperature change of 0.73°C (no adjustment) and 0.65°C (standard nightlight adjustment).

In the Appendix we show that pitch dark stations have sufficient global distribution to allow comparison of temperature change analysis pegged to dark stations (our standard analysis) with analysis pegged to pitch dark stations. Results are practically identical, offering further evidence that urban effects have negligible effect on our global analysis.

Global temperature change in this paper, unless indicated otherwise, is based on the standard nightlight adjustment. We conclude, based on results reported here and the other papers we referenced, that unaccounted for urban effects on global temperature change are small in comparison to the $\sim 0.8^\circ\text{C}$ global warming of the past century. Extensive confirmatory evidence (such as glacier retreat and borehole temperature profiles) is provided by *IPCC* [2007].

5. Alternative Ocean Data Sets

Yasunaka and Hanawa [2010] compare several datasets for sea surface temperature. They find reasonably good agreement among the data sets, ascribing differences to the use of different methods to interpolate between measurements, instrumental bias correction methods, treatments of satellite-derived temperatures, and other factors. The Global Climate Observing System (GCOS) SST Working Group has an intercomparison web site (<http://ghrsst.nodc.noaa.gov/intercomp.html>) that allows comparison of SST data sets. A peer-reviewed Community White Paper [*Rayner et al.*, <http://www.oceanobs09.net/blog/?p=121>)]

discusses changing biases in the SST data sets and some reasons for differences among them. A paper describing the differences in greater detail and recommending adjustments is in preparation [Kennedy, J., manuscript in preparation].

Here we illustrate how our analysis of global temperature change is altered by the choice of SST data set. Figure 6 compares global temperature change for three choices of ocean surface temperature: (1) HadISST1 [Rayner *et al.*, 2003] for 1880-1981 and OISST.v2 [Reynolds *et al.*, 2002] for 1982-present, (2) ERSST.v3b [Smith *et al.*, 2008] for the full period 1880-present, (3) HadISST1 for 1880-2008. In all three cases the land data are based on the GISS analysis of GHCN and Antarctic (SCAR) data.

The first of the ocean data sets, the combination HadISST1+OISST.v2, is used in the GISS analysis for our standard land-ocean temperature index. Results based on HadISST1 alone and HadISST1+OISST.v2 are in close agreement in the 1982-present period during which they might differ (Figure 6b). We use HadISST1 with concatenated OISST.v2 for our monthly updates because OISST is available in near real time.

ERSST, a newer SST analysis covering the period 1880 to the present, is also available in near real time. SST values in data sparse regions in ERSST are filled in by NCDC using statistical methods, dividing the SST anomaly patterns into low-frequency (decadal scale) anomalies, by averaging and filtering available data points, and high-frequency residual anomalies [Smith *et al.*, 2008]. The concept is that the SST reconstruction may be improved by constraining temperature anomaly fields toward realistic modes of variability. (HadISST1 also uses a reconstruction approach, while HadSST2 defines temperature only in gridboxes with actual observations.)

A newer Met Office Hadley Centre SST data set, HadSST2 (Rayner *et al.*, 2006) has cooler SSTs in 1908-1912, comparable to the lower temperatures in ERSST. Both HadSST2 and ERSST use newer versions of the International Comprehensive Ocean-Atmosphere Data Set [ICOADS; Worley *et al.*, 2005]. Presumably the newer ICOADS data set is superior, as it is based on more data with better geographical coverage. However, the HadSST2 resolution ($5^{\circ}\times 5^{\circ}$) is too crude for our purposes. HadISST1 and ERSST resolutions are $1^{\circ}\times 1^{\circ}$ and $2^{\circ}\times 2^{\circ}$.

The global warming trend in ERSST is intermediate between the trends in HadISST1 and HadSST2. ERSST yields 0.04°C greater global warming (based on linear trends) than HadISST1 or HadISST1+OISST over the period 1880-2009 (Figure 6a). Over 1980-2009 ERSST yields global warming about 0.03°C greater than either HadISST1 or HadISST1+OISST (Figure 6b). The warming trend in HadSST2 exceeds the warming in ERSST by about as much as the ERSST warming exceeds that in HadISST1, as illustrated in the Appendix.

Figure 7 illustrates the geographical distribution of the differences between ERSST and the ocean data used in the current standard GISS analysis (HadISST1+OISST). It is apparent that the greater warming in ERSST on the century time scale occurs primarily in the South Pacific and South Atlantic oceans. The difference between the two reconstructions is large enough in some parts of the Pacific Ocean, of the order of 1°C , that it may be possible to discriminate between them based on focused examination of all available data for those regions. Even proxy temperature measures, e.g., from corals [Beck *et al.*, 1992], have the potential to resolve differences of that magnitude.

In recent years, as shown for 2000-2008 in the lower right part of Figure 7, the greater warming in ERSST occurs especially in the Eastern Pacific Ocean. Both data sets include use of in situ data, so it is unclear why the difference is large in this region.

HadISST1+OISST has been used in the standard GISS global temperature analysis for the past decade. However, we now also make a second analysis, ERSST+OI, using a new

procedure for concatenating satellite data with pre-satellite era data. This new procedure is designed to incorporate (in the satellite era) the merits of both the high resolution satellite data and the greater stability of modern in situ measurements.

The new concatenation procedure allows the in situ data to appropriately constrain the long-term temperature trend, but the full information from the high spatial resolution satellite data is also incorporated. The merit of these characteristics is suggested by the following facts: (1) the satellite SST measurements have a bias in their trend because of effects of aerosols and clouds that cannot be fully removed [Reynolds *et al.*, 2002, 2010], (2) the satellite data are much superior in both spatial coverage and spatial resolution. The value of the latter characteristic of the satellite data has been shown by Hughes *et al.* [2009], who found that HadISST1+OISST provides much better agreement than ERSST with high temporal resolution in situ temperature measurements made at research sites in the North Atlantic Ocean.

Our new concatenation procedure compares the temperature anomalies in ERSST and OISST each month at all gridboxes containing in situ measurements (as judged by the HadSST2 uninterpolated gridded data). The OISST anomaly is then adjusted by the difference between the ERSST and OISST anomalies averaged over the area containing in situ measurements (the adjustment is a single global number each month). This adjusted OISST data thus provides a high spatial resolution SST data set from which the small cold bias described by Reynolds *et al.*, [2002, 2010] has been removed in global mean.

We anticipate that this new concatenation procedure yields a better data product for the satellite era (1992-present) than either unadjusted OISST data or SST data sets reconstructed from only in situ measurements (reconstruction refers to methods, such as employed by ERSST and HadISST1, in which a complete ocean SST data set is defined from irregularly spaced in situ data by constraining temperature anomaly fields toward realistic modes of variability). Validation with high-frequency in situ observations at many locations, such as the data at research stations in North Atlantic Ocean [Hughes *et al.*, 2009], is needed to judge the merits of the resulting data set.

Concerning the pre-satellite era, it is not clear which data set is most useful for our purposes. None of the available data sets corrects as yet for the possible data inhomogeneity at the end of World War II [Thompson *et al.*, 2008], a matter that will need assessment when proposed adjustments have been made available.

Until there is a demonstrably superior ocean data set, we will retain HadISST1+OISST (concatenated as in our analyses for the past several years) in our standard analysis. However, we also make available ERSST+OI with the new concatenation procedure. As expected, the global mean temperature trend using ERSST+OI with the new concatenation is practically the same as obtained with ERSST employed for the entire period 1880-present (Figure 6).

We compare alternative ocean data sets in more detail in the Appendix. Note that differences among the data sets are less than uncertainties estimated by the data providers. Also note that the differences are small enough that the choice of ocean data set does not alter the conclusions drawn in this paper about global temperature change.

6. Current GISS Surface Temperature Analysis

The results in this section use HadISST1+OISST with the switch to OI in 1982. A smooth concatenation is achieved by making the 1982-1992 OISST mean anomaly at each gridbox identical to the HadISST1 1982-1992 mean for that gridbox. Results with our ERSST+OI analysis are very similar, but with slightly larger global warming. Both results are available on our website.

Figure 8 shows the global surface temperature anomalies for the past four decades, relative to the 1951-1980 base period. On average, successive decades warmed by 0.17°C. The warming of the 1990s (0.13°C relative to the 1980s) was reduced by the temporary effect of the 1991 Mount Pinatubo volcanic eruption. (El Chichon, in the prior decade, produced a global average aerosol optical depth only about half as large as Pinatubo.)

Warming in these recent decades is larger over land than over ocean, as expected for a forced climate change [*Hansen et al.*, 2007; *IPCC*, 2007; *Sutton et al.*, 2007], in part because the ocean responds more slowly than the land due to the ocean's large thermal inertia. Warming during the past decade is enhanced, relative to the global mean warming, by about 50 percent in the United States, a factor of 2-3 in Eurasia, and a factor of 3-4 in the Arctic and the Antarctic Peninsula.

Warming of the ocean surface has been largest over the Arctic Ocean, second largest over the Indian and Western Pacific Oceans, and third largest over most of the Atlantic Ocean. Temperature changes have been small and variable in sign over the North Pacific Ocean, the Southern Ocean, and the regions of upwelling off the west coast of South America.

Figure 9a updates the GISS global annual and 5-year mean temperatures through 2009. Results differ slightly from our prior papers because of our present use of the global nightlights to adjust for urban effects, but the changes are practically imperceptible. The nightlight adjustment reduces the 1880-2009 global temperature change by an insignificant 0.004°C relative to the prior population-based urban adjustment.

Global temperature in the past decade was about 0.8°C warmer than at the beginning of the 20th century (1880-1920 mean). Two-thirds of the warming has occurred since 1975.

Figure 9a has become popular, eagerly awaited by some members of the public and the media. An analogous graph, often in the form of a histogram is made available each year by the Met Office Hadley Centre/University of East Anglia Climatic Research Unit and by the NOAA National Climatic Data Center.

We suggest, however, that a more informative and convenient graph is the simple 12-month running mean global temperature (Figure 9b). From a climate standpoint there is nothing special about the time of year at which the calendar begins. The 12-month running mean removes the seasonal cycle just as well at any time of year. Note that the use of temperature anomalies itself does not fully remove the seasonal cycle: substantial seasonality remains, with global temperature variability and average temperature trends larger in Northern Hemisphere winter than in Northern Hemisphere summer. This residual seasonality is removed via the 12-month (or $n \times 12$ month) running mean.

The 12-month running mean temperature anomaly (Figure 9b) provides an improved measure of the strength and duration of El Ninos, La Ninas, and the response to volcanic eruptions. In contrast, use of the calendar year, as in Figure 9a, can be misleading, because one El Nino may coincide well with a calendar year while another is split between two calendar years. The 12-month running mean also provides a better measure of the longevity of an event (a positive or negative temperature excursion).

We note that more sophisticated filters are useful for specific purposes. Also, if our flat (equal weight) running mean were used for frequency analyses it could have undesirable effects. However, there is no need for it to be used for that purpose, as full resolution data are available from our web site. We suggest only that the simple 12-month running mean provides a powerful and intuitive visual presentation that usefully complements the usual annual-mean graph, and it is probably the best filter for use in communication with the public. But we do not discourage others from using our full resolution data with more sophisticated filters and spectral analyses,

despite the shortness of the record, its non-uniformity over time, and the sorry record of spectral studies for such a short period in revealing anything new about the physical system.

A clearer view of temporal variations is provided by Figure 10, which covers the shorter period 1950-2010. The top curve is the monthly mean global temperature anomaly and the second curve is the 12-month running mean. The red-blue Nino index is the 12-month running mean of the temperature anomaly (relative to 1951-1980) averaged over the Nino 3.4 area (longitudes 120W-170W, latitudes 5N-5S) in the Eastern Pacific Ocean [Philander, 2006]. Because the monthly Nino index is much smoother than monthly global temperature, we can usefully extend the Nino index to the present. The final five points in the Nino 3.4 curve are 10, 8, 6, 4 and 2-month means.

The well-known strong correlation of global surface temperature with the Nino index is apparent in Figure 10. The correlation is maximum (at 30%) with 12-month running mean global temperature lagging the Nino index by 4 months. Global cooling due to large volcanoes in 1963 (Agung), 1982 (El Chichon) and 1991 (Pinatubo) is also apparent.

The data in Figures 9 and 10, and knowledge that tropical SSTs are now (July 2010) moving into La Nina conditions, permit several conclusions. The 12-month running mean global temperature in 2010 has reached a new record level for the period of instrumental data. It is likely that the 12-month mean will begin to decline in the second half of 2010. The subsequent minimum in the 12-month running mean is likely to be in 2011-2012 and not as deep as the 2008 minimum. The next maximum, likely to be in 2012-2014, will probably bring a new record global temperature, because of the underlying warming trend.

As for calendar year 2010, the first half of the year is warm enough that the 2010 global surface temperature in the GISS analysis likely will be a record for the period of instrumental data, or at least so close to the prior warmest year (2005) that it must be declared in a statistical dead heat. It is still conceivable that record global temperature for the calendar year will not occur, if tropical SSTs deteriorate rapidly into a very deep La Nina in the latter half of 2010.

7. Comparison of GISS, NCDC and HadCRUT Analyses

The likelihood of a record global temperature in the GISS analysis for 2010 raises the question about differences among the several global surface temperature analyses. For example, GISS and NCDC have 2005 as the warmest year in their analyses, while HadCRUT has 1998 as the warmest year. Here we investigate differences arising from two factors that are likely to be important: (1) the way that temperature anomalies are extrapolated, or not extrapolated, into regions without observing stations, and (2) the ocean data sets that are employed.

Figure 11 compares the GISS, NCDC, and HadCRUT analyses. The characteristic causing most interest and concern in the media and with the public is their different results for the warmest year in the record, as noted above.

A likely explanation for discrepancy in identification of the warmest year is the fact that the HadCRUT analysis excludes much of the Arctic, where warming has been especially large in the past decade, while GISS and NCDC estimate temperature anomalies throughout most of the Arctic. The difference between GISS and HadCRUT results can be investigated quantitatively using available data defining the area that is included in the HadCRUT analysis.

Figure 12 shows maps of GISS and HadCRUT 1998 and 2005 temperature anomalies relative to base period 1961-1990 (the base period used by HadCRUT). The temperature anomalies are at a 5°×5° (latitude-longitude) resolution in Figure 12 for the GISS data to match the resolution of the HadCRUT analysis. In the lower two maps we display the GISS data masked to the same area (and resolution) as the HadCRUT analysis.

Figure 13 shows time series of global temperature for the GISS and HadCRUT analyses, as well as for the GISS analysis masked to the HadCRUT data region. With the analyses limited to the same area, the GISS and HadCRUT results are similar. The GISS analysis finds 1998 as the warmest year, if analysis is limited to the masked area. This figure reveals that the differences that have developed between the GISS and HadCRUT global temperatures during the past decade are due primarily to the extension of the GISS analysis into regions excluded from the HadCRUT analysis.

The question is then: how valid are the extrapolations and interpolations in the GISS analysis? The GISS analysis assigns a temperature anomaly to many gridboxes that do not contain measurement data, specifically all gridboxes located within 1200 km of one or more stations that do have defined temperature anomalies. The rationale for this aspect of the GISS analysis is based on the fact that temperature anomaly patterns tend to be large scale, especially at middle and high latitudes.

The HadCRUT analysis also makes an (implicit) assumption about temperature anomalies in regions remote from meteorological stations, if the HadCRUT result is taken as a global analysis. The HadCRUT approach area-weights temperature anomalies of the regions in each hemisphere that have observations; then the means in each hemisphere are weighted equally to define the global result [Brohan *et al.*, 2006]. Thus HadCRUT implicitly assumes that areas without observations have a temperature anomaly equal to the hemispheric mean anomaly. Given the pattern of large temperature anomalies in the fringe Arctic areas with data (Figure 12), this implicit estimate surely understates the effect of Arctic temperature anomalies.

Qualitative support for the greater Arctic anomaly of the GISS analysis is provided by Arctic temperature anomaly patterns in the GISS analysis: regions warmer or cooler than average when the mean anomaly is adjusted to zero are realistic-looking meteorological patterns. More quantitative support is provided by satellite observations of infrared radiation from the Arctic [Comiso, 2006, 2010]. Although we have not yet attempted to integrate this infrared data record, which begins in 1981, into our temperature record, the temperature anomaly maps of Comiso [2006, 2010] have the largest positive temperature anomalies (several degrees Celsius) during the first decade of this century over the interior of Greenland and over the Arctic Ocean in regions where sea ice cover has decreased. Because there are no weather stations in central Greenland with long records or within the sea ice region, our analysis may understate warming in these regions. An exception is the station on Sakhalin Island, which is located in a region of decreasing sea ice cover and which does show relatively large warming in the past decade.

Comiso [2010] uses AVHRR (Advanced Very High Resolution Radiometer) data, because it provides the longest available satellite record. However, ongoing satellite observations with the MODIS (MODerate resolution Infrared Spectrometer) are also available (http://neo.sci.gsfc.nasa.gov/Search.html?datasetId=MOD_LSTAD_M) and provide another opportunity to check the validity of our extrapolations into regions without observing stations. cursory preliminary examination of data for Greenland suggests that Greenland temperature anomalies may be larger on that continent than those that we obtain via extrapolation from surface observation sites.

An independent method to extrapolate into regions without measurement stations is to use a global model in "reanalysis" mode, i.e., to insert available meteorological observations into the model, which then produces complete global "data" fields. Simmons *et al.* [2010] have carried out such reanalysis study, concluding that 2005 was likely the warmest year in the record.

Another analysis aimed specifically at evaluating the magnitude of Arctic warming in the past decade has been carried out by Chylek *et al.* [2010] using only surface air temperature

measurements at meteorological stations. They find about 0.25°C less Arctic warming during the past decade than in the GISS analysis, a difference that they attribute to our method of interpolating and extrapolating data, especially into the Arctic Ocean regions where no station data are available. We agree with Chylek et al. that the Arctic temperature change is uncertain and must be regarded with caution, but we make two observations. First, the Arctic covers about 4 percent of the global area, so an overestimate of its warming by 0.25°C would cause an overestimate of global temperature by only about 0.01°C and it would affect neighboring years as well as 2005, so it would not alter our conclusion about 2005 being the warmest year. Second, regions in the Arctic Ocean that changed from sea ice covered to open water in the past decade likely experienced a large increase in surface air temperature that would not be captured fully in an analysis using only land-based meteorological stations. Also greater warming in the interior of Greenland would not be captured, because there are no long-record interior stations in the analyses. Thus, although it is possible that the GISS analysis overstates the magnitude of Arctic warming in regions where data are extrapolated, it is also possible that the *Chylek et al.* [2010] analysis underestimates the warming in the Arctic Ocean.

Inadequate spatial coverage of stations is just one of the significant sources of uncertainty in global temperature change [*Karl et al.*, 1994; *Jones and Wigley*, 2010], but it is especially important in comparing nearby years, in contrast with slowly changing factors such as urban warming. We obtain a quantitative estimate of uncertainty (likely error) in the GISS global analysis due to incomplete spatial coverage of stations using a time series of global surface temperature generated by a long run of the GISS climate model runs [*Hansen et al.*, 2007]. We sample this data set at meteorological station locations that existed at several times during the past century. We then find the average error when the model's data for each of these station distributions are used as input to the GISS surface temperature analysis program.

Table 1 shows the derived error. As expected, the error is larger at early dates when station coverage was poorer. Also the error is much larger when data are available only from meteorological stations, without ship or satellite measurements for ocean areas. In recent decades the $2\text{-}\sigma$ uncertainty (95 percent confidence of being within that range, thus $\sim 2\text{-}3$ percent chance of being outside that range in a specific direction) has been about 0.05°C . Incomplete coverage of stations is the primary cause of uncertainty in comparing nearby years, for which the effect of more systematic errors such as urban warming is small.

Additional sources of error, including urban effects, become important when comparing temperature anomalies separated by longer periods [*Brohan et al.*, 2006; *Folland et al.*, 2001; *Smith and Reynolds*, 2002]. *Hansen et al.* [2006] estimated the additional error, by factors other than incomplete spatial coverage, as being $\sigma \approx 0.1^{\circ}\text{C}$ on time scales of several decades to a century, but that estimate is necessarily partly subjective. If that estimate is realistic, the total uncertainty in global mean temperature anomaly with land and ocean data included is similar to the error estimate in the first line of Table 1, i.e., the error due to limited spatial coverage when only meteorological stations are available. However, biases due to changing practices in ocean measurements may cause even greater uncertainty on a century time scale [*Rayner et al.*, 2006].

Occasionally measurement biases might change rapidly, as has been suggested for SSTs right after World War II. However, in most cases the error in comparing nearby years to each other is probably dominated by the substantial error from incomplete spatial coverage of measurements, rather than changes of measurement practices or urban warming. Under that assumption, let us consider whether we can specify a rank among the recent global annual temperatures, i.e., which year is warmest, second warmest, etc. Figure 9a shows 2009 as the second warmest year, but it is so close to 1998, 2002, 2003, 2006, and 2007 that we must declare

these years as being in a virtual tie as the second warmest year. The maximum difference among these years in the GISS analysis is $\approx 0.03^{\circ}\text{C}$ (2009 being the warmest among those years and 2006 the coolest). This total range is approximately equal to our 1σ uncertainty of $\approx 0.025^{\circ}\text{C}$.

Year 2005 is 0.06°C warmer than 1998 in the GISS analysis. How certain is it that 2005 was warmer than 1998? Given $\sigma \approx 0.025^{\circ}\text{C}$ for nearby years, we estimate the chance that 1998 was warmer than 2005 as follows. Actual 1998 and 2005 temperatures are specified by normal probability distributions about our calculated values. For each value of actual 1998 temperature there is a portion of the probability function for the 2005 temperature that has 2005 cooler than 1998. Integrating successively over the two distributions we find that the chance that 1998 was warmer than 2005 is 0.05, i.e., there is 95 percent confidence that 2005 was warmer than 1998.

The NCDC analysis finds 2005 to be the warmest year, but by a smaller amount. Thus a similar probability calculation for their results would estimate a greater chance that 1998 was actually warmer than 2005. NCDC reports 2009 as being the fifth warmest year (<http://www.ncdc.noaa.gov/sotc/?report=global>). Although the latter result seems to disagree with the GISS conclusion that 2009 tied for the second warmest year, this is mainly a consequence of the GISS preference to describe as statistical ties those years with global temperature differing by only about one standard deviation or less.

8. Weather variability versus climate trends

Public opinion about climate change is affected by recent and ongoing weather. North America had a cool summer in 2009, perhaps the largest negative temperature anomaly on the planet (Figure 14a). Northern Hemisphere winter (Dec-Jan-Feb) of 2009-2010 was unusually cool in the United States and northern Eurasia (Figure 14b). The cool weather contributed to increased public skepticism about the concept of "global warming", especially in the United States. These regional extremes occurred despite the fact that Jun-Jul-Aug 2009 was second warmest (behind Jun-Jul-Aug 1998) and Dec-Jan-Feb 2009-2010 was second warmest (behind Dec-Jan-Feb 2006-2007).

Northern Hemisphere winter of 2009-2010 was characterized by an unusual exchange of polar and mid-latitude air. Arctic air rushed into both North America and Eurasia, and, of course, was replaced in the polar region by air from middle latitudes. Penetration of Arctic air into middle latitudes is related to the Arctic Oscillation (AO) index [*Thompson and Wallace, 2000*], which is defined by surface atmospheric pressure. When the AO index (Figure 15a) is positive surface pressure is low in the polar region. This helps the middle latitude jet stream blow strongly and consistently from west to east, thus keeping cold Arctic air locked in the polar region. A negative AO index indicates relatively high pressure in the polar region, which favors weaker zonal winds, and greater movement of frigid polar air into middle latitudes.

December 2009 had the most extreme negative Arctic Oscillation since the 1970s. There were ten cases between the early 1960s and mid 1980s with negative AO index more extreme than -2.5, but no such extreme cases since then until December 2009. It is no wonder that the public had become accustomed to a reduction in extreme winter cold air blasts. Then, on the heels of the December 2009 anomaly, February 2010 had an even more extreme AO (Figure 15), the most negative AO index in the record that extends back to 1821 [<http://www.cru.uea.ac.uk/~timo/datapages/naoi.htm>].

The AO index and United States surface temperature anomalies are shown with monthly resolution in Figure 15b for two 6-year periods: the most recent years (2005-2010) and the period (1975-1980) just prior to the rapid global warming of the past three decades. Extreme negative temperature anomalies, when they occur, are usually in a winter month. Note that

winter cold anomalies in the late 1970s were more extreme than the recent winter cold anomalies.

Maps of monthly temperature anomalies are shown in Figure 16. Figure 16a is a polar projection for 24-90°N with 1200 km smoothing to provide full coverage, while 250 km resolution without further smoothing is possible for the contiguous United States (Figure 16b) because of high station density there. Monthly mean negative anomalies exceeded five degrees Celsius over large areas in the cold months of the late 1970s, while negative temperature excursions are more limited in 2009-2010.

Monthly temperature anomalies in the United States and Europe in winter are positively correlated with the AO index (Figures 15 and 17), with any lag between the index and temperature less than the monthly temporal resolution (Figure 15). The correlation of temperature anomalies and the AO index is especially large (correlation coefficient 61 percent) for Europe in winter. Winter correlation for the United States is 41 percent. The correlation in summer is 29 percent for the United States and 25 percent for Europe.

Thompson and Wallace [2000], *Shindell et al.* [2001] and others point out that increasing carbon dioxide causes the stratosphere to cool, in turn causing on average a stronger polar jet stream and thus a tendency for a more positive Arctic Oscillation. There is an AO tendency of the expected sense (Figure 15), but the change is too weak to account for the temperature trend. Indeed, Figure 17 shows that the warming trend of the past few decades has led to mostly positive seasonal temperature anomalies, even when the AO is negative. In the United States 16 of the past 20 winters and 15 of the past 20 summers were warmer than the 1951-1980 climatology, a frequency consistent with the expected "loading of the climate dice" [*Hansen*, 1997] due to global warming. Notable change of these probabilities is a result of the fact that local seasonal-mean temperature change due to long-term trends (Figure 18) is now comparable to the magnitude of local interannual variability of seasonal mean temperature (Figure 14b).

Such change of probability of a warm season is typical of the rest of the world. The right half of Figure 17 compares the AO index with temperature anomalies over "Europe", where, for convenience, we have approximated European borders by latitudes 36N-70N and longitudes 10W-30E. In this region 16 of the past 20 winters and 19 of the past 20 summers have been warmer than the 1951-1980 average.

Monthly temperature anomalies are typically 1.5 to 2 times greater than seasonal anomalies. So loading of the climate dice is not as easy to notice in monthly mean temperature. Daily weather fluctuations are even much larger than global mean warming. Yet it is already possible for an astute observer to detect the effect of global warming in daily data by comparing the frequency of days with record warm temperature to days with record cold temperature. The number of days with record high temperature now exceed the number of days with record cold by about a two to one ratio [*Meehl et al.*, 2009].

9. Data and analysis flaws

Figure 19 shows the effect of an error that came into the GISS analysis with the changes in the analysis described by *Hansen et al.* [2001]. One of the changes was use of an improved version of the USHCN station data records including adjustments developed by NCDC to correct for station moves and other discontinuities [*Easterling et al.*, 1996]. Our error was failure to recognize that the updates to the records for these stations obtained from NCDC electronically each month did not contain these adjustments. Thus there was a discontinuity in 2000 in those station records, as prior years contained the adjustment while later years did not.

The error was readily corrected, once it was recognized and we were notified, which occurred in 2007. Figure 19 shows the global and United States temperatures with and without the error. The error averaged 0.15°C over the contiguous 48 states, which is detectable in Figure 19b but much smaller than the 1-2°C interannual and interdecadal variations of United States temperature. Because the 48 states cover only about 1½ percent of the globe, the error in global temperature was about 0.003°C, which is insignificant and undetectable in Figure 19a.

This error was widely reported in the media, frequently with the assertion that NASA had intentionally exaggerated the magnitude of global warming, and with a further assertion that correction of the error made 1934 the warmest year in the record rather than 1998 [http://www.rushlimbaugh.com/home/daily/site_080907/content/01125106.guest.html]. Initial confusion between global and United States temperature was conceivably inadvertent, but attempts to correct this misstatement were ineffectual. As Figure 19b shows, the reality is that United States temperatures in 1934 and 1998 (and 2006) were and continue to be too similar to conclude that one year was warmer than the other.

Uncertainty in comparing United States temperatures for years separated by more than half a century is larger than 0.1°C. As shown above, the GISS adjustment for urban effects in the United States by itself approaches that magnitude, even when the adjustment is made on USHCN records that are already adjusted by NCDC to account for several sources of bias (inhomogeneity). The NCDC pairwise comparison of urban and rural stations is expected to remove much of the urban effect [Menne *et al.*, 2009].

Temperature records in the United States are especially prone to uncertainty, not only because of high energy use in the United States but also other unique problems such as the bias due to systematic change in the time at which observers read 24-hour max-min thermometers. These problems and adjustments to minimize their effect have been described in numerous papers by NCDC researchers [Karl *et al.* 1986, 1987, 1988, 1989; Quayle *et al.*, 1991; Easterling *et al.*, 1996; Peterson *et al.*, 1997, 1998a, 1998b].

When alterations, improvements, or adjustments occur in any of the input data streams (from meteorological stations, ocean measurements, or Antarctic research stations), the results of the GISS global temperature analysis change accordingly. Monthly updates of the GHCN (including USHCN) data records include not only an additional month of data but late station reports for previous months and sometimes corrections of earlier data. Thus slight changes in the GISS analysis can occur every month, but these changes are small in comparison with the global and United States temperature changes of the past century or the past three decades.

Occasionally changes of input data occur that are detectable in graphs of the data. This has occurred especially for United States meteorological station data, as NCDC has worked to improve the quality and homogeneity of these records. Figure 20 illustrates the effect of changes in the USHCN data that occurred when we switched (in November 2009) from USHCN version 1 to USHCN version 2, the latter being a NCDC update of their homogenization of USHCN stations). The effect of this revision of USHCN data is noticeable in the United States temperature (Figure 20), but it is small compared to century time scale changes. The effect on global temperature is imperceptible, of the order of a thousandth of a degree.

Based on our experience with the data flaw illustrated in Figure 19, we made two changes to our procedure. First, we now (since April 2008) save the complete input data records that we receive from the three near-real-time data sources every month. Although the three input data streams are publicly available from their sources, and a record is presumably maintained by the providing organizations, they are now also available from GISS. Second, we published the computer program used for our temperature analysis, making it available on our web site.

An additional data flaw occurred in November 2008. Although the flaw was only present in our data set for a few days (November 10-13), it resulted in additional lessons learned. The GHCN records for many Russian stations for November 2008 were inadvertently a repeat of October 2008 data. The GHCN records are not our data, but we properly had to accept the blame for the error, because the data were used in our analysis. Occasional flaws in input data are normal, and the flaws are eventually noticed and corrected if they are substantial. We have an effective working relationship with NCDC – reporting to them questionable data that we or our data users discover.

This specific data flaw was a case in point. The quality control program that NCDC runs on the data from global meteorological stations includes a check for repetition of data: if two consecutive months have identical data, the data are compared with that at the nearest stations. If it appears that the repetition is likely to be an error, the data are eliminated until the original data source has verified the data. The problem in the November 2008 data evaded this quality check because a change in their computer program inadvertently bypassed that quality check.

This data flaw led to another round of fraud accusations on talk shows and other media. Another lesson learned. Since then, to minimize misinformation, we first put our monthly analyzed data up on a site that is not visible to the public. This allows several scientists to examine graphs of the data for potential flaws. If anything seems questionable, we report it back to the data providers for their resolution. This process can delay availability of our data analysis to users for up to several days, and has resulted in a criticism that we now "hide" our data.

It should be noted that the data flaws discussed here are all either minor or temporary or both, and they do not compromise the integrity of the overall data product. The flaws are important, though, and efforts must be made to minimize them, because they can be used to cast doubt on the entire scientific enterprise.

It is impossible to entirely eliminate data flaws or satisfy all conflicting user demands. We believe that the steps we take now to check the data are a good compromise between assurance of data integrity and prompt availability, and also reasonable from the standpoint of the use of our time and resources. But we continue to seek ways to improve the data, we welcome user suggestions, and we appreciate it when problems are brought to our attention directly.

10. Summary Discussion

Human-made climate change has become an issue of surpassing importance to humanity, and global warming is the first order manifestation of increasing greenhouse gases that are predicted to drive climate change. Thus it is understandable that analyses of ongoing global temperature change are now subject to increasing scrutiny and criticisms that are different than would occur for a purely scientific problem.

Our comments here about communication of this climate change science to the public are our opinion. Other people may have quite different opinions. We offer our opinion because it seems inappropriate to ignore the vast range of claims appearing in the media and in hopes that open discussion of these matters may help people distinguish the reality of global change sooner than would otherwise be the case. However these comments, even though based on experience over a few decades, are only opinion. Our primary contribution is quantitative results discussed in the numbered paragraphs below.

Communication of the status of global warming to the public has always been hampered by weather variability. Lay people's perception tends to be strongly influenced by the latest local fluctuation. This difficulty can be alleviated by stressing the need to focus on the frequency and

magnitude of warm and cold anomalies, which change noticeably on decadal time scales as global warming increases.

Other obstacles to public communication include the media's difficulty in framing long-term problems as 'news', a preference for sensationalism, a generally low level of familiarity with basic science, and a preference for 'balance' in every story. The difficulties are compounded by the politicization of reporting of global warming, a perhaps inevitable consequence of economic and social implications of efforts required to alter the course of human-made climate change.

The task of alleviating the communication obstacle posed by politicization is formidable, and it is made more difficult by attacks on the character and credibility of scientists [for example: <http://mediamatters.org/blog/200911290004> and <http://mediamatters.org/mmtv/201001290037>]. Polls indicate that the attacks have been effective in causing many members of the public to doubt the reality or seriousness of global warming [<http://www.gallup.com/poll/126560/americans-global-warming-concerns-continue-drop.aspx>].

Given this situation, the best hope may be repeated clear description of the science and passage of sufficient time to confirm validity of the description. A problem with that prescription is the danger that the climate system could pass tipping points that cause major climate changes to proceed largely out of humanity's control [*Hansen et al.*, 2008]. Yet continuation of careful scientific description of ongoing climate change seems to be essential for the sake of minimizing the degree of future climate change, even while other ways are sought to draw attention to the dangers of continued greenhouse gas increases.

One lesson we have learned is that making our global data analysis immediately available, with data use by ourselves and others helping to reveal flaws in the input data, has a practical disadvantage: it allows any data flaws to be interpreted and misrepresented as machinations. Yet the data are too useful for scientific studies to be kept under wraps, so we will continue to make the data available on a monthly basis. But we are making special efforts to make the process as transparent as possible, including availability of the computer program that does the analysis, the data that goes into the analysis (also available from original sources), and detailed definition of urban adjustment of meteorological station data.

Our principal task remains the scientific one – describing ongoing global temperature change with as much clarity and insight as we can. Contributions of the present paper include:

(1) Insight into why the GISS analysis yields 2005 as the warmest calendar year, while the HadCRUT analysis has 1998 as the warmest year. The main factor is our inclusion of estimated temperature change for the Arctic region. We note that SST change cannot be used as a measure of surface air temperature change in regions of sea ice, and that surface air temperature change is the quantity of interest both for its practical importance to humans and for comparison with the results that are usually reported in global climate model studies.

(2) Twelve-month (and $n \times 12$ -month) running mean temperatures provide more information than the usual graphs with calendar-year mean temperature. The magnitude and duration of global temperature effects of volcanoes and the Southern Oscillation can be seen much more clearly in a 12-month running mean graph such as Figure 10. The simplicity of the running mean, compared to filtered time series, is helpful for public communications.

(3) The 12-month running mean global temperature in the GISS analysis has reached a new record in 2010. The new record temperature in 2010 is particularly meaningful because it occurs when the recent minimum of solar irradiance (*Frohlich*, 2006; data at <http://www.pmodwrc.ch/pmod.php?topic=tsi/composite/SolarConstant>) is having its maximum cooling effect. At the time of this writing (July 2010) the tropical Pacific Ocean is

changing from El Nino to La Nina conditions in the tropical Pacific Ocean. It is likely that global temperature for calendar year 2010 will reach a record level for the period of instrumental data, but that is not certain if La Nina conditions deepen rapidly.

(4) The cool weather anomalies in the United States in Jun-Jul-Aug 2009 and in both the United States and northern Eurasia in the following Dec-Jan-Feb are close to the cool extreme of the range of seasonal temperatures that are now expected (Figure 17) given the warming of the past few decades. Although comparably cool conditions could occur again sometime during the next several years, the likelihood of such event is low in any given year and it will continue to decrease as global warming continues to increase.

(5) We suggest a new procedure for use of satellite SST data that takes advantage of the high spatial resolution and broad coverage of satellite observations but avoids the bias in the temperature trend in satellite data [Reynolds *et al.*, 2002, 2010]. We adjust the satellite data by a small constant such that the monthly temperature anomalies of satellite and in situ data are equal over their common area. This procedure is used in our current ERSST+OI analysis. We continue to also provide our HadISST1+OI analysis, without such adjustment, as our standard data product. Because of a cold bias in unadjusted OI data, global warming in ERSST+OI exceeds that in HadISST1+OI by about 0.04°C in 2010. Further study is needed to verify which of these data products is superior. Other improvements of the ocean data sets may become available in the near future. For example, none of the publicly available global data sets corrects as yet for a discontinuity in ocean data that has been suggested to exist near the end of World War II [Thompson *et al.*, 2008]. However, note that none of these adjustments or uncertainties is large enough to alter any of our major conclusions in this paper.

(6) Global warming on decadal time scales is continuing without letup. Figure 8, showing decadal mean temperature anomalies, effectively illustrates the monotonic and substantial warming that is occurring on decadal time scales. But because it is important to draw attention to change as soon as possible, we need ways to make the data trends clear without waiting for additional decades to pass. Figure 21 shows the 60-month (5-year) and 132-month (11-year) running means of global temperature. The 5-year mean is sufficient to minimize ENSO variability, while the 11-year mean also minimizes the effect of solar variability. Figure 21 gives the lie to the frequent assertion that "global warming stopped in 1998". Of course it is possible to find almost any trend for a limited period via judicious choice of start and end dates of a data set that has high temporal resolution, but that is not a meaningful exercise. Even a more moderate assessment, "the trend in global surface temperature has been nearly flat since the late 1990s despite continuing increases in the forcing due to the sum of the well-mixed greenhouse gases" [Solomon *et al.*, 2009], is not supported by our data. On the contrary, we conclude that there has been no reduction in the global warming trend of 0.15-0.20°C/decade that began in the late 1970s.

Acknowledgments. We thank Wayne Hamilton, Phil Jones, Tom Karl, Nick Rayner, Dick Reynolds, Alan Robock and Gavin Schmidt for helpful suggestions on a draft version of the paper.

References

- Beck, J.W., R.L. Edwards, E. Ito, F.W. Taylor, J. Recy, F. Rougerie, P. Joannot, and C. Henin, Sea-surface temperature from coral skeletal strontium/calcium ratios, *Science*, **257**, 644-647, 1992.
- Brohan, P., J.J. Kennedy, I. Harris, S.F.P. Tett, and P.D. Jones, Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850. *J. Geophys. Res.*, **111**, D12106, doi:10.1029/2005JD006548, 2006.
- Christy, J.R., W.B. Norris, and R.T. McNider, Surface temperature variations in East Africa and possible causes. *J. Clim.*, **22**, 3342-3356, 2009.
- Chylek, P., C.K. Folland, D.E. Parker, G. Lesins, M. Wang, M.K. Dubey, and L.A. Chylek, How large is recent Arctic warming? *Geophys. Res. Lett.* (in review).
- Comiso, J.C., Arctic warming signals from satellite observations, *Weather*, **61**, 70-76, 2006.
- Comiso, J.C., *Polar Oceans from Space*, Springer, New York, 507 pp., 2010.
- Easterling, D.R., T.C. Peterson, and T.R. Karl, On the development and use of homogenized climate data sets, *J. Clim.*, **9**, 1429-1434, 1996.
- Folland, C.K., N.A. Rayner, S.J. Brown, T.M. Smith, S.S.P. Shen, D.E. Parker, I. Macadam, P.D. Jones, R.N. Jones, N. Nichols, and D.M.H. Sexton, Global temperature change and its uncertainties, *Geophys. Res. Lett.*, **28**, 2621-2624, 2001GL012877, 2001.
- Foster, G., J.D. Annan, P.D. Jones, M.E. Mann, B. Mullan, J. Renwick, J. Salinger, G.A. Schmidt, and K.E. Trenberth, Comment on "Influence of the Southern Oscillation on tropospheric temperature" by J.D. McLean, C.R. de Freitas, and R.M. Carter, *J. Geophys. Res.*, **115**, doi:10.1029/2009JD012960.
- Fröhlich, C., Solar irradiance variability since 1978: revision of the PMOD composite during solar cycle 21. *Space Science Rev.*, **125**, 53-65., 2006
- Hansen, J., D. Johnson, A. Lacis, S. Lebedeff, P. Lee, D. Rind, and G. Russell: Climate impact of increasing atmospheric carbon dioxide. *Science*, **213**, 957-966, 1981.
- Hansen, J.E., and S. Lebedeff, Global trends of measured surface air temperature. *J. Geophys. Res.*, **92**, 13345-13372, 1987.
- Hansen, J., Public understanding of global climate change, pp. 247-253 in *Carl Sagan's Universe*, eds. Y. Terzian and E. Bilson, Cambridge University Press, 282 pp., 1997.
- Hansen, J., R. Ruedy, J. Glascoe, and M. Sato, GISS analysis of surface temperature change. *J. Geophys. Res.*, **104**, 30997-31022, 1999.
- Hansen, J.E., R. Ruedy, M. Sato, M. Imhoff, W. Lawrence, D. Easterling, T. Peterson, and T. Karl, A closer look at United States and global surface temperature change. *J. Geophys. Res.*, **106**, 23947-23963, 2001.
- Hansen, J., M. Sato, R. Ruedy, K. Lo, D.W. Lea, and M. Medina-Elizade, Global temperature change. *Proc. Natl. Acad. Sci.*, **103**, 14288-14293, 2006.

Hansen, J., M. Sato, R. Ruedy, P. Kharecha, A. Lacis, R.L. Miller, L. Nazarenko, K. Lo, G.A. Schmidt, G. Russell, et al., Climate simulations for 1880-2003 with GISS modelE. *Clim. Dynam.*, **29**, 661-696, 2007 doi:10.1007/s00382-007-0255-8.

Hansen, J., M. Sato, P. Kharecha, D. Beerling, R. Berner, V. Masson-Delmotte, M. Pagani, M. Raymo, D.L. Royer, and J.C. Zachos, Target atmospheric CO₂: where should humanity aim? *Open Atmos. Sci. J.*, **2**, 217-231, 2008.

Hansen, J., *Storms of My Grandchildren*, Bloomsbury USA, New York, 304 pp., 2009.

Hughes, S.L., N.P. Holliday, E. Colbourne, V. Ozhigin, H. Valdimarsson, S. Osterhus, and K. Wilshire, Comparison of *in situ* time-series of temperature with gridded sea surface temperature datasets in the North Atlantic, ICES j. *Mar. Sci. Adv. Access*, 15 March 2009 doi:10.1093/icesjms/fsp041

Imhoff, M.L., Lawrence, W.T., Stutzer, D.C. and Elvidge, C.D., A technique for using composite DMSP-OLS "city lights" satellite data to map urban area. *Remote Sens. Envir.*, **61**, 361-370, 1997.

IPCC, *Climate Change 2007: The Physical Science Basis*, eds. S. Solomon et al., Cambridge University Press, 996 pp., 2007.

Jones, P.D., T.M.L. Wigley, and P.M. Kelly, Variations in surface air temperature: Part 1: Northern Hemisphere, 1881-1980. *Mon. Wea. Rev.*, **110**, 59-70, 1982.

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

Jones, P.D., M. New, D.E. Parker, S. Martin, and I.G. Rigor, Surface air temperature and its variations over the last 150 years, *Rev. Geophys.*, **37**, 173-199, 1999.

Jones, P.D. and T.M.L. Wigley, Estimation of global temperature trends: What's important and what isn't, *Clim. Change*, **100**, 59-69, 2010.

Karl, T.R., C.N. Williams, P.J. Young, and W.M. Wendland, Model to estimate the time of observation bias associated with monthly mean maximum, minimum and mean temperatures for the United States, *J. Clim Appl. Meteorol.*, **25**, 145-160, 1986.

Karl, T.R., and C.N. Williams, An approach to adjusting climatological time series for discontinuous inhomogeneities, *J. Clim. Appl. Meteorol.*, **26**, 1744-1763, 1987.

Karl, T.R., H.F. Diaz, and G. Kukla, Urbanization: Its detection and effect in the United States climate record, *J. Clim.*, **1**, 1099-1123, 1988.

Karl, T.R., J.D. Tarplay, R.G. Quayle, H.F. Diaz, D.A. Robinson, and R.S. Bradley, The recent climate record: What it can and cannot tell us, *Rev. Geophys.*, **27**, 405-430, 1989.

Karl, T.R., C.N. Williams, and F.T. Quinlan, United States Historical Climatology Network (HCN) Serial Temperature and Precipitation Data. ORNL/CDIAC-30, NDP-019/R1. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tennessee, 1990.

Karl, T.R., R.W. Knight, and J.R. Christy, Global and hemispheric temperature trends related to inadequate spatial sampling, *J. Clim.*, **7**, 1144-1163, 1994.

Lugina, K.M., P. Ya. Groisman, K.Ya. Vinnikov, V.V. Koknaeva, and N.A. Sperandkaya, Monthly surface air temperature time series area-averaged over the 30-degree latitudinal belts of the globe, 1881-2005, in *Trends: A Compendium of Data on Global Change*, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tennessee doi:10.3334/CDIAC/cli.003, 2006.

Meehl, G.A., C. Tebaldi, G. Walton, D. Easterling, and L. McDaniel, Relative increase of record high maximum temperatures compared to record low minimum temperatures in the U.S., *Geophys. Res. Lett.*, **36**, L23701, 2009, doi:10.1029/2009GL040736.

Menne, M.J., C.N. Williams, and R.S. Vose, The U.S. historical climatology network monthly temperature data, version 2, *Bull. Amer. Meteorol. Soc.*, **90**, 993-1007, 2009.

Oke, T.R., The micrometeorology of urban forest, *Phil. Trans. Roy. Soc., Biolog. Sci.*, **324(B)**, 335-349, 1989.

Parker, E.E., C.K. Folland, and M. Jackson, Marine surface temperature: observed variations and data requirements, *Clim. Change*, **31**, 559-600, 1995

Parker, D.E., Large-scale warming is not urban, *Nature*, **432**, 290, 2004.

Parker, D.E., Urban heat island effects on estimates of observed climate change, *Wiley Interdisciplinary Reviews: Climate Change*, **1**, 123-133, doi:10.1002/wcc.o21, 2010.

Peterson, T.C. and R. Vose, An overview of the global historical climatology network temperature database, *Bull. Am. Meteorol. Soc.*, **78**, 2837-2850, 1997.

Peterson, T.C., D.R. Easterling, T.R. Karl and 18 co-authors, Homogeneity adjustments of in situ atmospheric climate data: A review, *Int. J. Climatol.*, **18**, 1493-1517, 1998a.

Peterson, T.C., R. Vose, R. Schmoyer, and V. Razuvaev, Global historical climatology network (GHCN) quality control of monthly temperature data, *Int. J. Climatol.*, **18**, 1169-1179, 1998b.

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

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

Philander, S.G., *Our Affair with El Niño: How We Transformed an Enchanting Peruvian Current into a Global Climate Hazard*, Princeton University Press, Princeton, 288 pp., 2006.

Quayle, R.G., D.R. Easterling, T.R. Karl, and P.Y. Hughes, Effects of recent thermometer changes in the cooperative station network, *Bull. Amer. Meteorol. Soc.*, **72**, 1718-1724, 1991.

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

- Rayner, N.A., P. Brohan, D.E. Parker, C.K. Folland, J.J. Kennedy, M. Vanicek, T.J. Ansell, and S.F.B. Tett, Improved analyses of changes and uncertainties in sea surface temperature measured in situ since the mid-nineteenth century: the HadSST2 dataset, *J. Clim.*, **19**, 446-469, 2006.
- Reynolds, R.W., N.A. Rayner, T. M. Smith, D.C. Stokes, and W. Wang, An improved *in situ* and satellite SST analysis for climate, *J. Clim.*, **15**, 1609-1625, 2002.
- Reynolds, R.W., C.L. Gentemann, and G.K. Corlett, Evaluation of AATSR and TMI satellite SST data, *J. Clim.*, **23**, 152-165, 2010.
- Shindell, D., G.A. Schmidt, R.L. Miller, and D. Rind, Northern Hemisphere winter climate response to greenhouse gas, volcanic, ozone and solar forcing, *J. Geophys. Res.*, **106**, 7193-7210, 2001.
- Simmons, A.J., K.M. Willett, P.D. Jones, P.W. Thorne, and D. Dee, Low-frequency variations in surface atmospheric humidity, temperature and precipitation: Inferences from reanalyses and monthly gridded observational datasets, *J. Geophys. Res.*, **115**, D01110, doi:10.1029/2009JD012442.
- Smith, T.M., and R.W. Reynolds, Bias corrections for historic sea surface temperatures based on marine air temperatures, *J. Clim.*, **15**, 73-87, 2002.
- Smith, T.M., R.W. Reynolds, and J. Lawrimore, Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880-2006), *J. Clim.*, **21**, 2283-2296, 2008.
- Solomon, S., K. Rosenlof, R. Portmann, J. Daniel, S. Davis, T. Sanford, and G.K. Plattner, *ScienceExpress*, 10.1126/science.1182488, 28 January 2010.
- Sutton, R.T., B. Dong, and J.M. Gregory, Land/sea warming ratio in response to climate change: IPCC AR4 model results and comparison with observations, *Geophys. Res. Lett.*, **34**, L02701.
- Thompson, D.W.J., and J.M. Wallace, Annular modes in the extratropical circulation: Part I: Month-to-month variability, *J. Clim.*, **13**, 1000-1016, 2000.
- Thompson, D.W.J., J.J. Kennedy, J.M. Wallace, and P.D. Jones, A large discontinuity in the mid-twentieth century in observed global-mean surface temperature, *Nature*, **453**, 646-649, 2008.
- Thompson, D.W.J., J.M. Wallace, P.D. Jones, and J.J. Kennedy, Identifying signatures of natural climate variability in time series of global-mean surface temperature: methodology and insights, *J. Clim.*, **22**, 6120-6141, 2009.
- Turner, J., S. R. Colwell, G. J. Marshall, T. A. Lachlan-Cope, A. M. Carleton, Phil D. Jones, V. Lagun, P. A. Reid, and S. Iagovkina, The SCAR READER project: Toward a high quality database of mean Antarctic meteorological observations. *J. Clim.*, **17**, 2890-2898, 2004.
- Worley, S.J., S.D. Woodruff, R.W. Reynolds, S.J. Lubker, and N. Lott, ICOADS release 2.1. data and products. *Int. J. Climatol.*, **25**, 823-842, 2005.
- Yasunaka, S., and K. Hanawa, Intercomparison of historical sea surface temperature datasets. *Int. J. Climatol.*, DOI: 10.1002/joc.2104, 2010

Appendix

Figure A1(a) shows the global distribution of pitch dark stations (nightlight radiance less than $1 \mu\text{W}/\text{m}^2/\text{sr}/\mu\text{m}$) and stations that are dark but not pitch dark (radiance between 1 and $32 \mu\text{W}/\text{m}^2/\text{sr}/\mu\text{m}$). Figure A1(b) compares the analyzed global temperature change for the case of our standard nightlight adjustment and the case in which pitch dark stations are used to adjust the long-term trend of all other stations. As the figure shows, adjustment using only pitch dark stations has very little effect on the result. Indeed, the global mean warming is slightly larger (by 0.01°C) using the stricter nightlight adjustment. Conceivably the slight warming is a result of the fact that the pitch dark requirement removes about three-quarters of the airport stations from those used. However, we have not investigated which specific stations cause the slight change in Figure A1, because the change is negligible in comparison with the total temperature change and its uncertainty.

The global temperature change obtained in our analysis depends on the ocean surface temperature data set(s) that we employ. We compare here the HadISST1, ERSST, and HadSST2 data sets, as well as the data sets that result when OI satellite data are concatenated with HadISST1 or ERSST. Here the OI concatenation procedure is the same for all data sets, the OI mean for 1982-1992 being equated to the mean of the other data set for the same period.

Figure A2 compares the HadISST1, HadSST2 and ERSST data sets after spatial averaging. The top row is the temperature anomaly averaged over only those regions where all three records have data. The lower graphs are estimates of the global ocean mean anomaly, obtained as follows: monthly SSTs are interpolated to $5^\circ \times 5^\circ$ grid (HadSST2 grid), mean anomalies are computed for the permanently ice-free ocean area with defined SST within each of four latitude zones ($90\text{-}25^\circ\text{S}$, $25\text{-}0^\circ\text{S}$, $0\text{-}25^\circ\text{N}$, $25\text{-}90^\circ\text{N}$), the global-ocean mean is computed as the average of these four zones with each zone weighted by the open-ocean area of that zone.

Variations among the resulting data sets are within the expected uncertainty ranges [Folland *et al.*, 2001; Rayner *et al.*, 2006; Reynolds *et al.*, 2002; Smith *et al.*, 2008]. It is notable that the largest differences occur in the past decade, when the most comprehensive observations exist. Figure A3 shows the geographical distribution of the differences between the data sets.

The largest difference between HadSST2 and HadISST1, as well as the largest difference between ERSST and HadISST1, occurs in the upwelling region just west of South America. Presumably this change from the earlier HadISST1 is a result of the more comprehensive data available from ICOADS release 2.1 [Worley *et al.*, 2005]. However, HadSST2 and ERSST are also warmer than HadISST1 during the past decade throughout most of the global ocean. The ubiquity of the recent differences suggests that they may be related to calibration of satellite measurements, which are a major data source in the past decade.

The effect on the GISS analysis of global temperature change caused by alternative choices for the ocean data record is reduced by the fixed contributions from meteorological stations on continents and islands, including extrapolation into the Arctic. Figure 6 compares the global temperature records that result with HadISST1+OI and ERSST+OI, the differences being as much as several hundredths of a degree Celsius.

The standard GISS global analysis uses the concatenated HadISST1+OISST data set, as described in the main text. Any of the alternative ocean data sets described here would yield slightly greater global warming, both in recent decades and on the century time scale.

Until improved assessments of the alternative SST data sets exist, the GISS global analysis will be made available for both HadISST1+OISST and ERSST+OISST. The HadISST1 concatenation with OISST will continue to be based on equating means for 1982-1992, as it always has been in the standard GISS analysis. The ERSST+OISST concatenation will be as

described in the main text of this paper. HadISST1+OISST will continue to be our standard product unless and until verifications show ERSST+OISST to be superior.

Figure A4 investigates a charge that has been bruited about frequently in the past year, specifically the claim that GISS has systematically reduced the number of stations used in its temperature analysis so as to introduce an artificial global warming. GISS uses all of the GHCN stations that are available, but the number of reporting meteorological stations in 2009 was only 2490, compared to ~6300 usable stations in the entire 130-year GHCN record. The reduced number of stations is in part a consequence of the fact that not all stations report in near-real-time, so their data may be included in future GHCN records. However, Figure A4 shows that the additional stations will not make much difference. When the GISS analysis uses only stations reporting in 2009 the temperature curve obtained for recent years is almost identical to curve obtained using all 6300 stations.

Table 1. Two-sigma (2σ) error estimate vs time for meteorological stations and land-ocean index

	1880-1900	1900-1950	1960-2008
Meteorological Stations	0.2	0.15	0.08
Land-Ocean Index	0.08	0.05	0.05

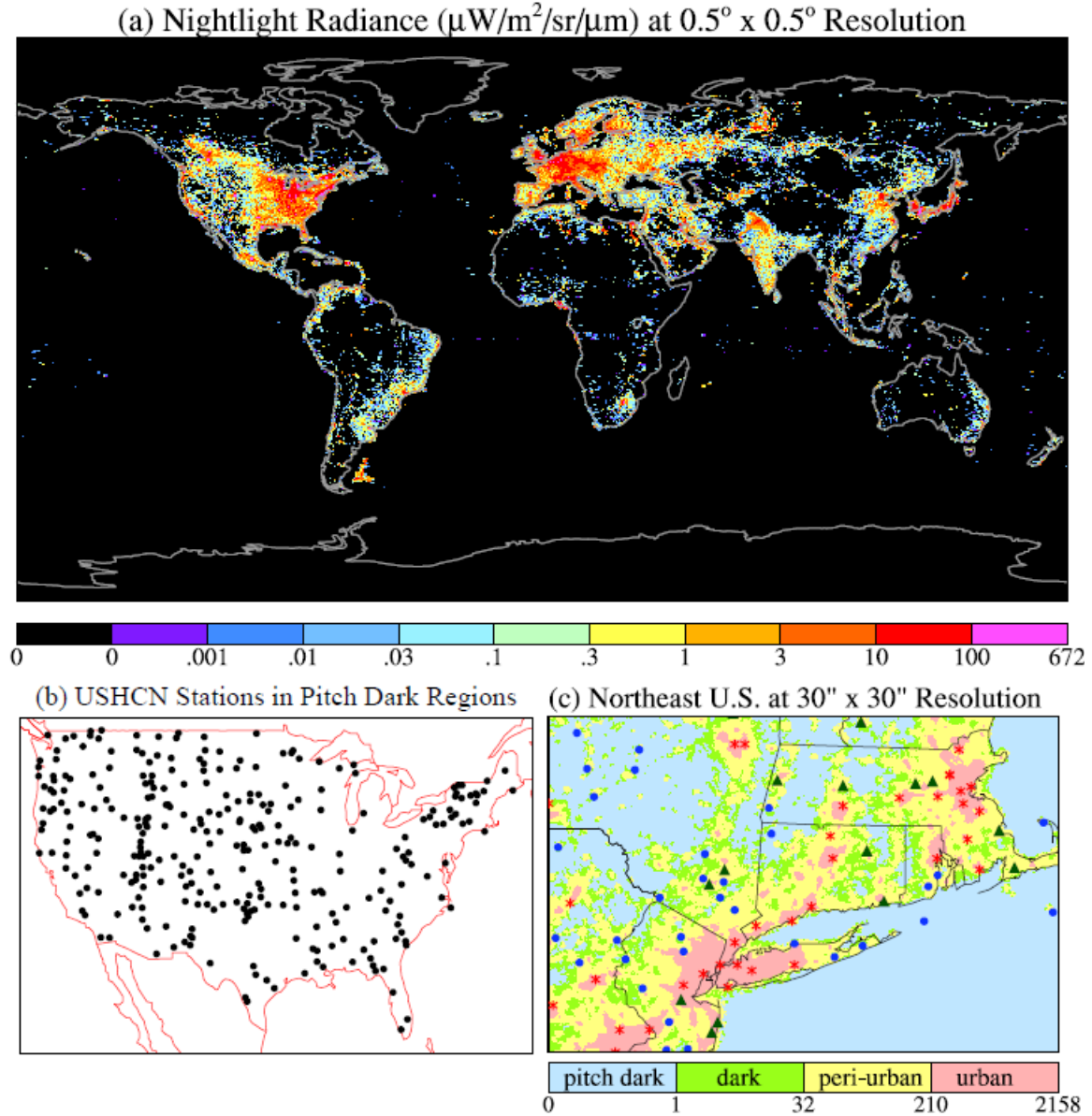


Figure 1. (a) Satellite-observed nightlight radiances at a spatial resolution of $0.5^\circ \times 0.5^\circ$, (b) locations of USHCN stations in extreme darkness, nightlight radiance less than $1 \mu\text{W}/\text{m}^2/\text{sr}/\mu\text{m}$, and (c) a region shown at the data resolution of $0.0083^\circ \times 0.0083^\circ$ [Imhoff *et al.*, 1997]. Light blue area is that seen as "pitch dark" from the satellite, i.e., with radiance less than $1 \mu\text{W}/\text{m}^2/\text{sr}/\mu\text{m}$; the green area is "dark", between 1 and $32 \mu\text{W}/\text{m}^2/\text{sr}/\mu\text{m}$. Blue dots, green triangles, and red asterisks indicate meteorological stations associated with towns having population less than 10,000, between 10,000 and 50,000 and greater than 50,000, respectively. In our new standard nightlight treatment stations in the yellow and pink regions are adjusted for urban effects.

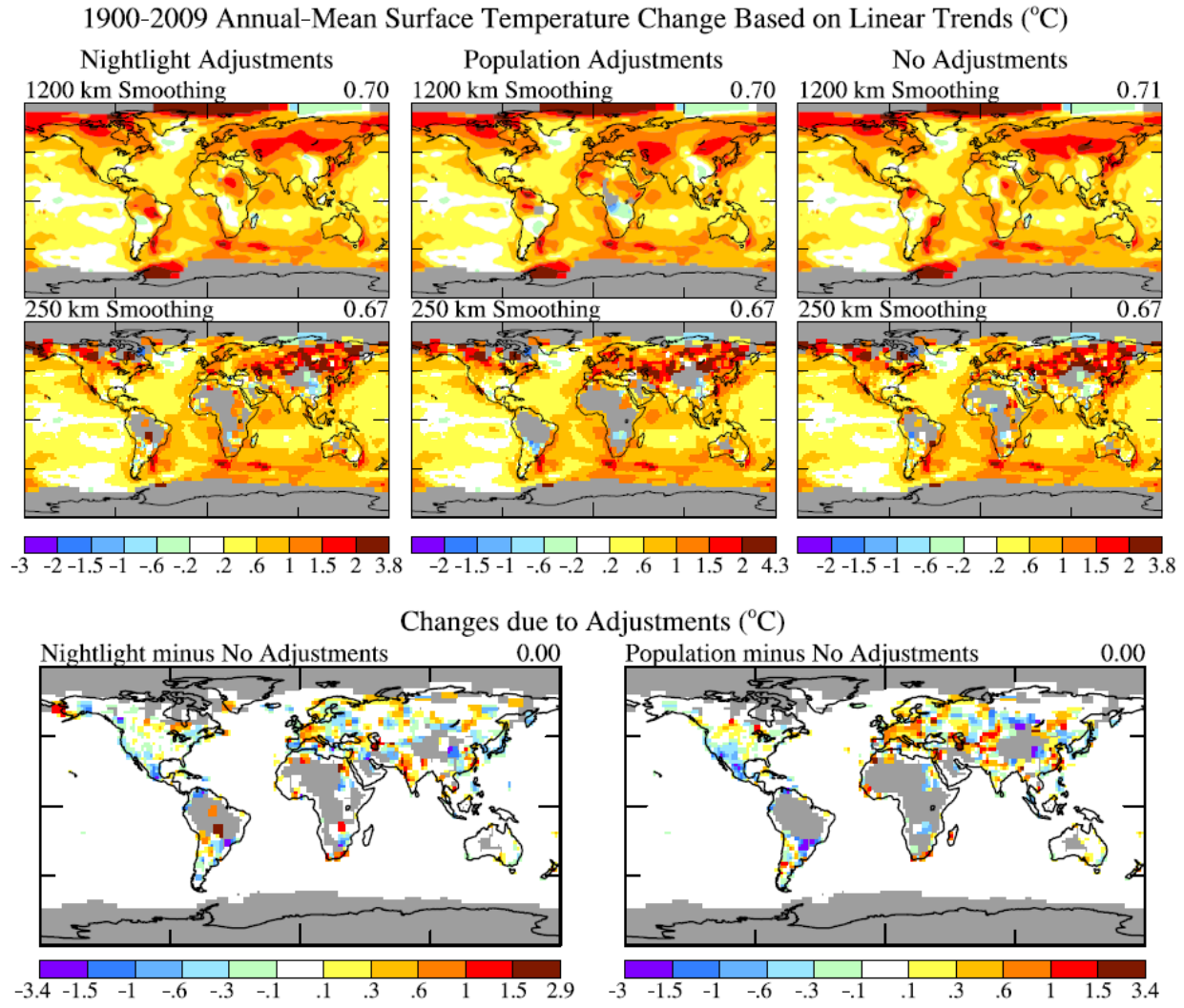


Figure 2. Comparison of alternative urban adjustments. The top row uses the standard 1200 km radius of influence for each station while the second row reduces the radius of influence to 250 km, so that the influence of adjustments on individual stations can be ascertained. The temperature change due to the adjustments is shown in the two bottom maps. Numbers on upper right corner of these and other maps are the global means.

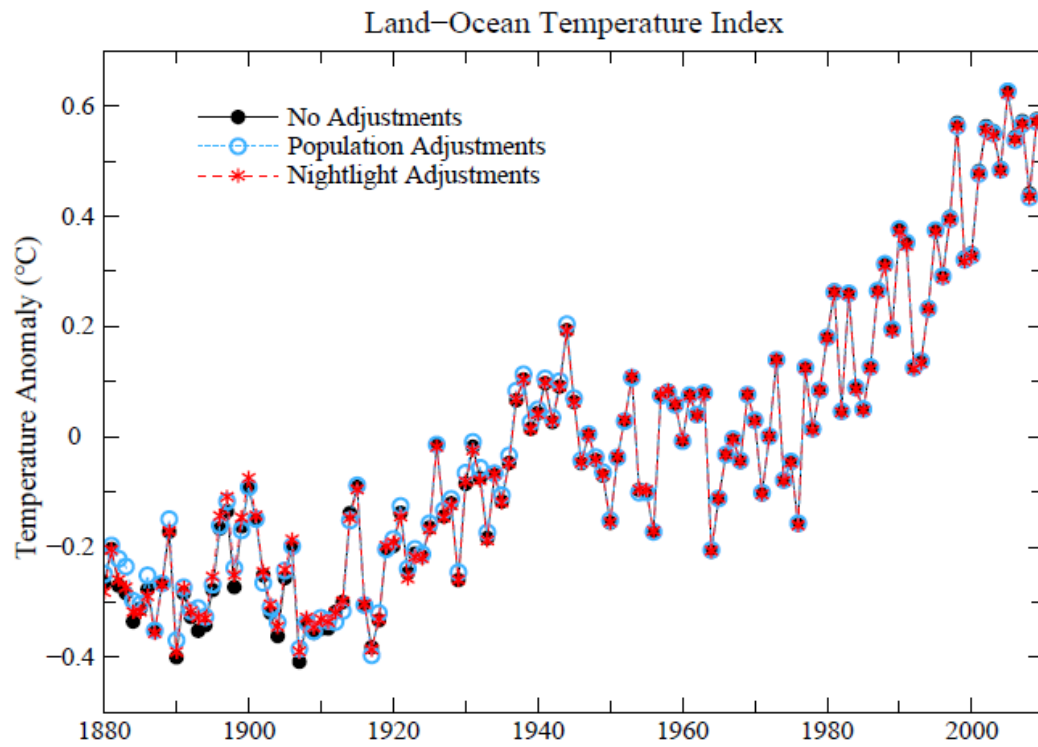


Figure 3. Global-mean annual-mean land-ocean temperature index for three alternative treatments of the urban adjustment.

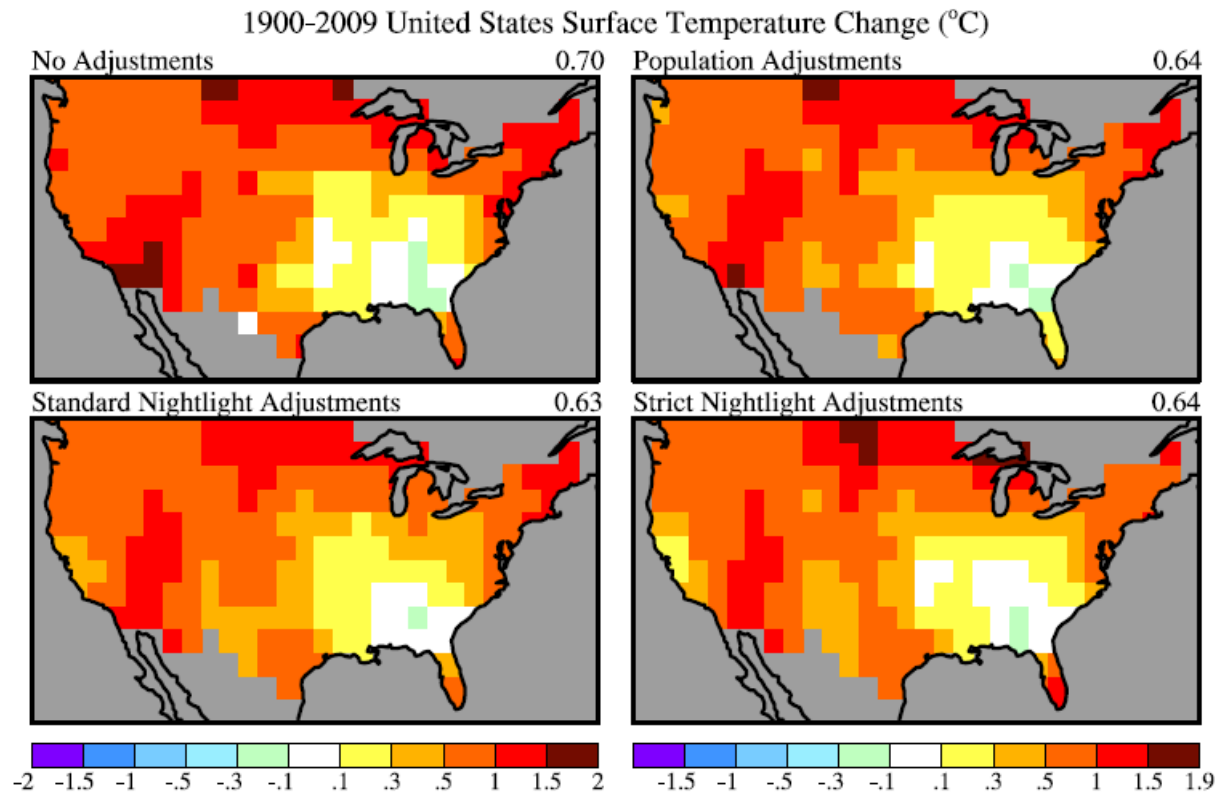


Figure 4. Temperature change in the United States for alternative choices of urban adjustment.

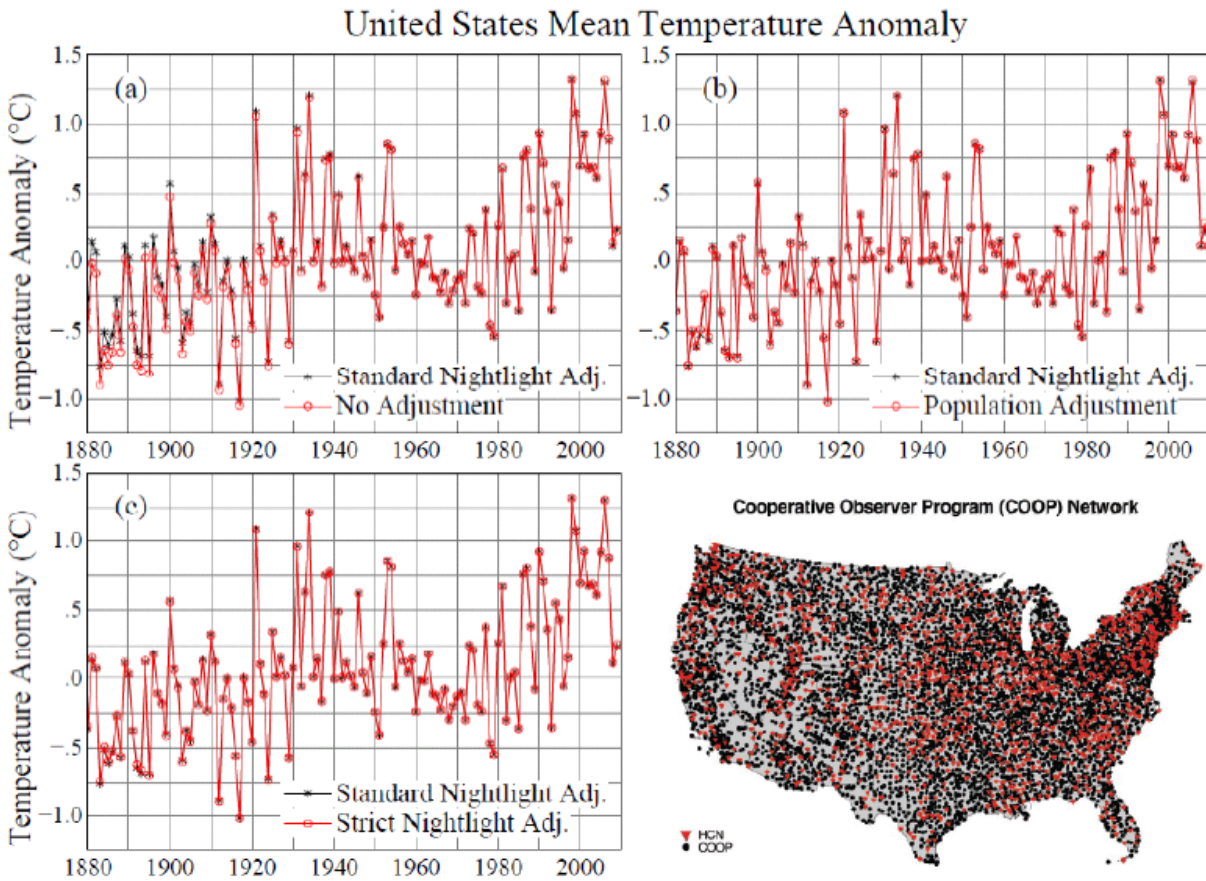


Figure 5. Comparisons of mean temperature anomalies in the contiguous 48 United States for the standard GISS nightlight adjustment and alternative. The map on the lower right shows the high density of meteorological stations in the United States, with red being stations in the U.S. Historical Climatology Network and black being Cooperative stations [Menne *et al.*, 2009].

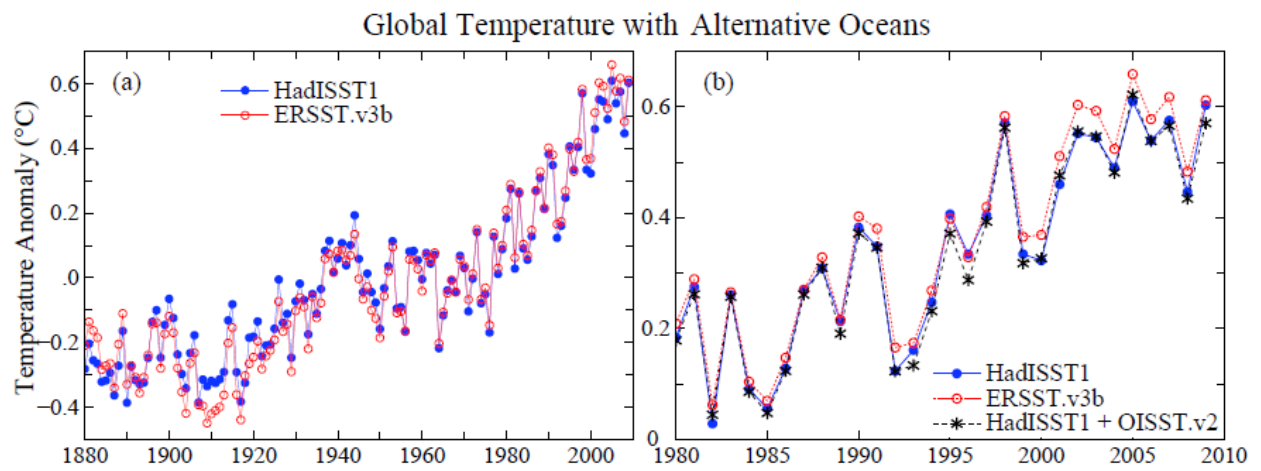


Figure 6. Global temperature change in the GISS global analysis for alternative choices of sea surface temperature.

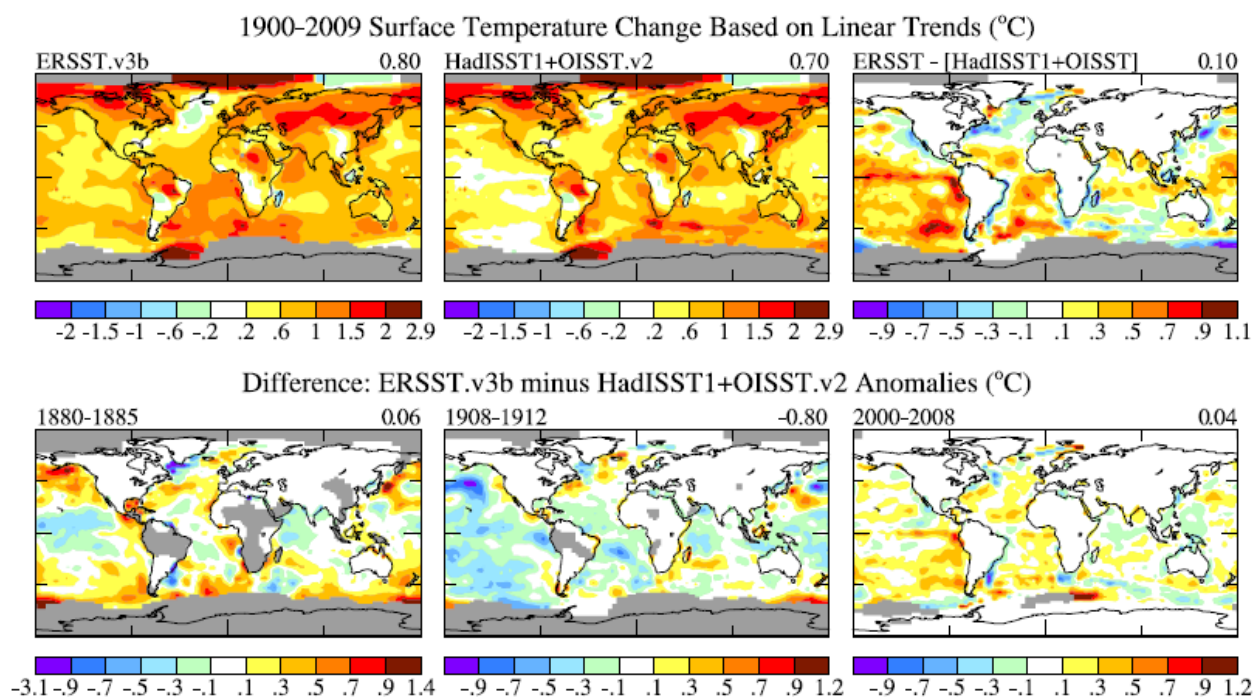


Figure 7. Temperature change in the GISS global analysis using ERSST and HadISST1+OISST and differences in specific periods. Numbers on upper right corners are global means.

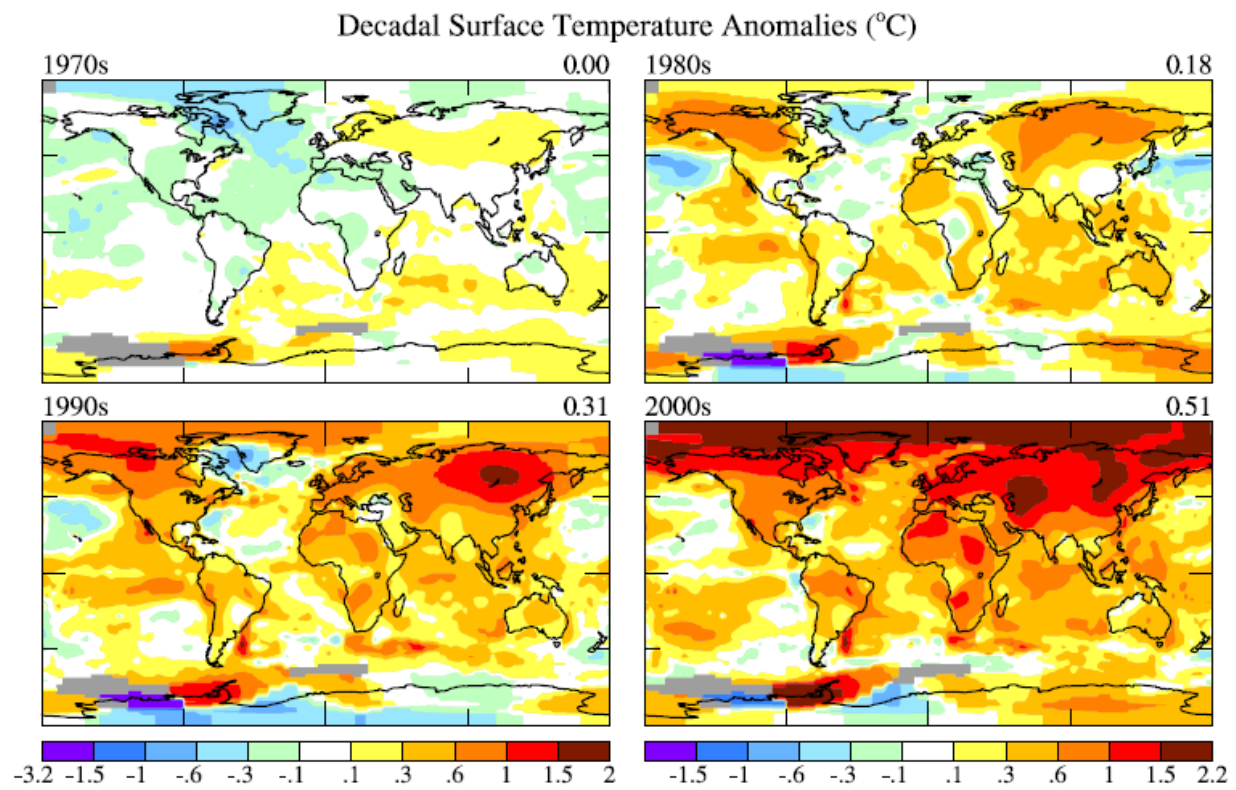


Figure 8. Decadal surface temperature anomalies relative to 1951-1980 base period.

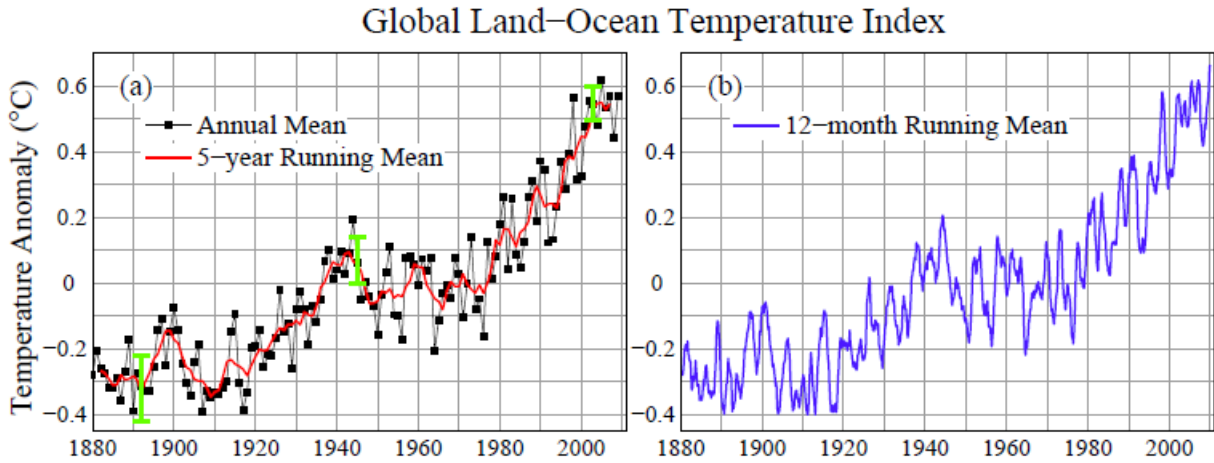


Figure 9. Global surface temperature anomalies relative to 1951-1980 mean for (a) annual and 5-year running means through 2009, and (b) 12-month running mean using data through June 2010. Green bars are 2- σ error estimate, as discussed in section 7.

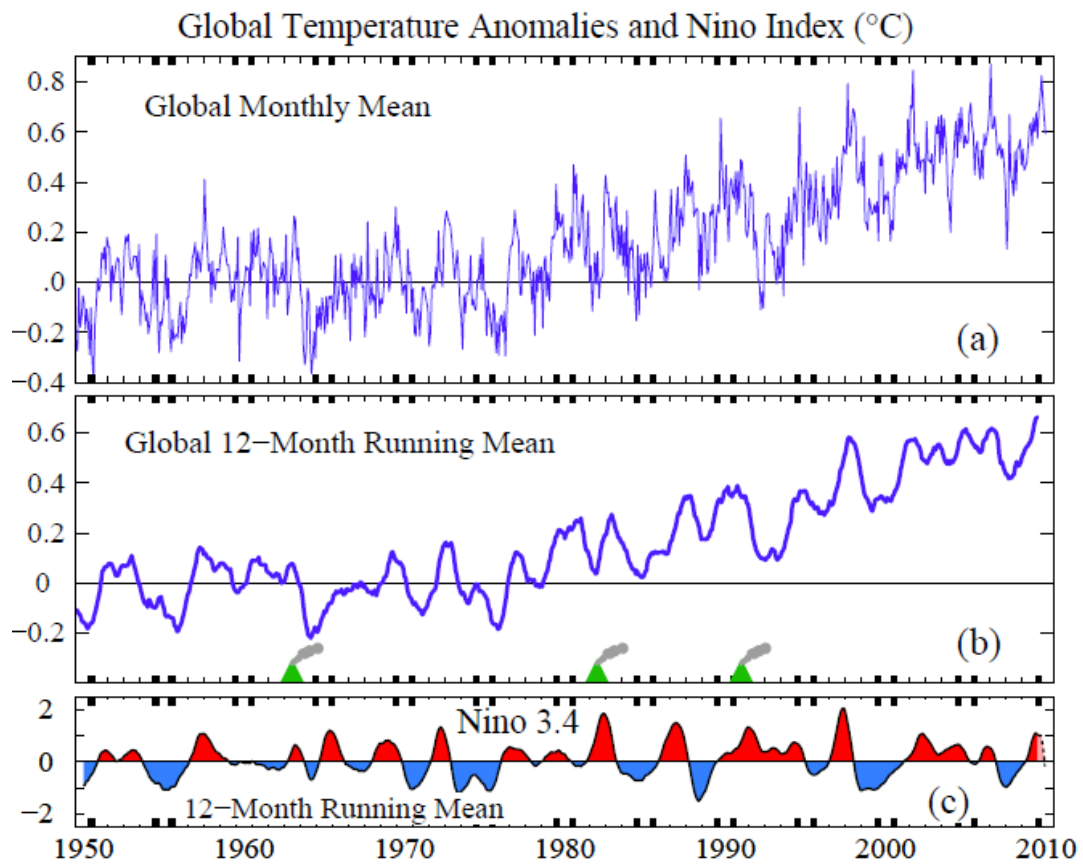


Figure 10. Global monthly and 12-month running mean surface temperature anomalies relative to 1951-1980 base period, and the Nino 3.4 index. Data extend through June 2010.

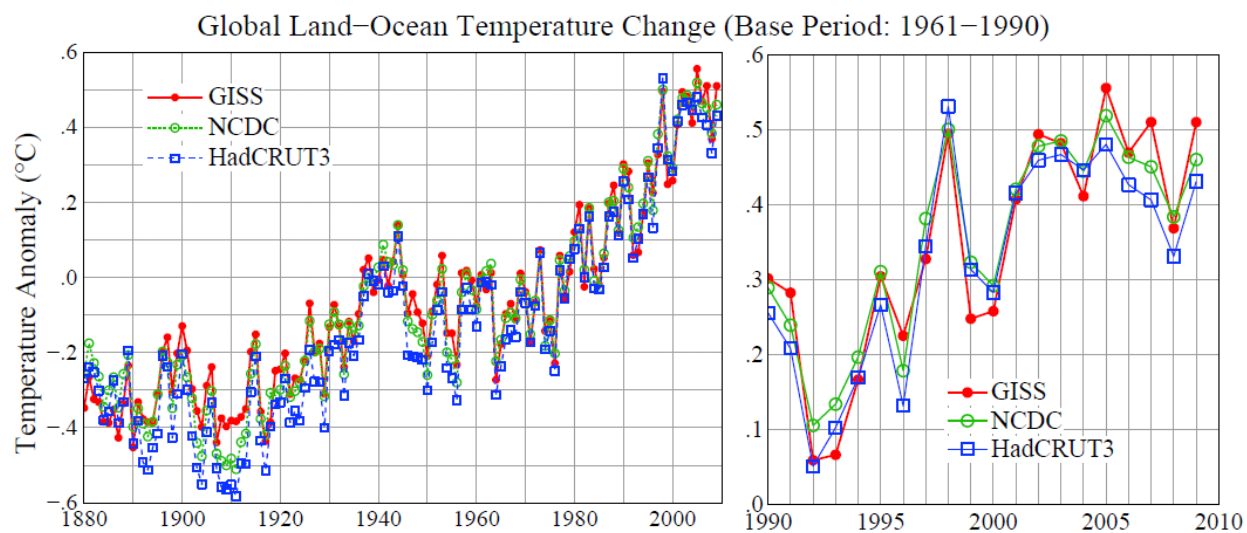


Figure 11. GISS, NCDC and HadCRUT global surface temperature anomalies. Base period is 1961-1990 for consistency with base periods used by NCDC and HadCRUT and for consistency with Figures 12 and 13. The last two decades are expanded on the right side.

Comparisons of GISS and HadCRUT Surface Temperature Anomalies (°C)

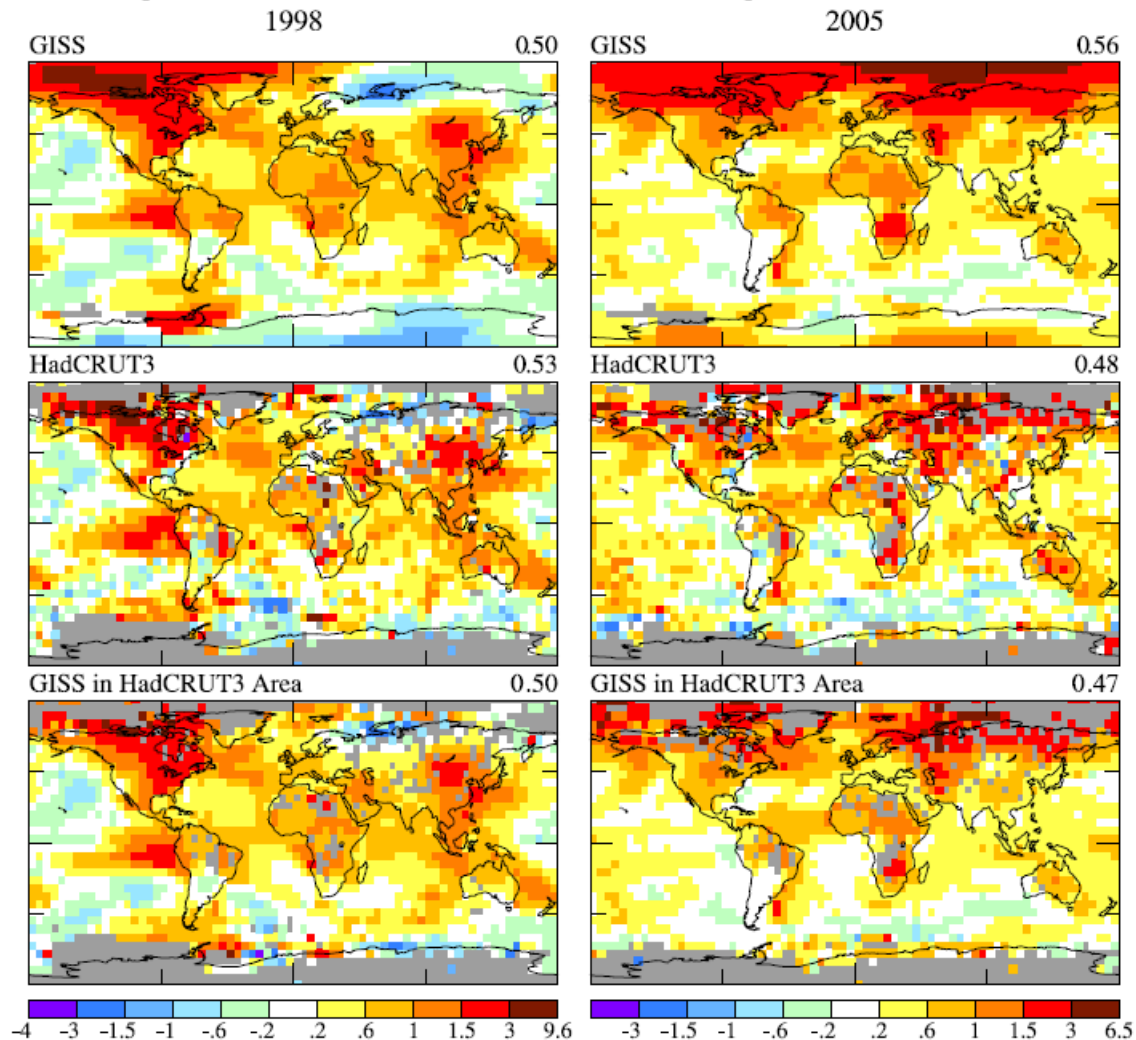


Figure 12. Temperature anomalies (°C) in 1988 (left) and 2005 (right). Top row is GISS analysis, middle row HadCRUT analysis, and bottom row the GISS analysis masked to the same area as the HadCRUT analysis. Areas without data are gray. "Global" means (upper right corner) are averages over area with data. [Base period is 1961-1990 and spatial resolution is 5 by 5 degrees.]

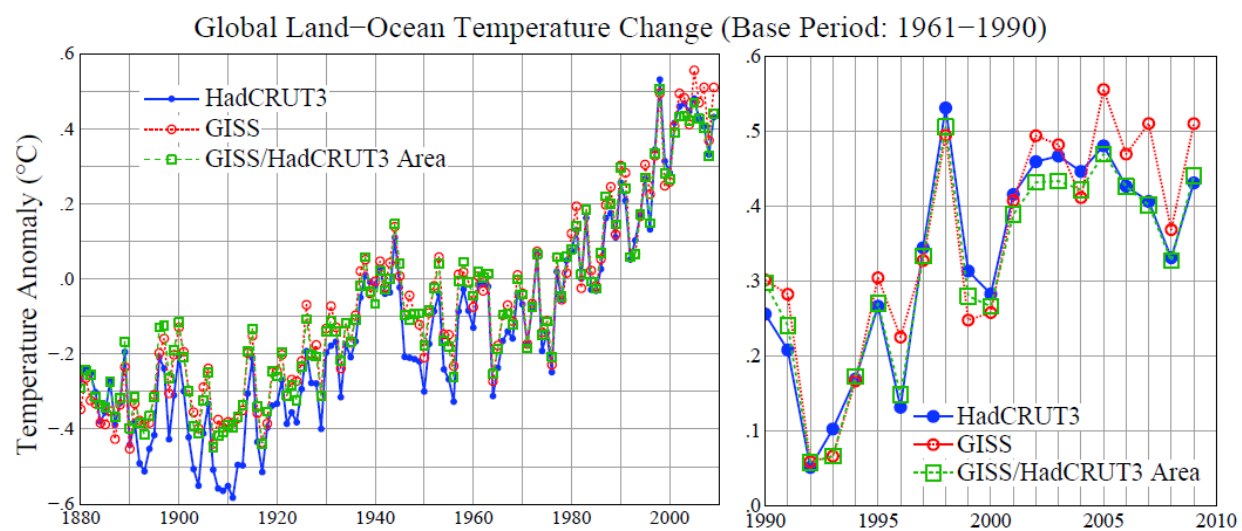


Figure 13. Global surface temperature anomalies ($^{\circ}\text{C}$) relative to 1961–1990 base period for three cases: HadCRUT3, GISS, and GISS anomalies limited to the HadCRUT area.

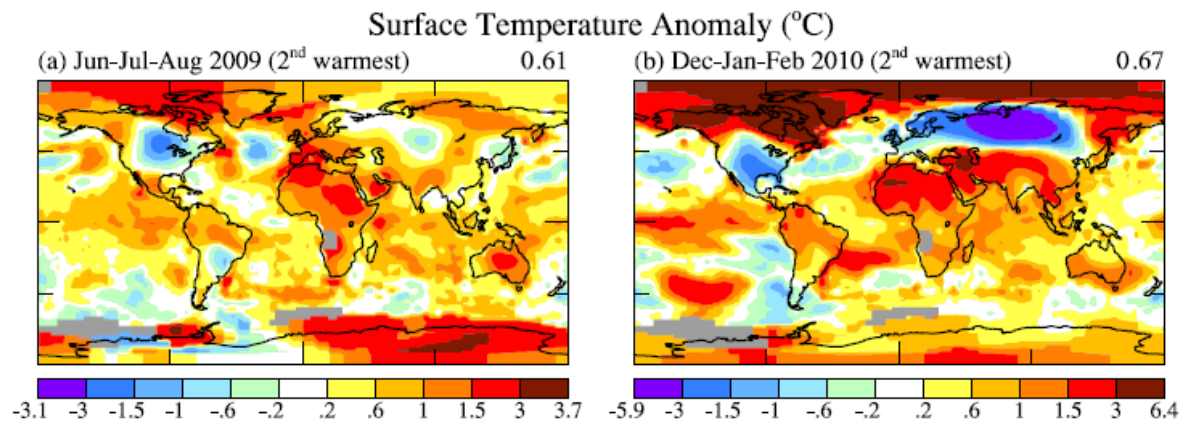


Figure 14. Jun-Jul-Aug 2009 and Dec-Jan-Feb 2010 surface temperature anomalies ($^{\circ}\text{C}$).

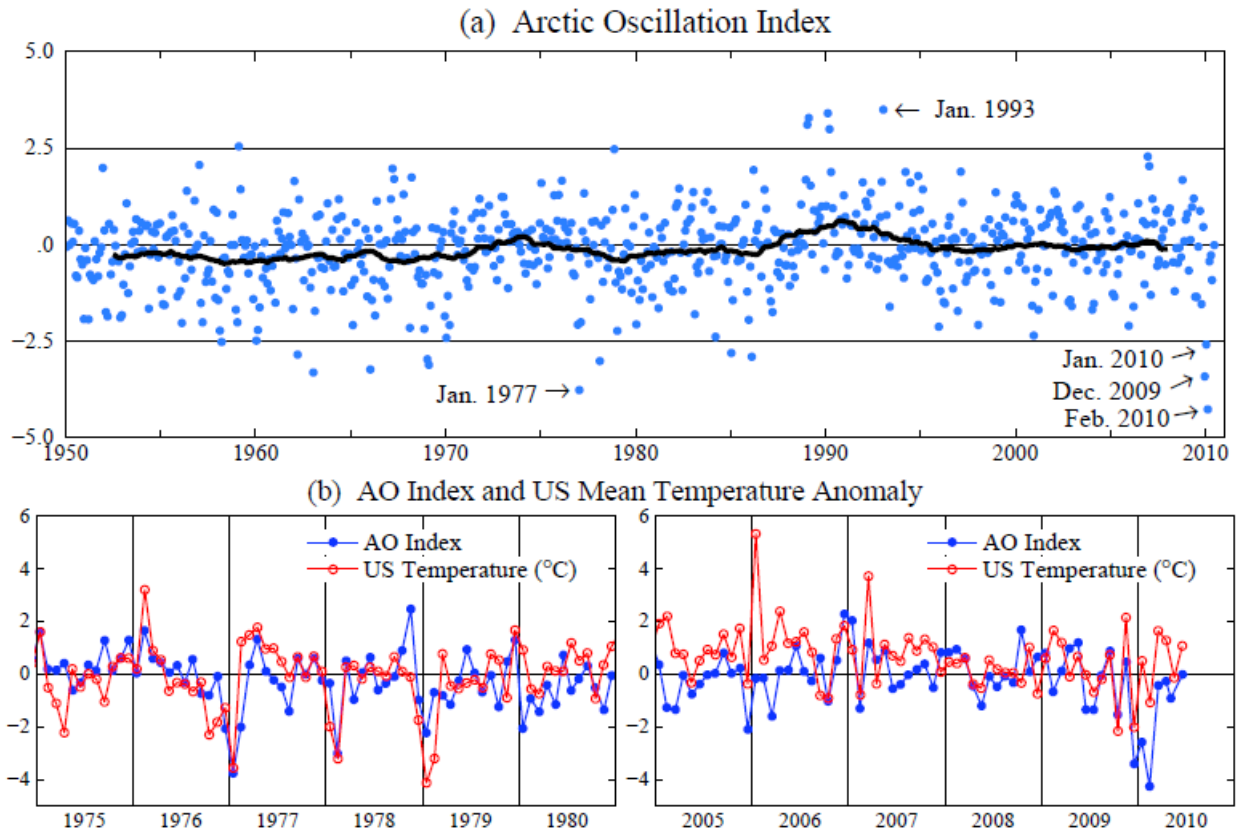


Figure 15. (a) Arctic oscillation (AO) index, data from (http://www.cpc.noaa.gov/products/precip/CWlink/daily_ao_index/monthly.ao.index.b50.current.ascii.table). Blue dots are monthly means and black curve is the 60-month (5-year) running mean. (b) AO index at higher temporal resolution and temperature anomaly in contiguous 48 United States.

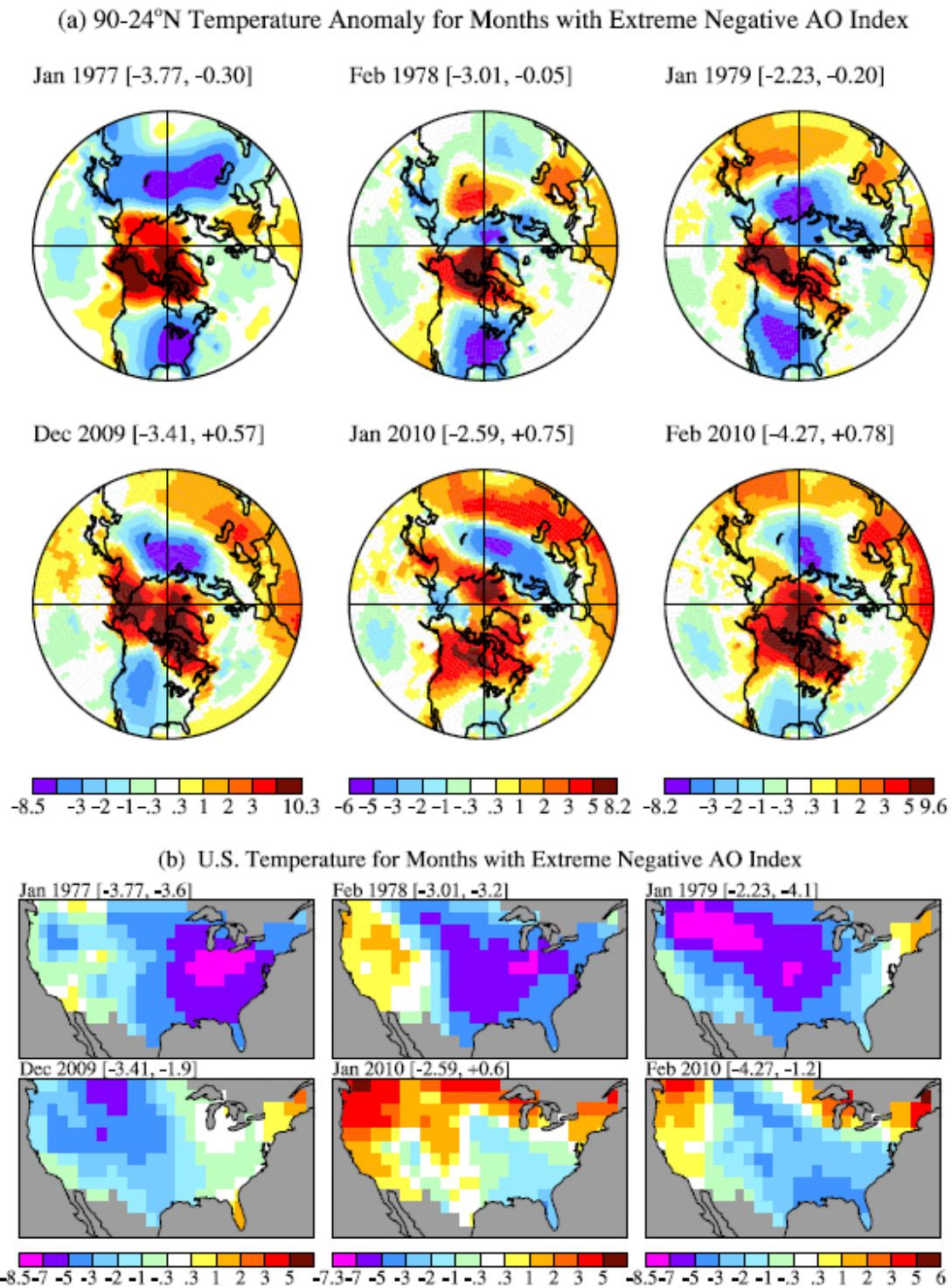


Figure 16. Temperature anomaly from GISS analysis for months with extreme negative AO index. Parenthetical numbers are the AO index and the mean temperature anomaly in °C for (a) 24-90°N and (b) the contiguous 48 states.

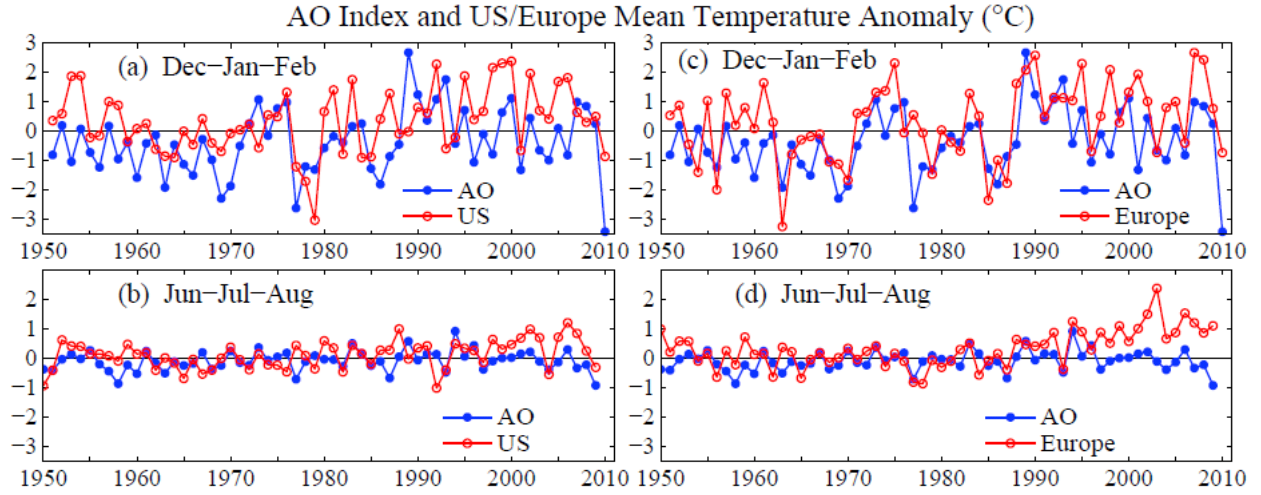


Figure 17. Arctic Oscillation index and United States (48 states) surface temperature anomaly for December-January-February (a) and June-July-August (b). Temperature zero line is 1951-1980 mean. Right side: same quantities for European region defined by 36N-70N and 10W-30E.

1950-2009 Seasonal Surface Temperature Changes Based on Linear Trends (°C)

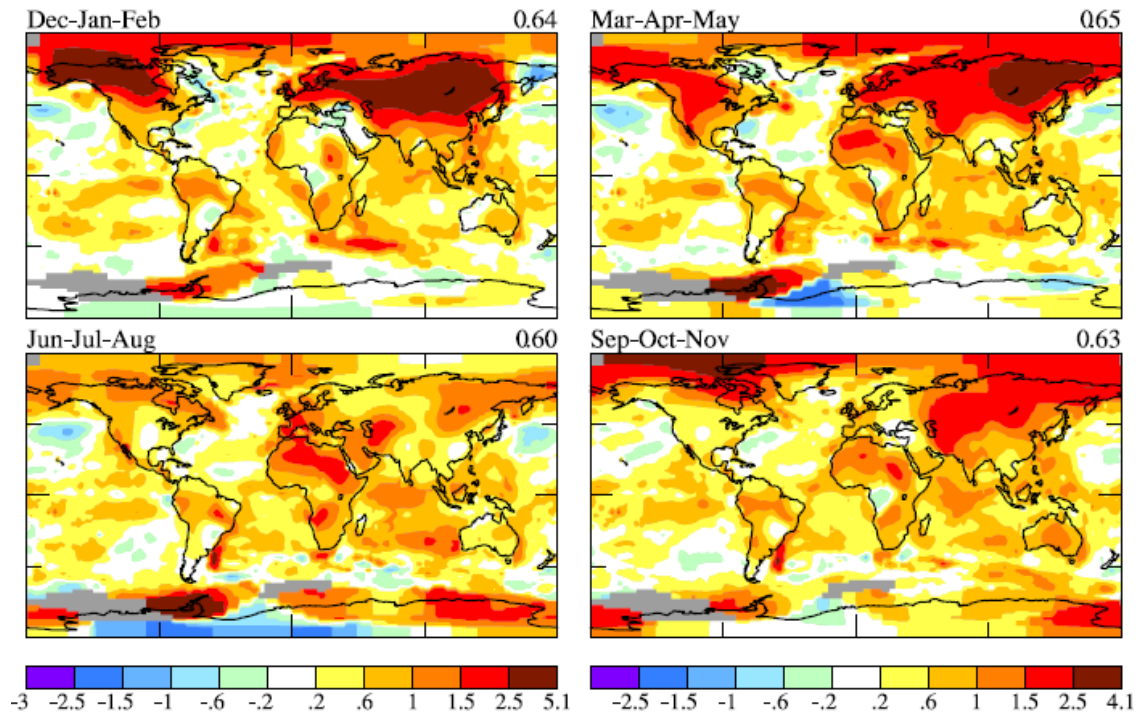


Figure 18. Global maps 4 season temperature anomaly trends (°C) for period 1950-2009.

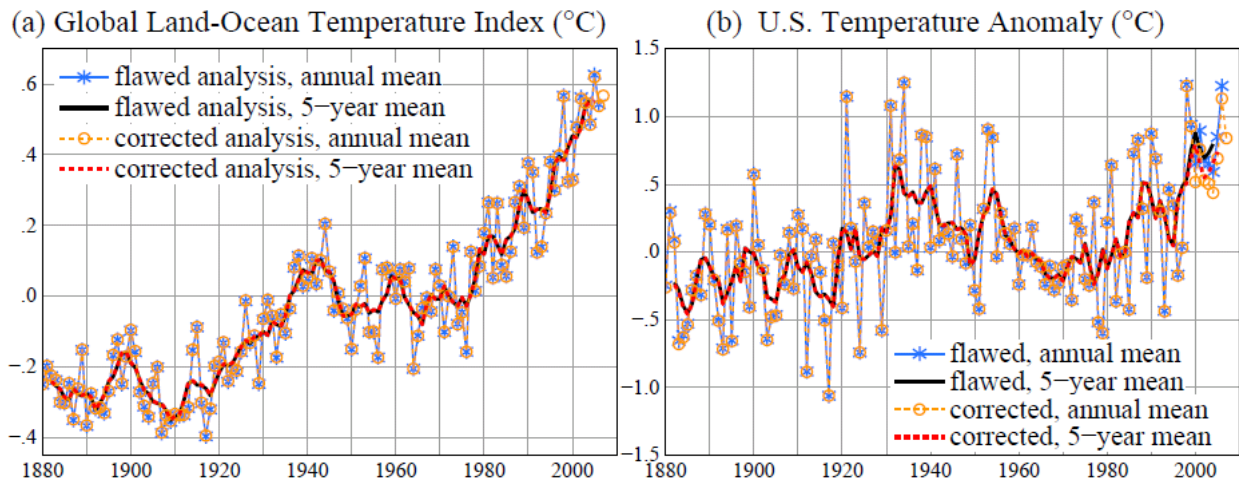


Figure 19. (a) Global and (b) United States analyzed temperature change in the GISS analysis before and after correction of a data flaw in 2007. Results are indistinguishable except post-2000 in the United States.

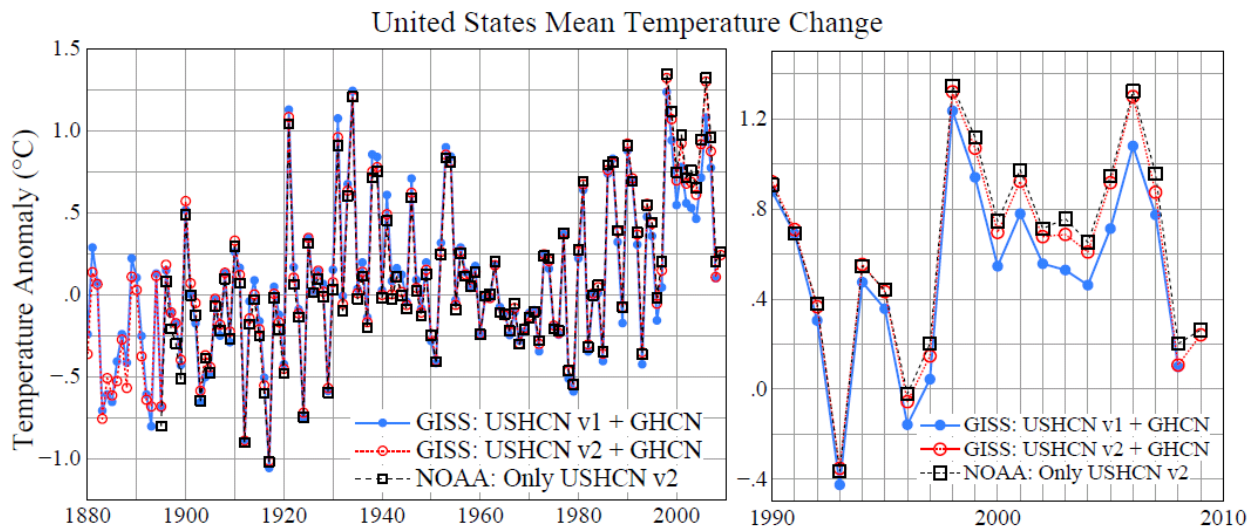


Figure 20. GISS analysis of United States temperature change (48 states) using USHCN.v1 (version 1) and USHCN.v2 (v2 became available from NCDC in July 2009; v2 replaced v1 in GISS analysis in November 2009) and NCDC analysis [<http://lwf.ncdc.noaa.gov/oa/ncdc.html>] for USHCN.v2. NCDC analysis uses only USHCN stations, while GISS analysis includes use of non-USHCN GHCN stations, with nighttime adjustments of all urban and periurban stations.

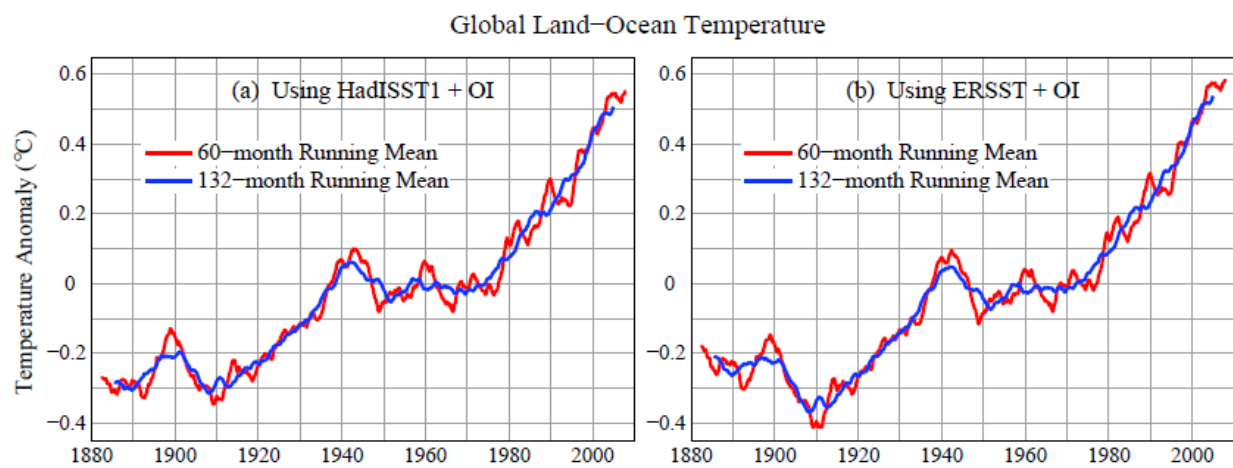
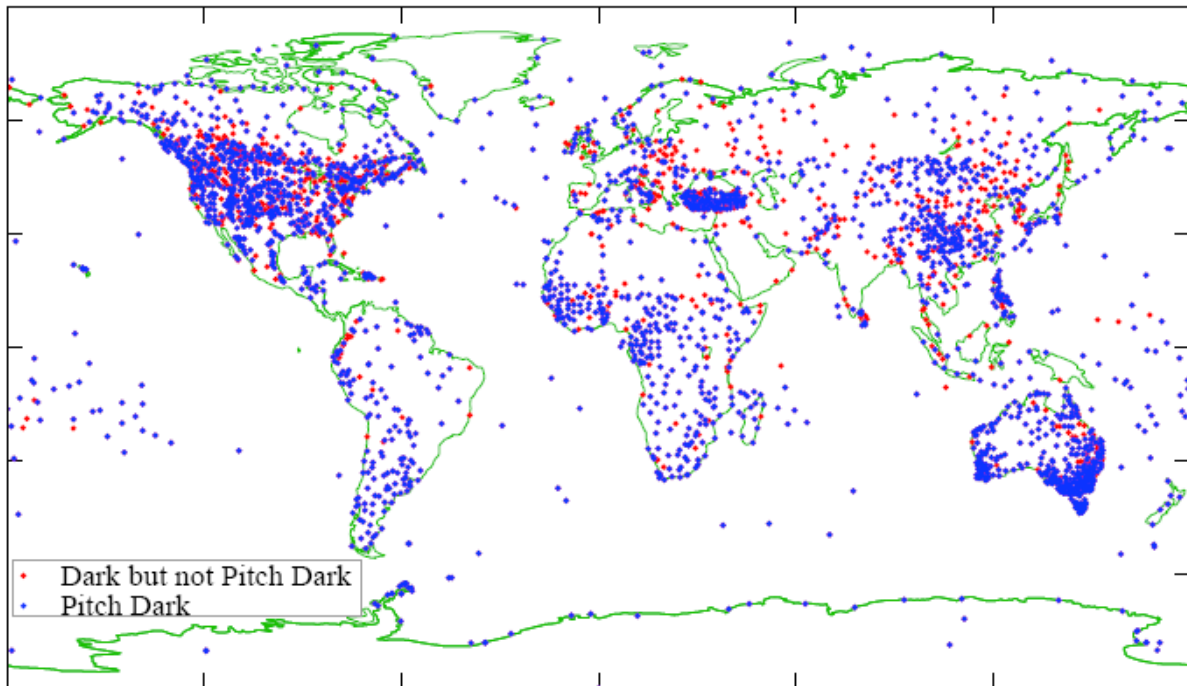


Figure 21. 60-month and 132-month running means using data through June 2010 for two alternative choices for the ocean data set.

(a) Dark and Pitch Dark GHCN Stations



(b) 1900-2009 Annual-Mean Surface Temperature Change ($^{\circ}\text{C}$)

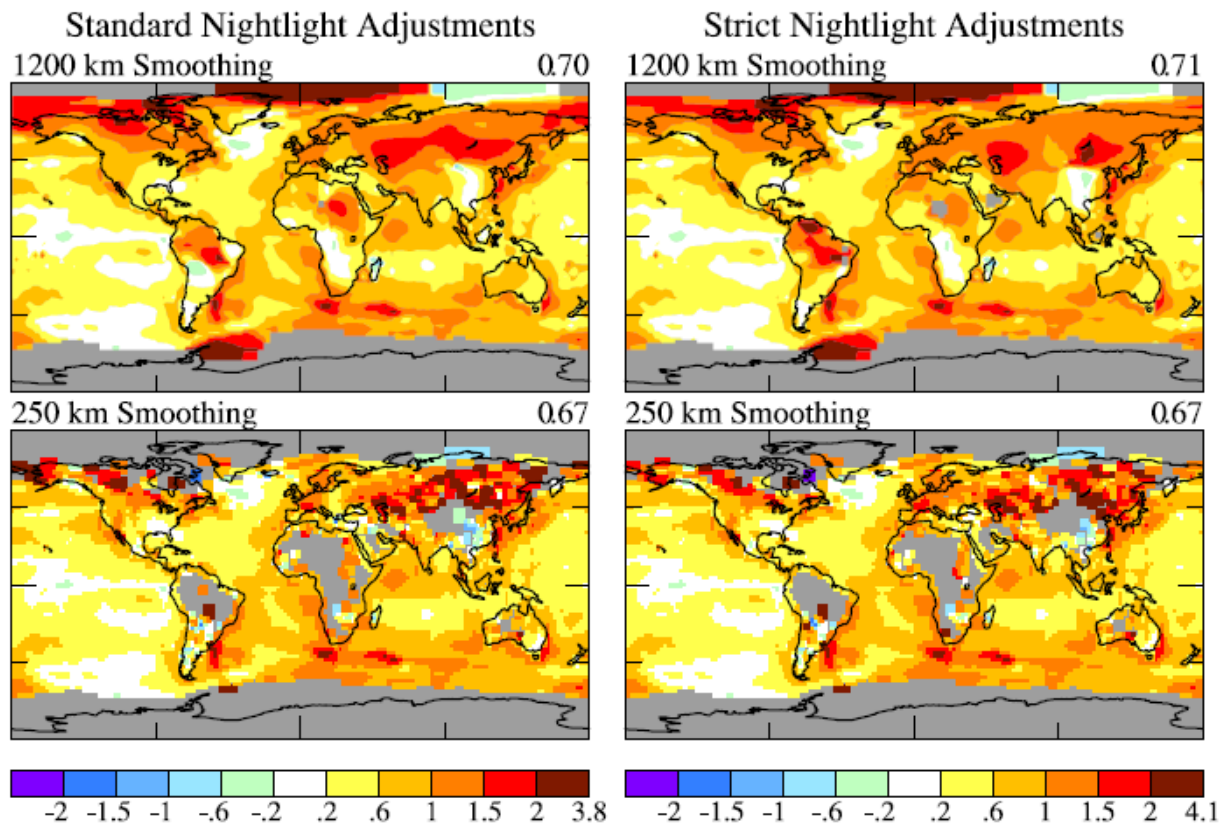


Figure A1. (a) Blue dots are stations with nightlight radiance less than $1 \mu\text{W}/\text{m}^2/\text{sr}/\mu\text{m}$, while red dots are stations with radiance between 1 and $32 \mu\text{W}/\text{m}^2/\text{sr}/\mu\text{m}$, (b) global maps of 1900-

2009 surface temperature change comparing results for standard nightlight adjustment and the strict nightlight adjustment (only pitch dark stations used to establish long-term trends).

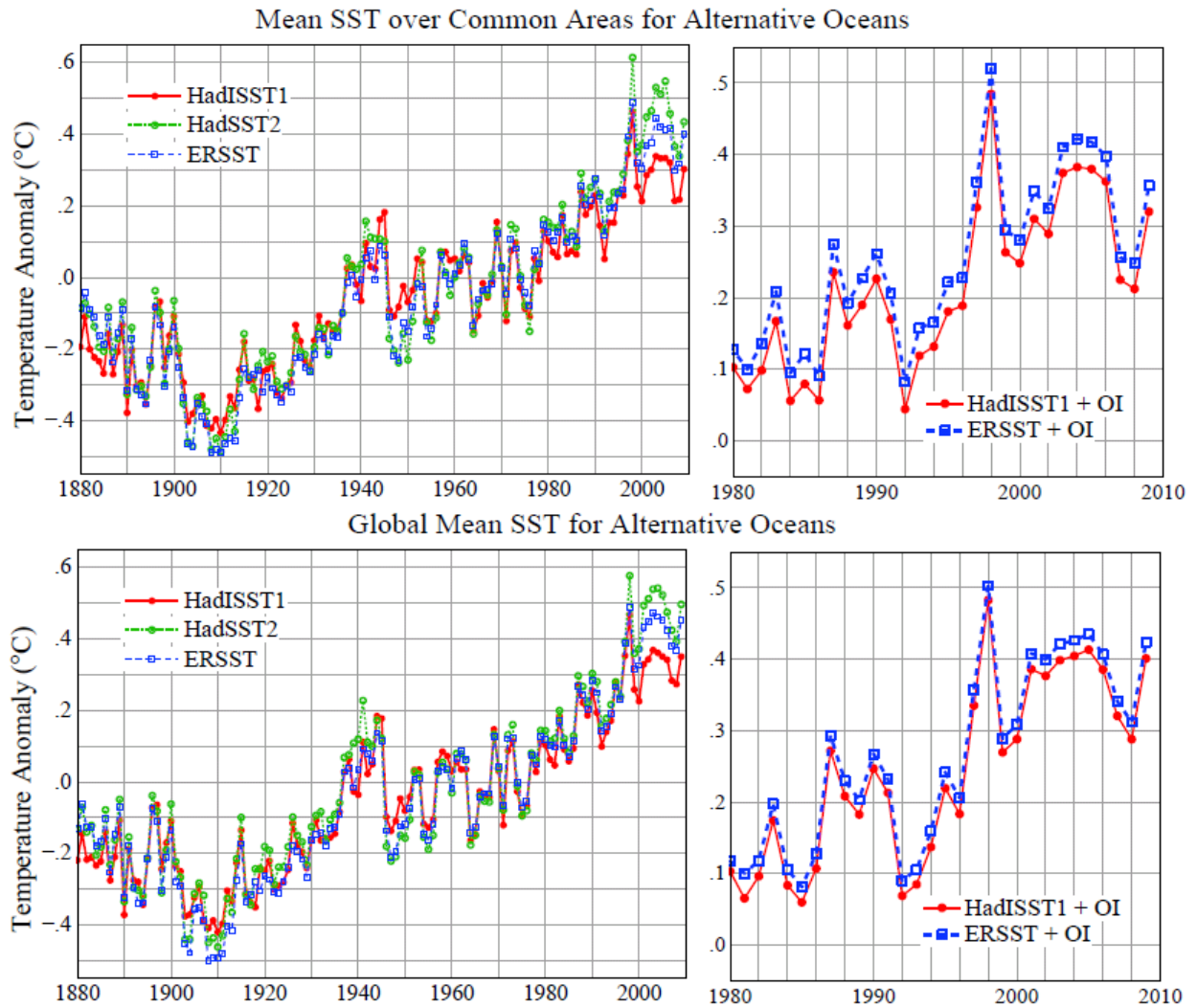


Figure A2. Ocean surface temperatures for alternative data sets, the temperature zero point in all cases being the 1951-1980 (base period) mean. In the graphs on the right OI is concatenated by equating its mean for 1982-1992 with the mean of the appended data set for the same period. The top row uses only areas that have data in all data sets; the temperature anomaly is averaged over this area, so it is not a true global ocean mean. The bottom row is the global ocean mean obtained for each of these data sets, as described in the text.

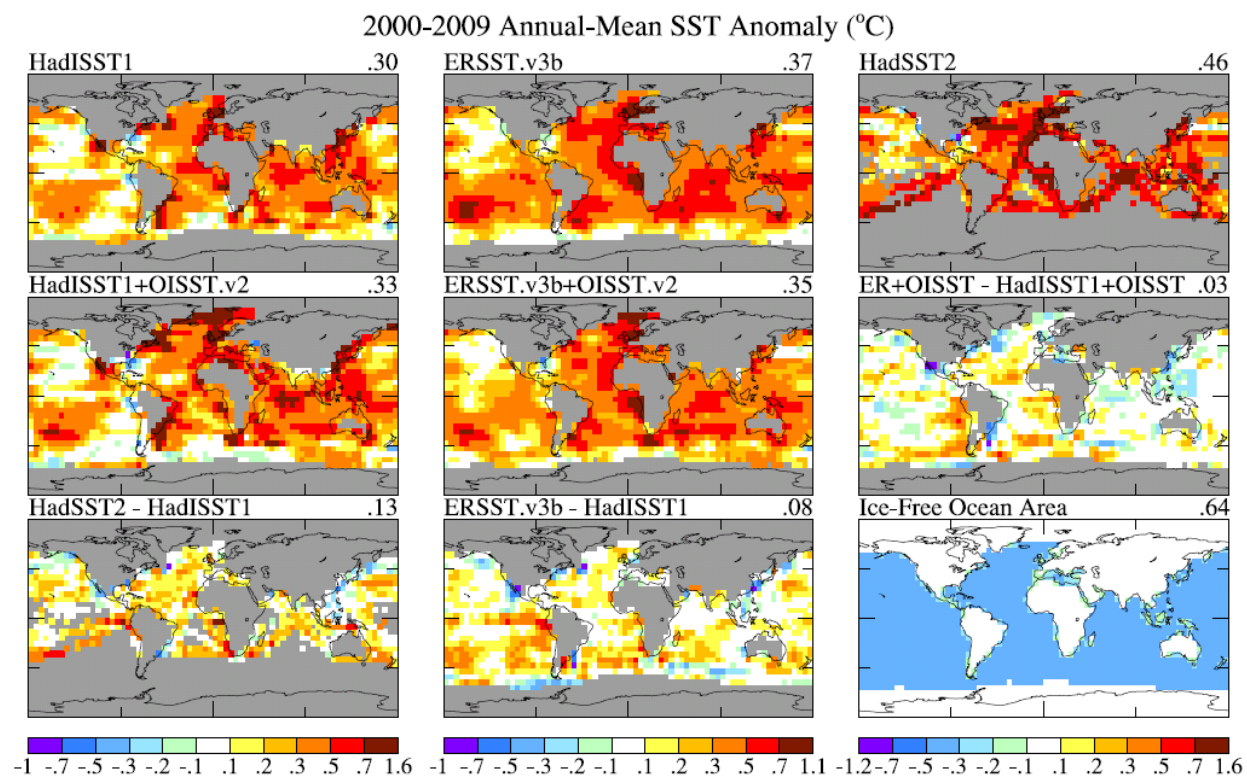


Figure A3. SST 2000-2009 anomalies for several data sets and their differences. The final map shows the permanently ice-free area used in computing the global ocean mean temperature.

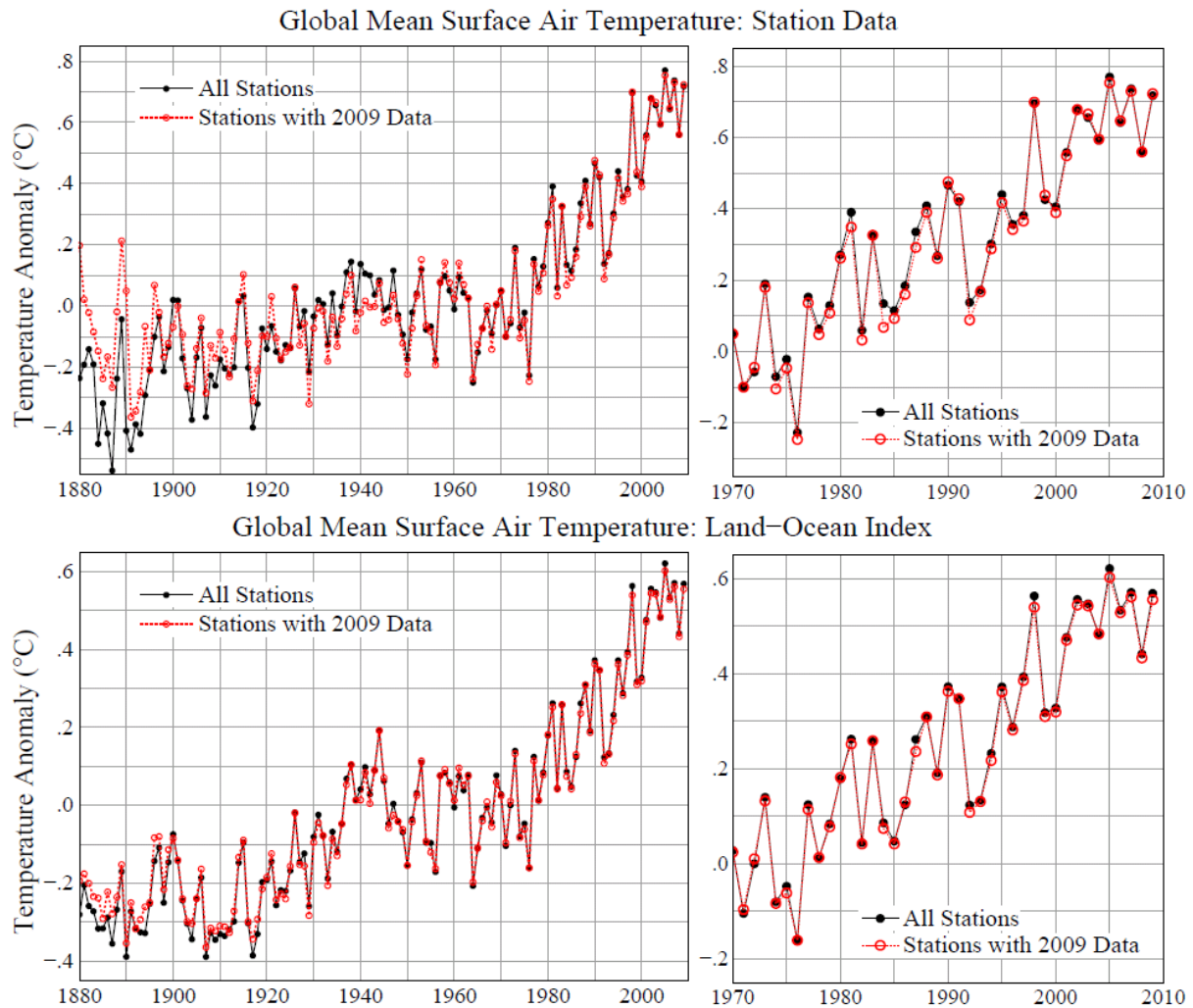


Figure A4. Comparison of global temperatures in the GISS analysis when all stations are used and when only the stations reporting in 2009 are employed in the analysis.