Workshop on creating surface temperature datasets to meet 21st Century challenges

The USHCN and GHCN Experience: Lessons Learned

Matthew Menne

Claude Williams

Jay Lawrimore

NOAA/National Climatic Data Center

Asheville, North Carolina, USA

Topics

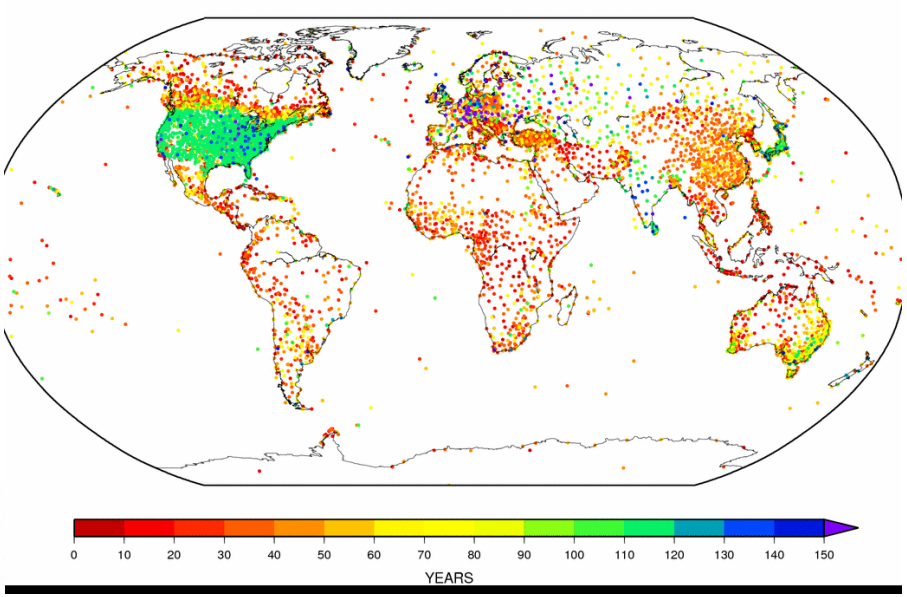
- The Datasets
- Seven Lessons Learned
- A few Take Away Comments

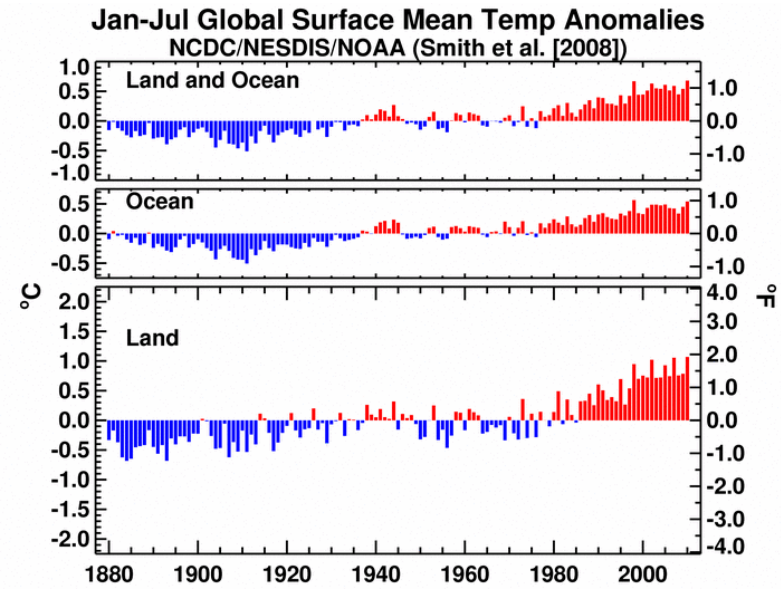
The U.S. Historical and Global Historical Climatology Networks (USHCN and GHCN)

GHCN Monthly

- Version 1 released in 1992
- Version 2 released in 1997
- Version 3 “Beta” to be released this month
  - 7000+ stations with mean monthly temperature records

Number of years of data for each station in GHCN Monthly mean temperature dataset



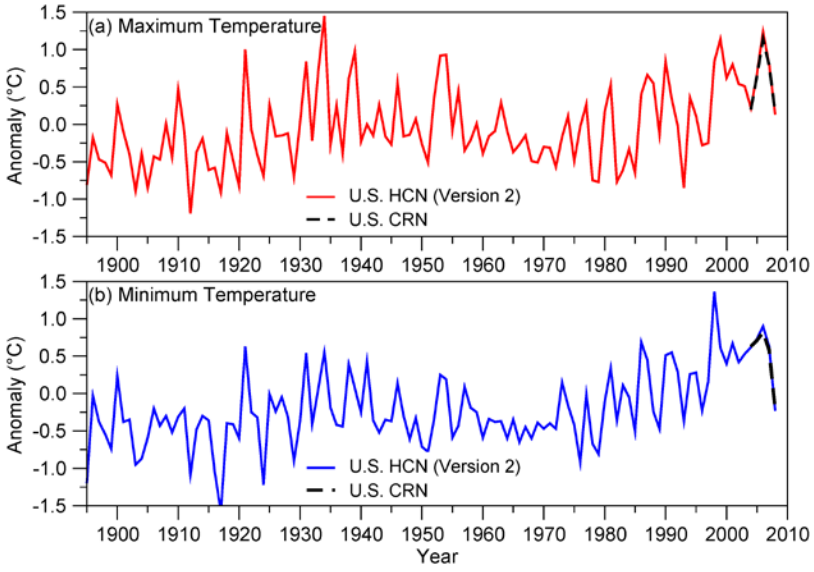


- Version 3 Beta is based on the quality control and homogenization approach used to produce the USHCN Version 2 (but without station histories for stations outside the USA)

USHCN V1

Released in 1987, with subsequent revisions

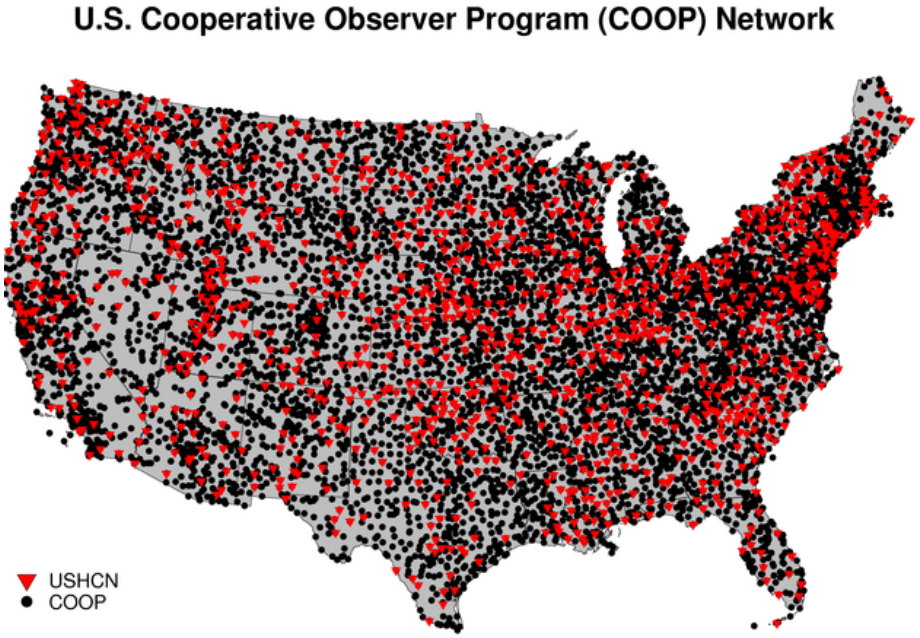
- Time of observation bias (Karl et al. 1986)
- Changes documented in the station history archives (Karl and Williams 1987)
- Urbanization (Karl et al. 1988)
- LiG to MMTS instrument change (Quayle et al. 1991)



USHCN V2

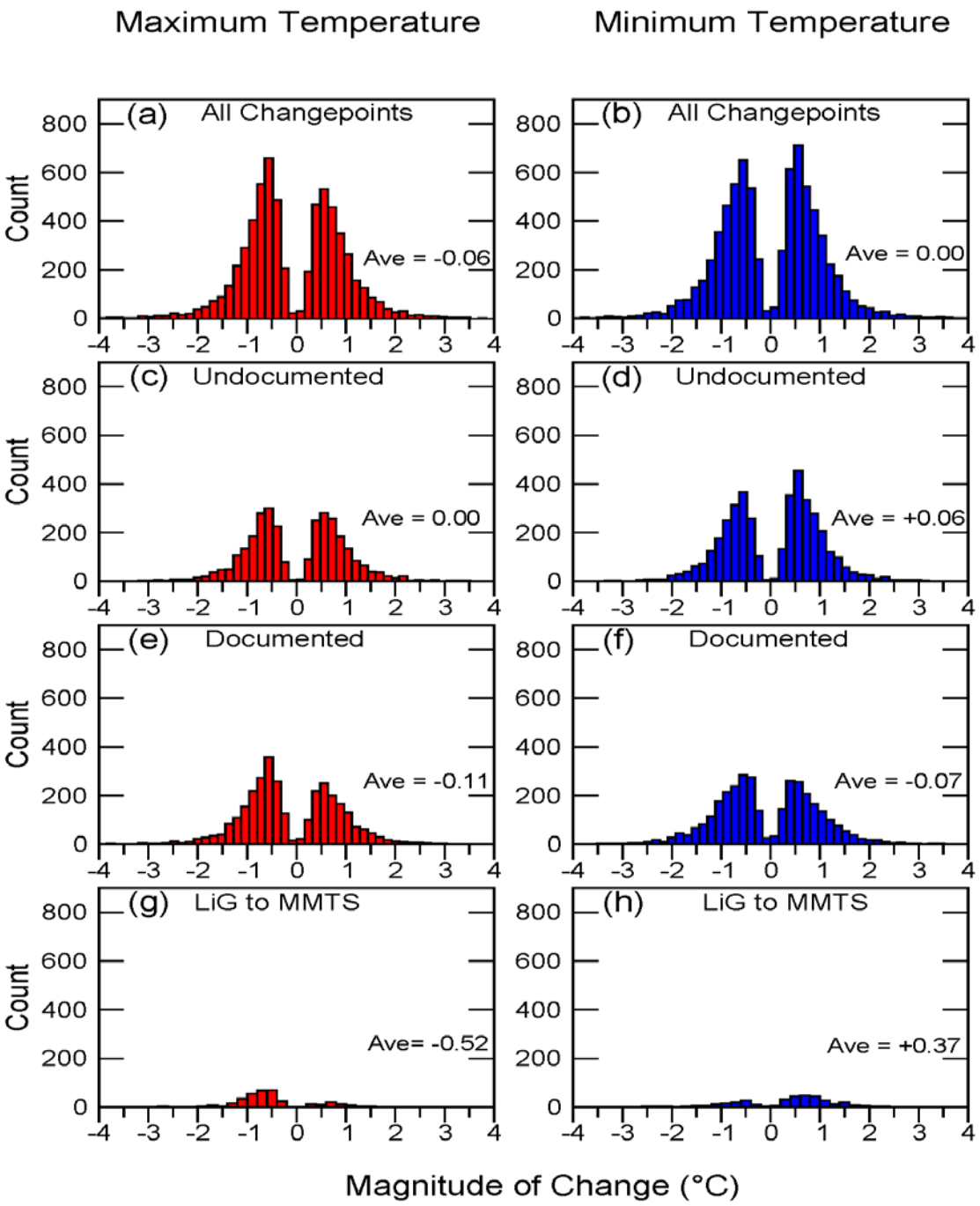
Released in 2009

- Time of observation bias (Karl et al. 1986)
- Documented and undocumented station changes (Menne and Williams 2009)



Lessons Learned

1) Undocumented changes can be as prevalent as documented changes even when extensive (digitized) metadata are available

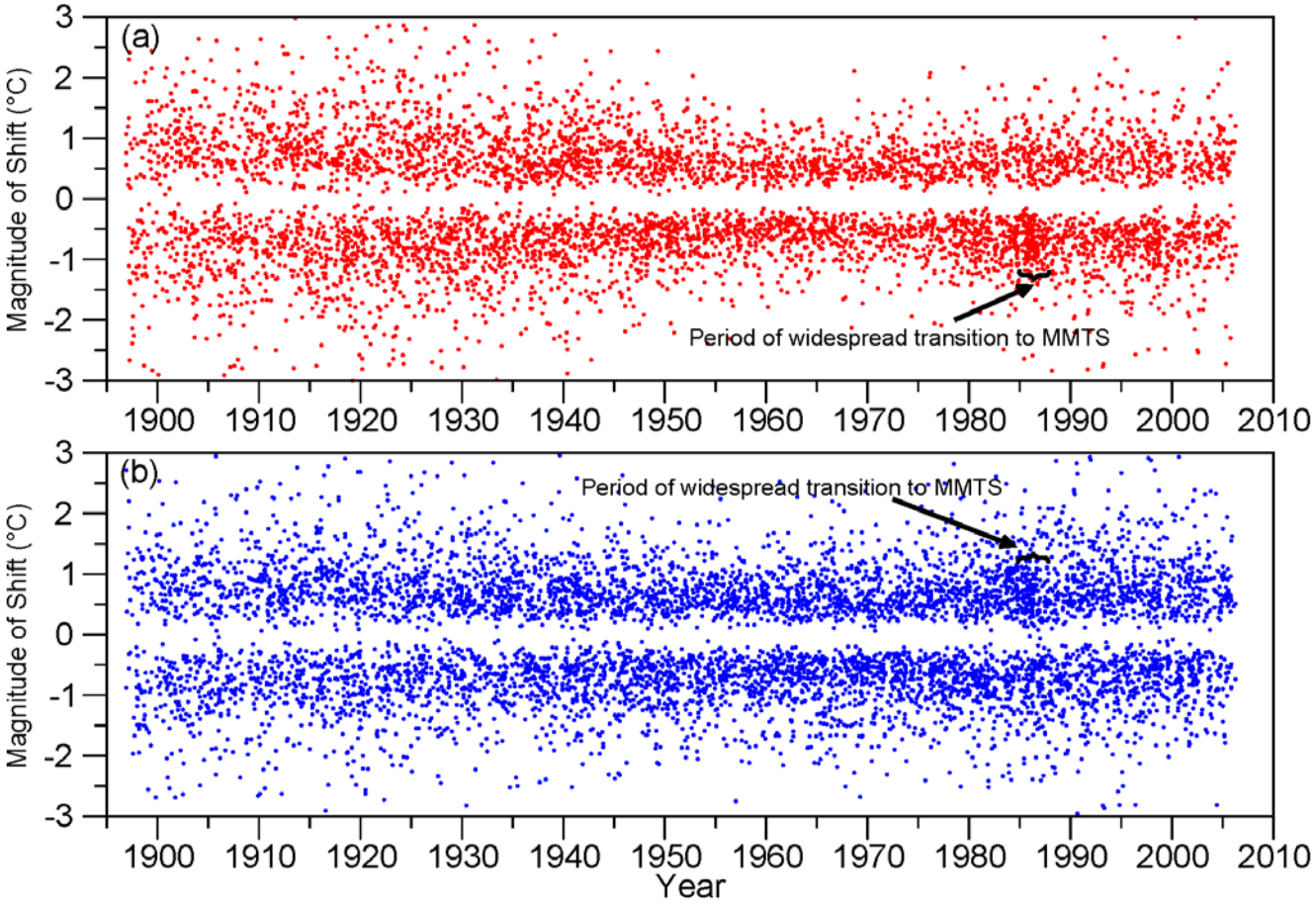


Summary shifts in USHCN mean monthly maximum and minimum temperature series. A negative value indicates that the inhomogeneity led to a decrease in the mean level of the series relative to preceding values.

(Source Menne et al. 2009)

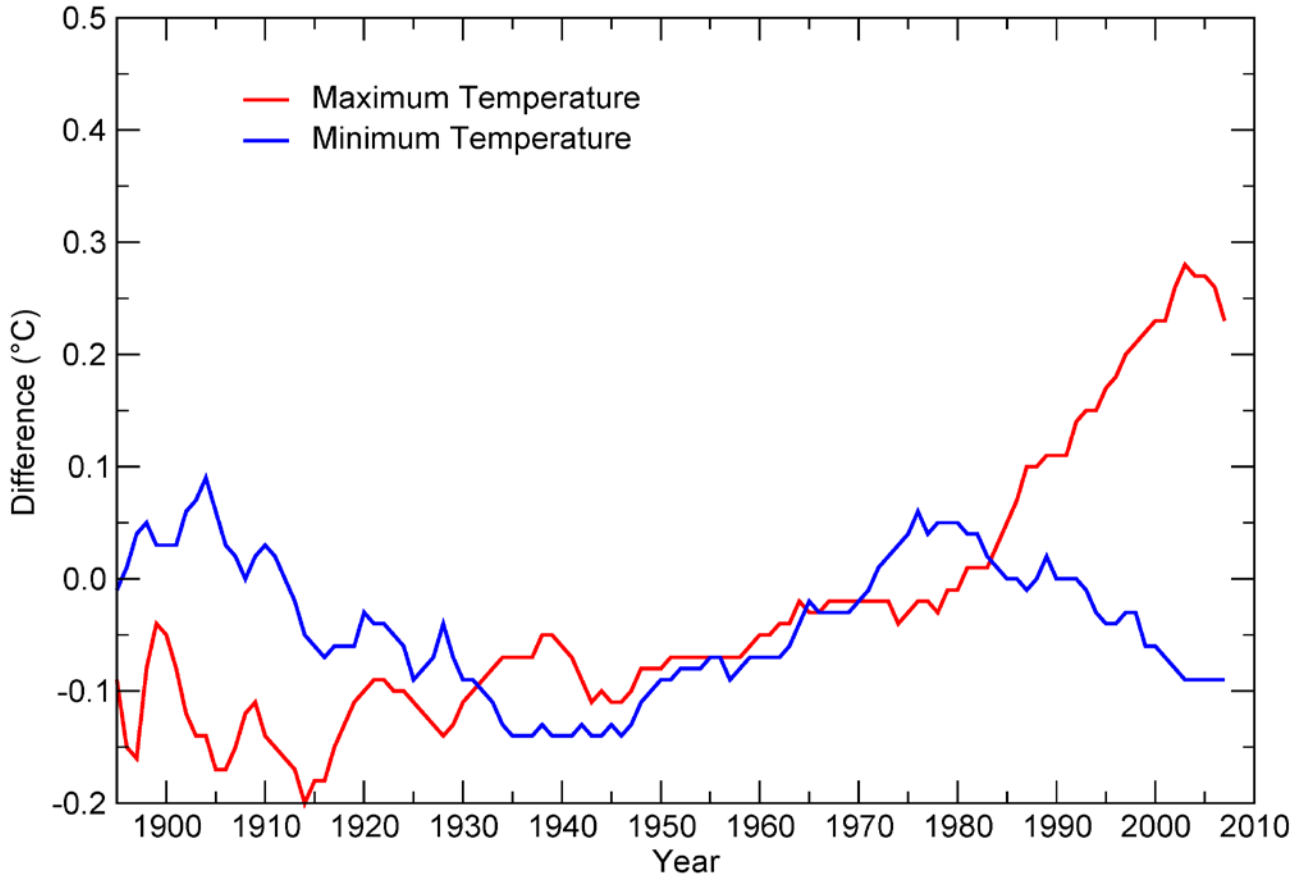
2) Collectively station changes often have nearly random impacts, but even slight deviations from random matter greatly

Timing and magnitude of shifts detected in the USHCN temperature series

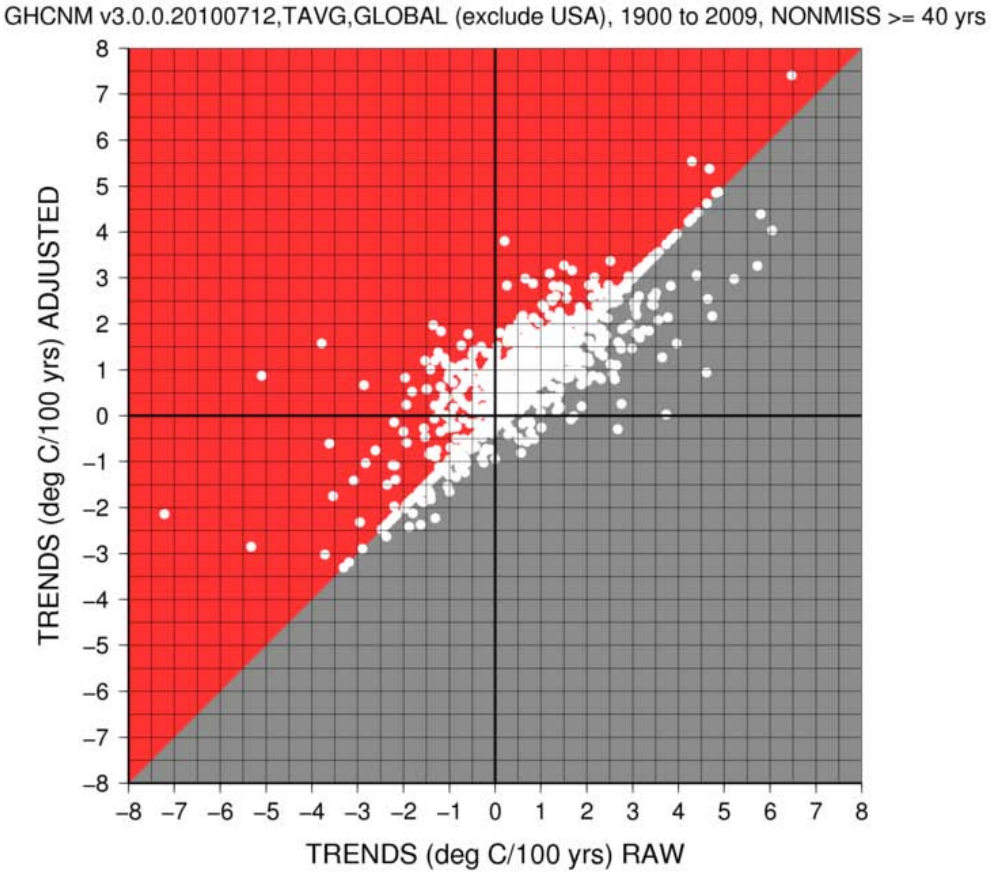


Detected shifts in U.S. HCN mean monthly maximum and minimum temperature series. A negative value indicates that the inhomogeneity led to a decrease in the mean level of the series relative to preceding values.

Average annual difference over the United States between the fully adjusted (homogenized) USHCN temperature data and the data adjusted only for the time of observation bias



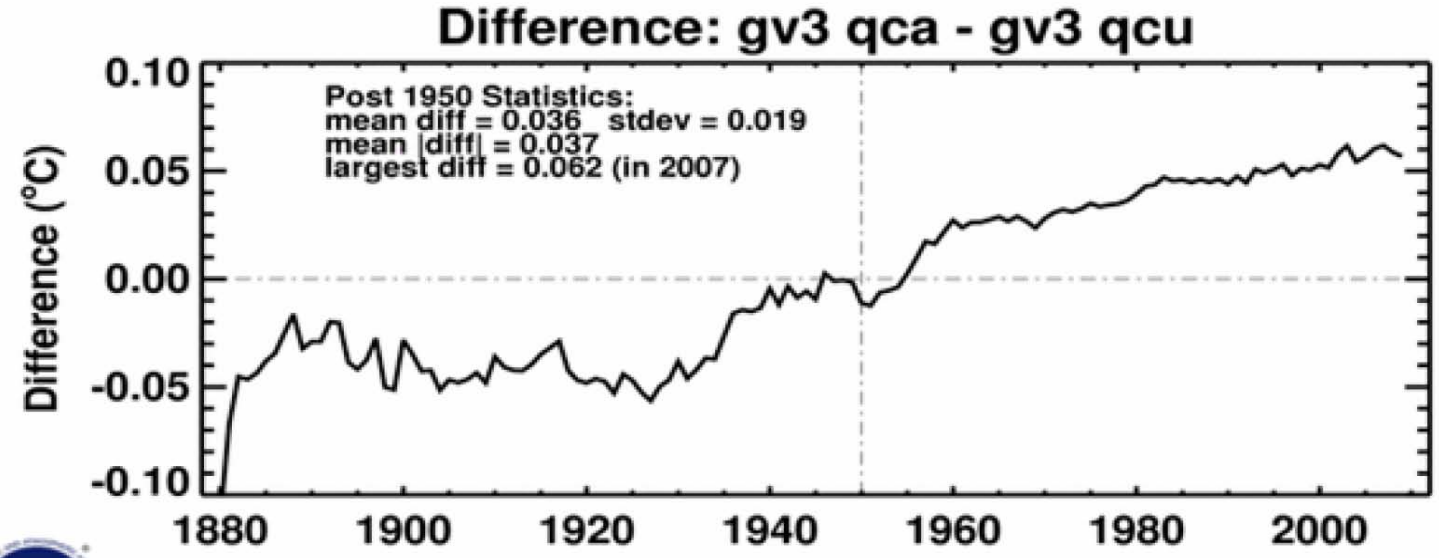
Outside of the USA ~60% of the GHCN Version 3 average temperature trends are larger following homogenization



Average difference between adjusted and unadjusted GHCN Version 3 stations

Trends (°C/decade)

Period	gv3 qca	gv3 qcu	diff
1900-2009	0.086	0.074	0.012
1950-2009	0.186	0.176	0.010
1979-2009	0.290	0.283	0.007

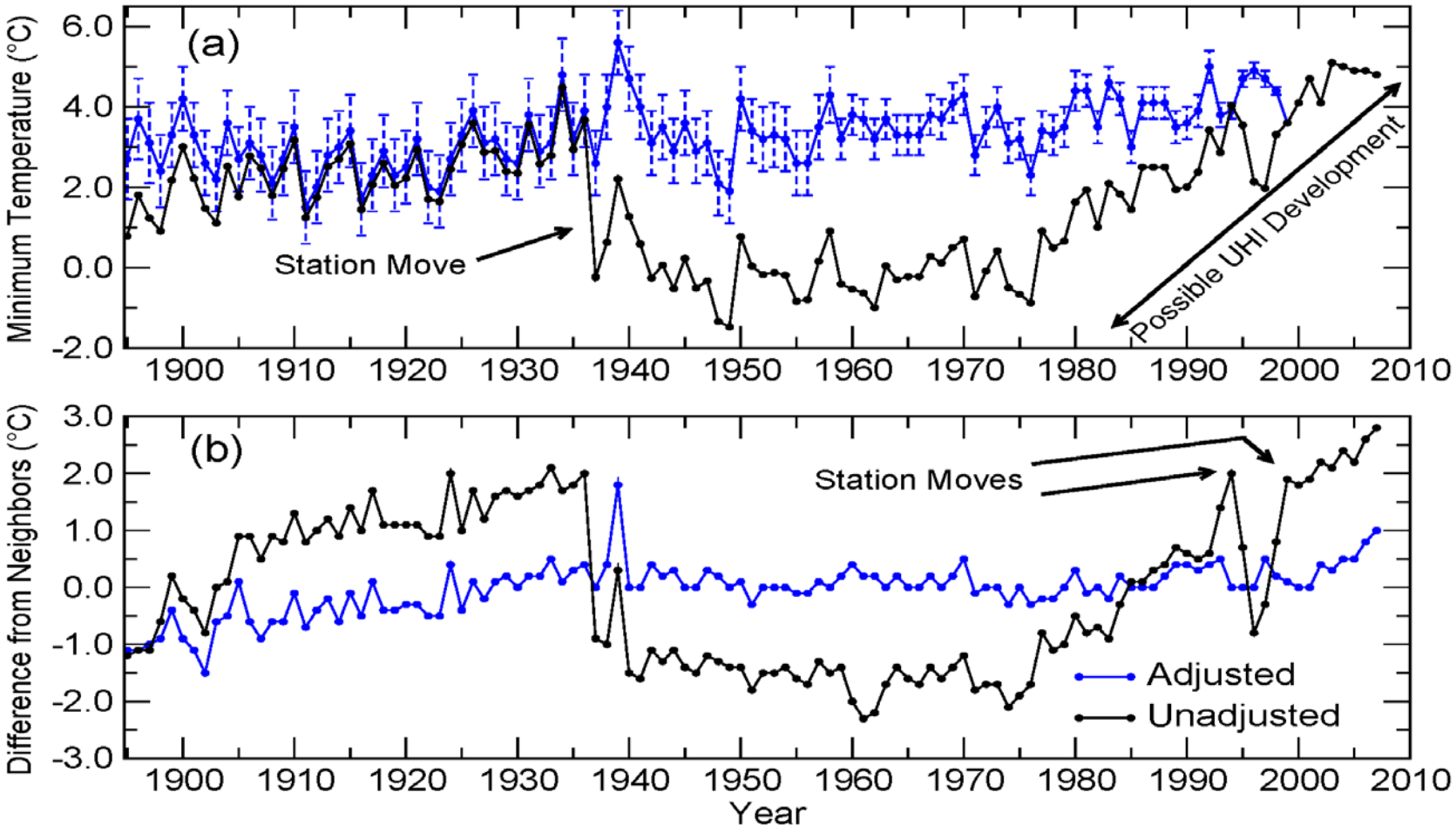


qca=adjusted
qcu=unadjusted

NCDC/NESDIS/NOAA

3) There is a need to identify gradual as well as abrupt changes in bias (but it is may be problematic to adjust for abrupt changes only)

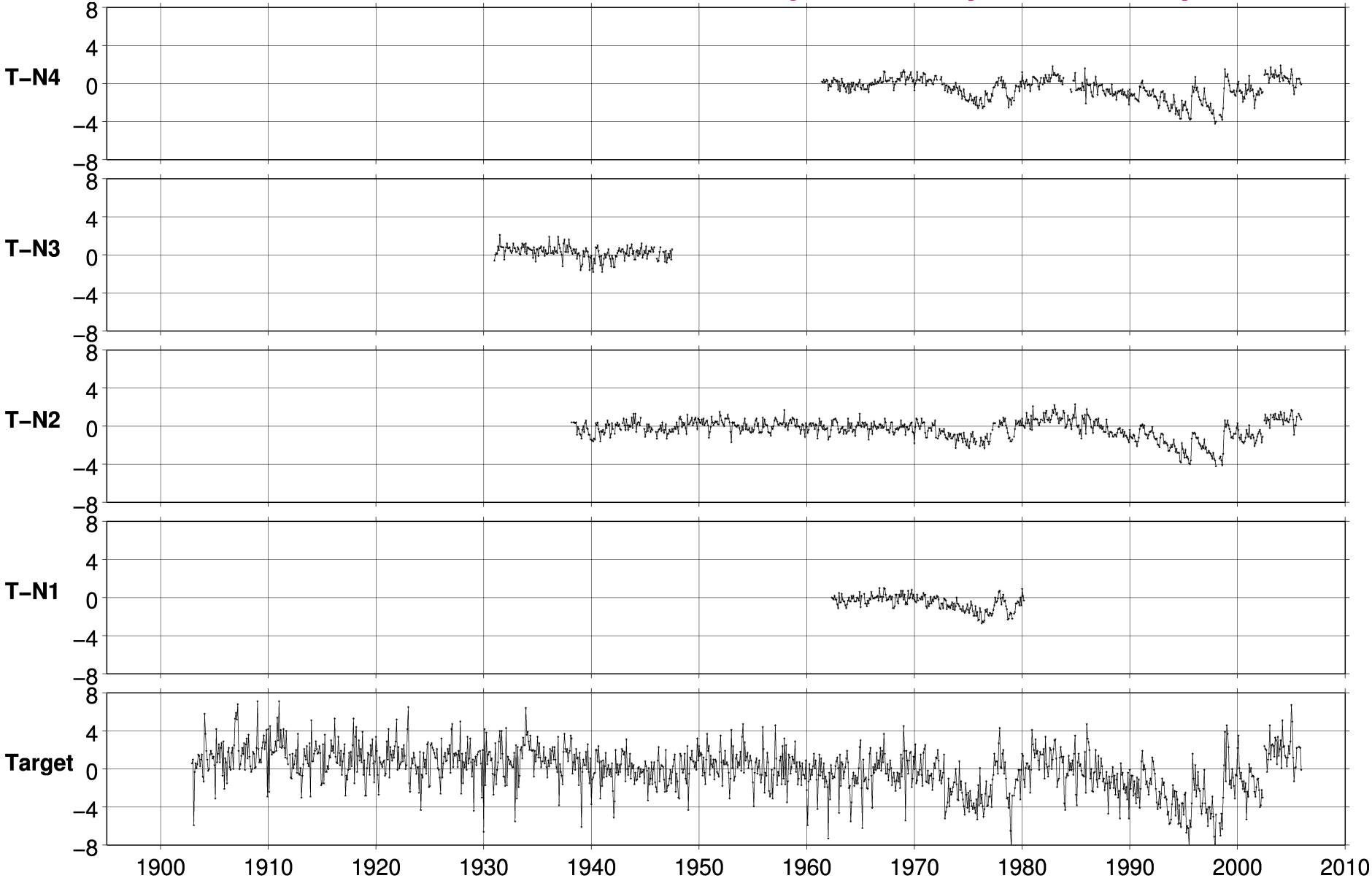
Annual Average Minimum Temperature at Reno, Nevada USA



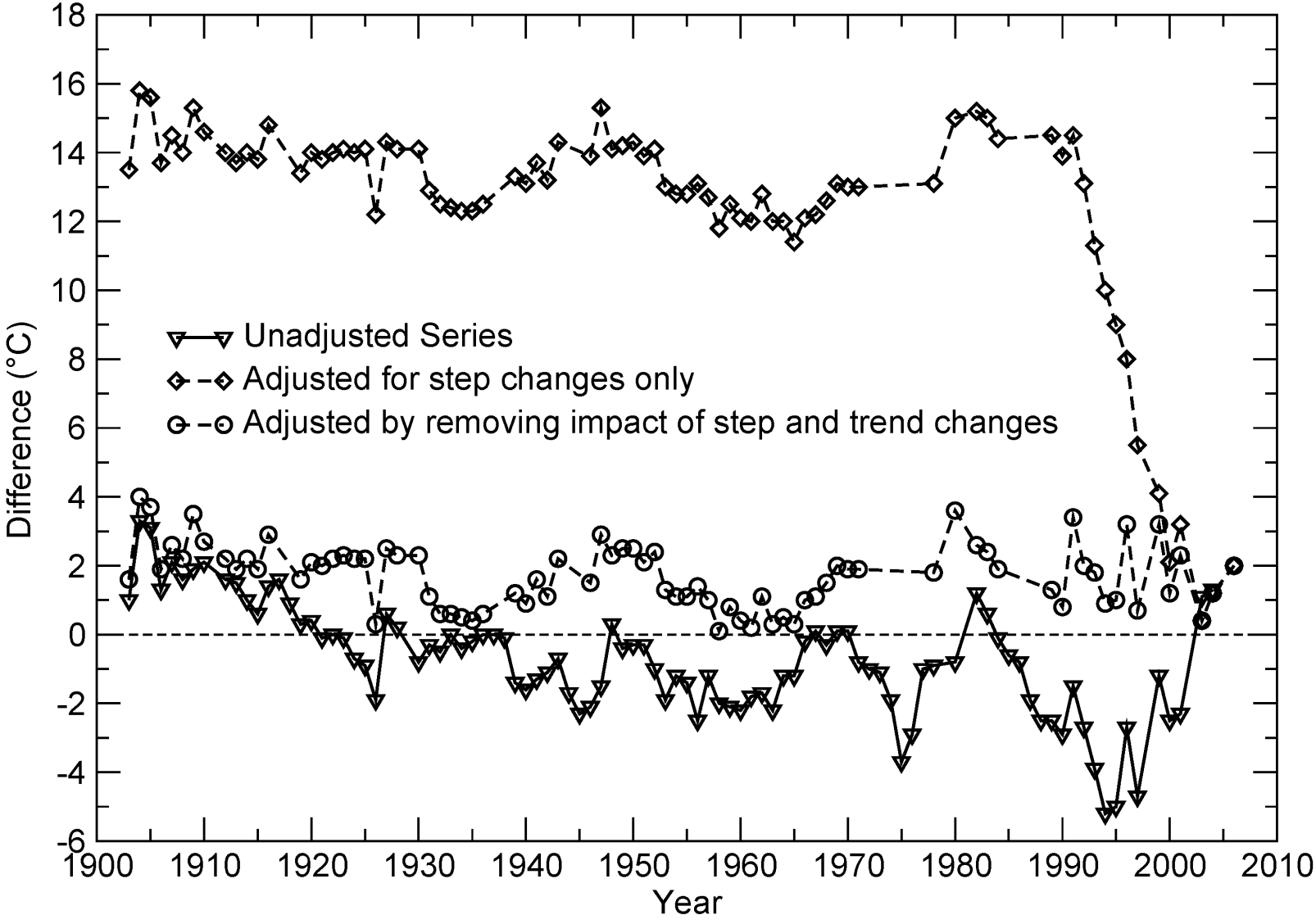
(a) Mean annual unadjusted and fully adjusted minimum temperatures at Reno, Nevada. Error bars indicate the magnitude of uncertainty in the adjustments (±1 standard error);

(b) Difference between minimum temperatures at Reno and the mean from its 10 nearest neighbors.

Mean monthly minimum temperature anomalies (°C) at Cheesman, Colorado USA (target) and differences between monthly temperature anomalies at Cheesman and four correlated series from nearby stations (T-N1 to T-N4)



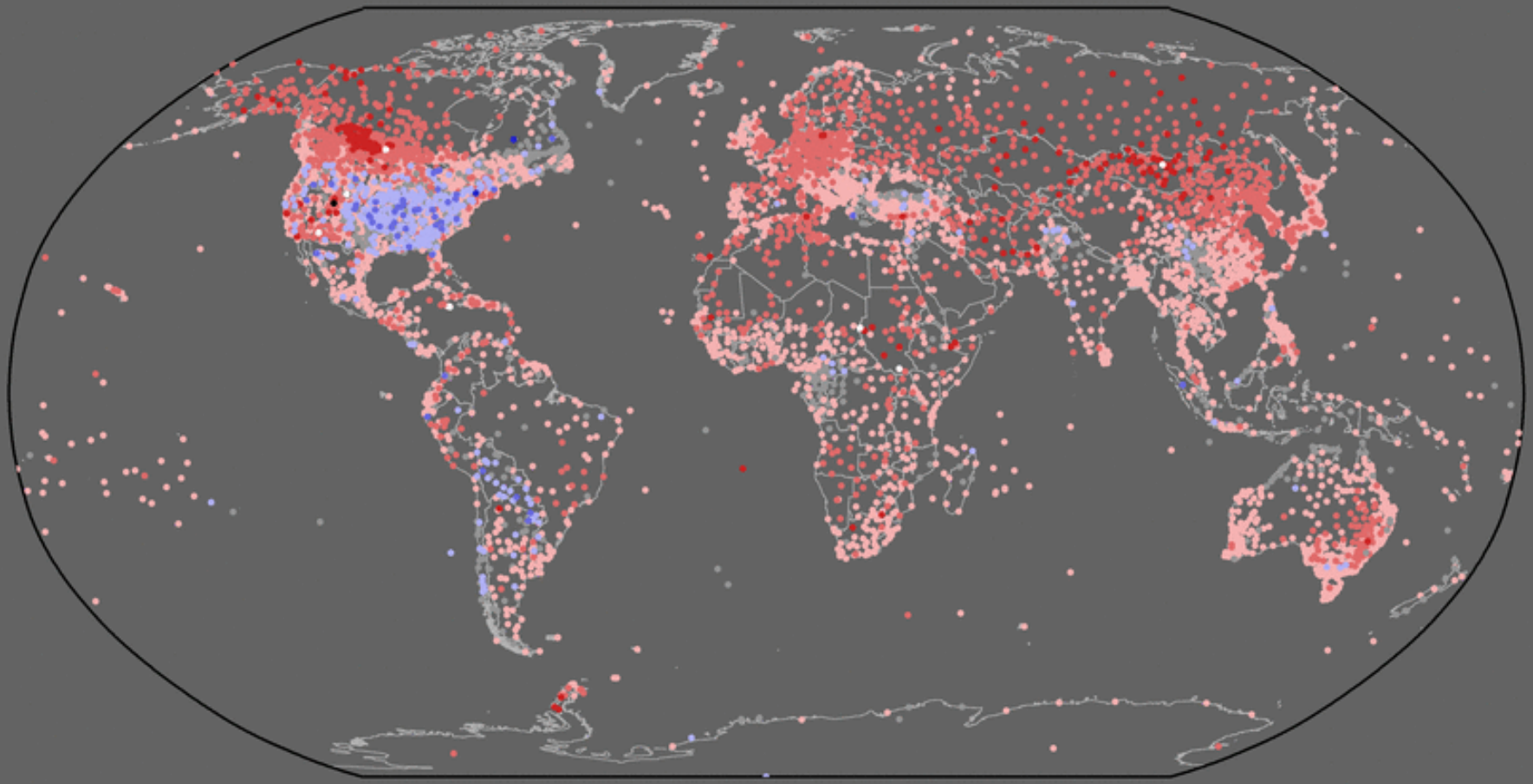
Differences between annual minimum temperatures at Cheesman, Colorado and 20 neighboring stations



4) Automation is the only realistic approach to deal with large datasets

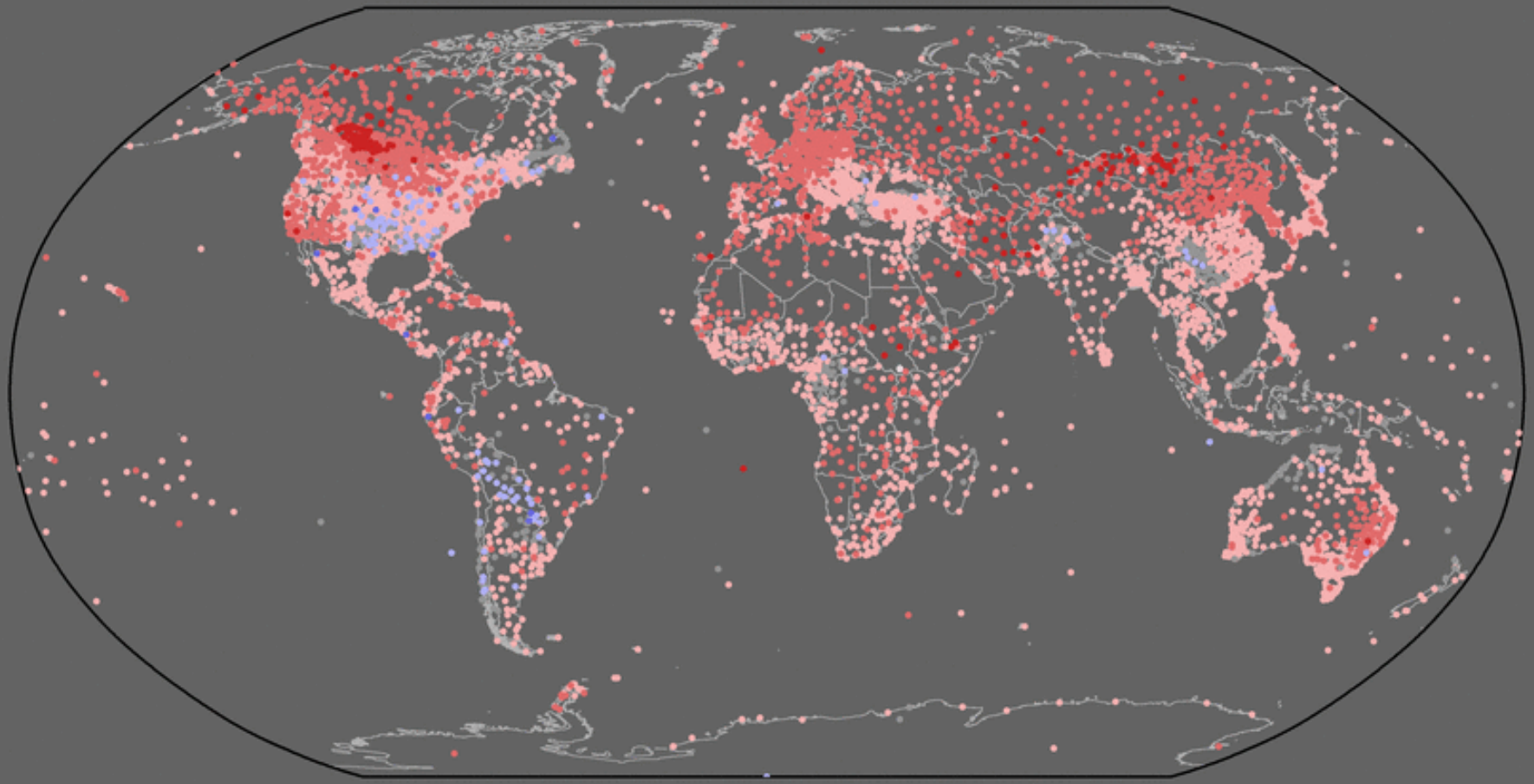
Average Annual Temperature Trends (1950 – 2009)

GHCN Monthly (Version 3) – Unadjusted



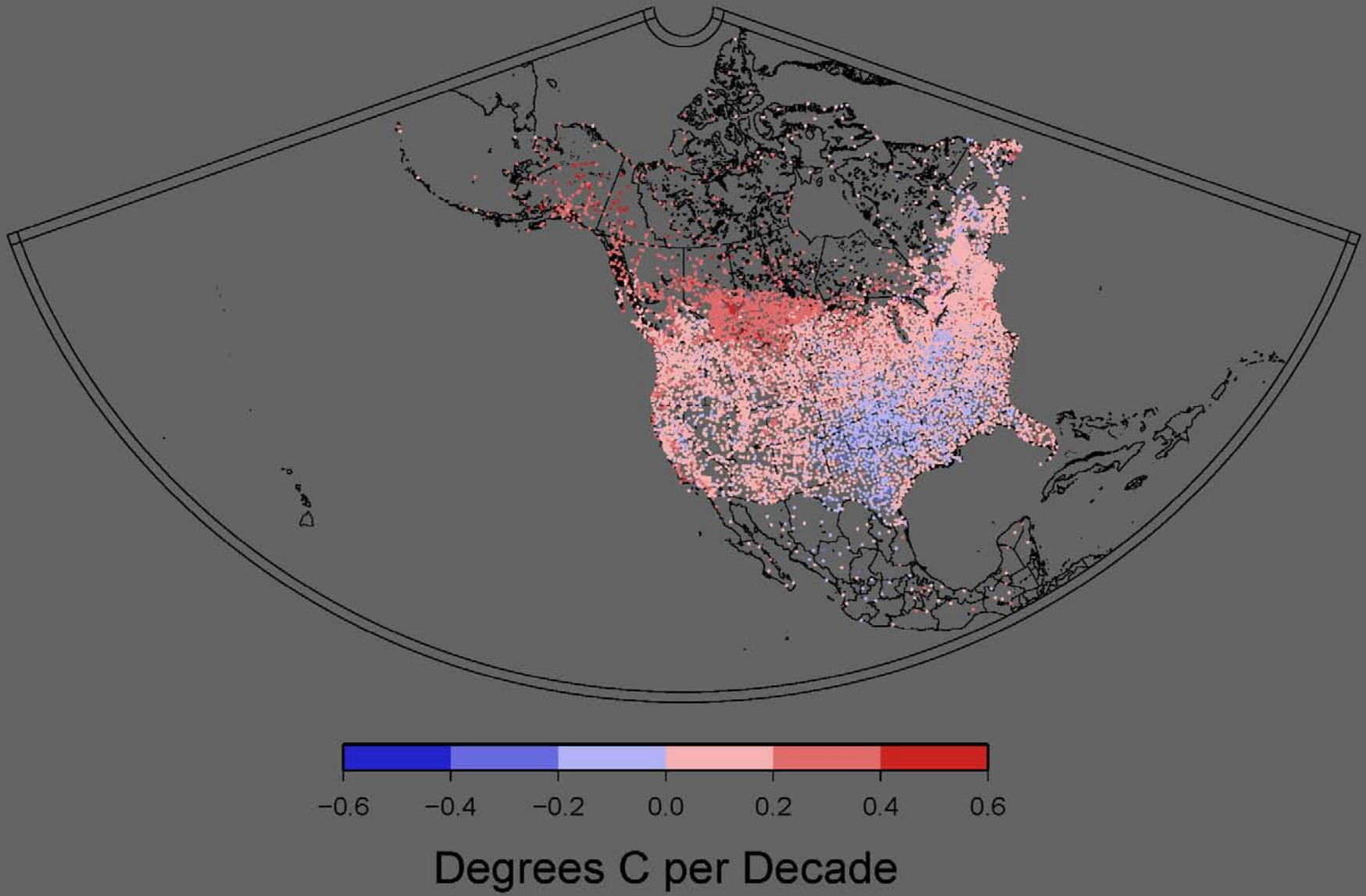
Average Annual Temperature Trends (1950 – 2009)

GHCN Monthly (Version 3) – Homogenized



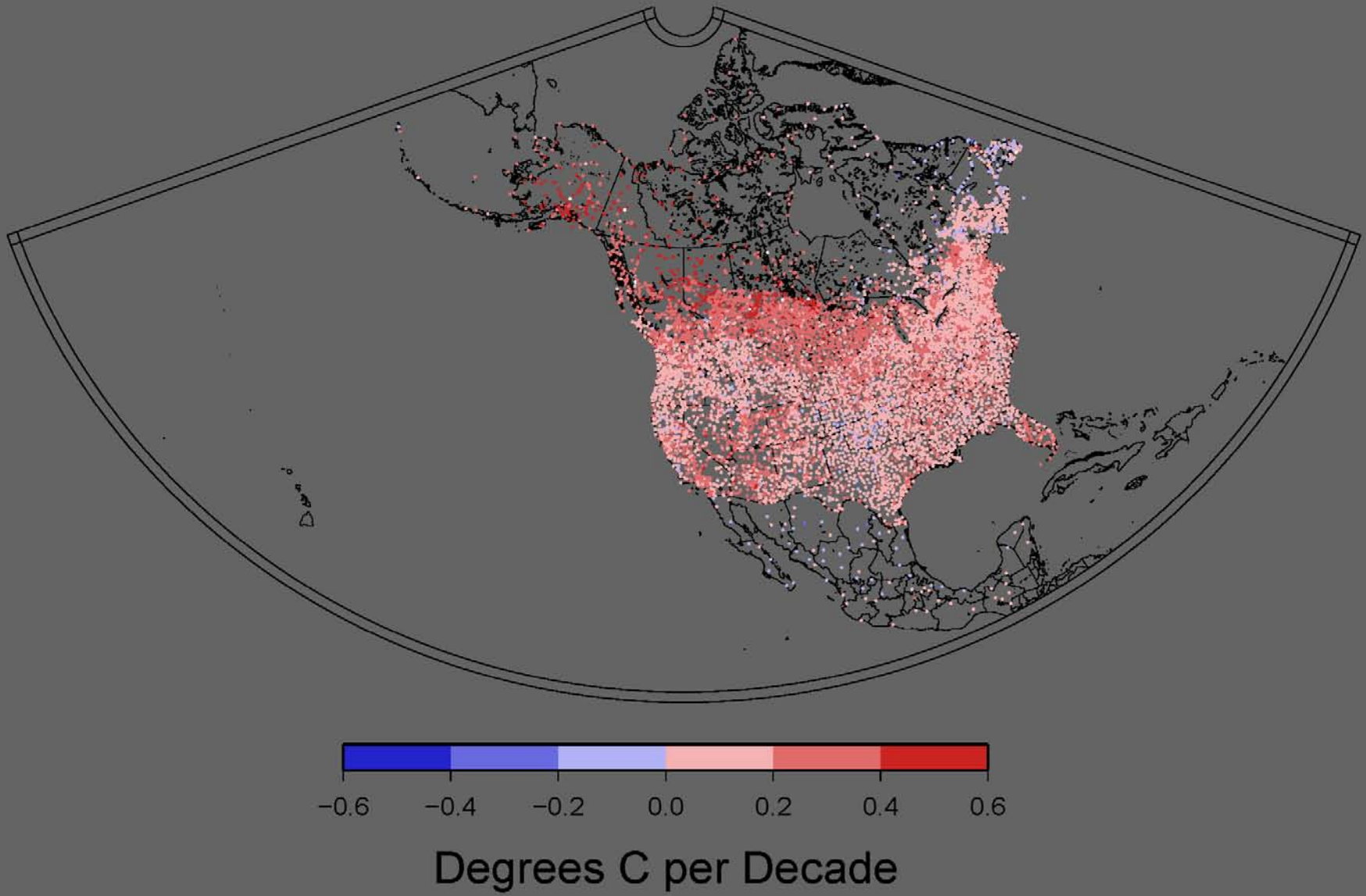
Maximum Temperature Trends (1950 to 2005)

GHCN Daily – Homogenized



Minimum Temperature Trends (1950 to 2005)

GHCN Daily – Homogenized

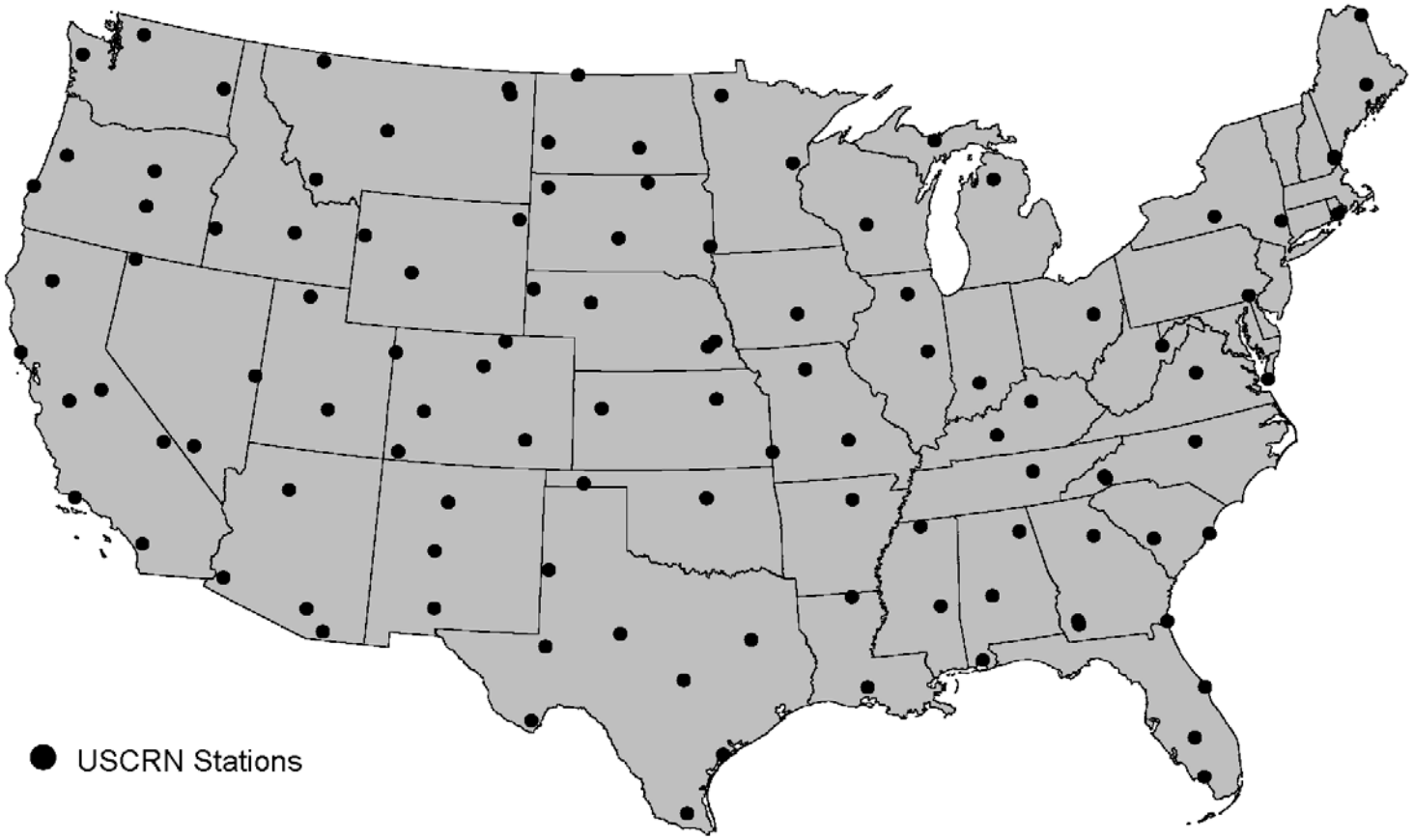


A Climate Data Scientist's Hippocratic Oath

- *First, do not flag good data as bad*
- *Then, do not make bias adjustments where none are warranted*

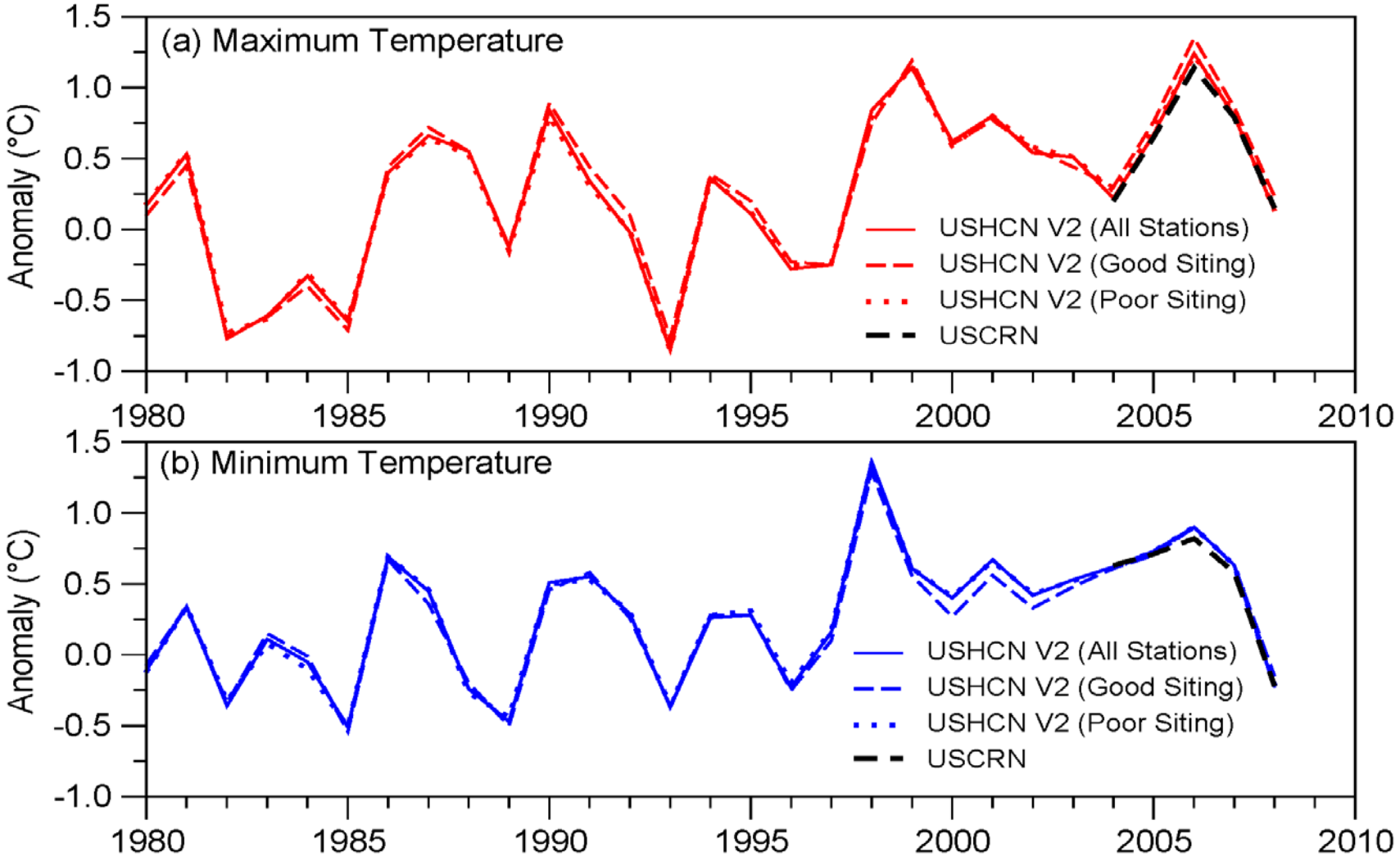
5) More work is required to assess and quantify uncertainties in bias adjustments

Example: Climate Reference Networks

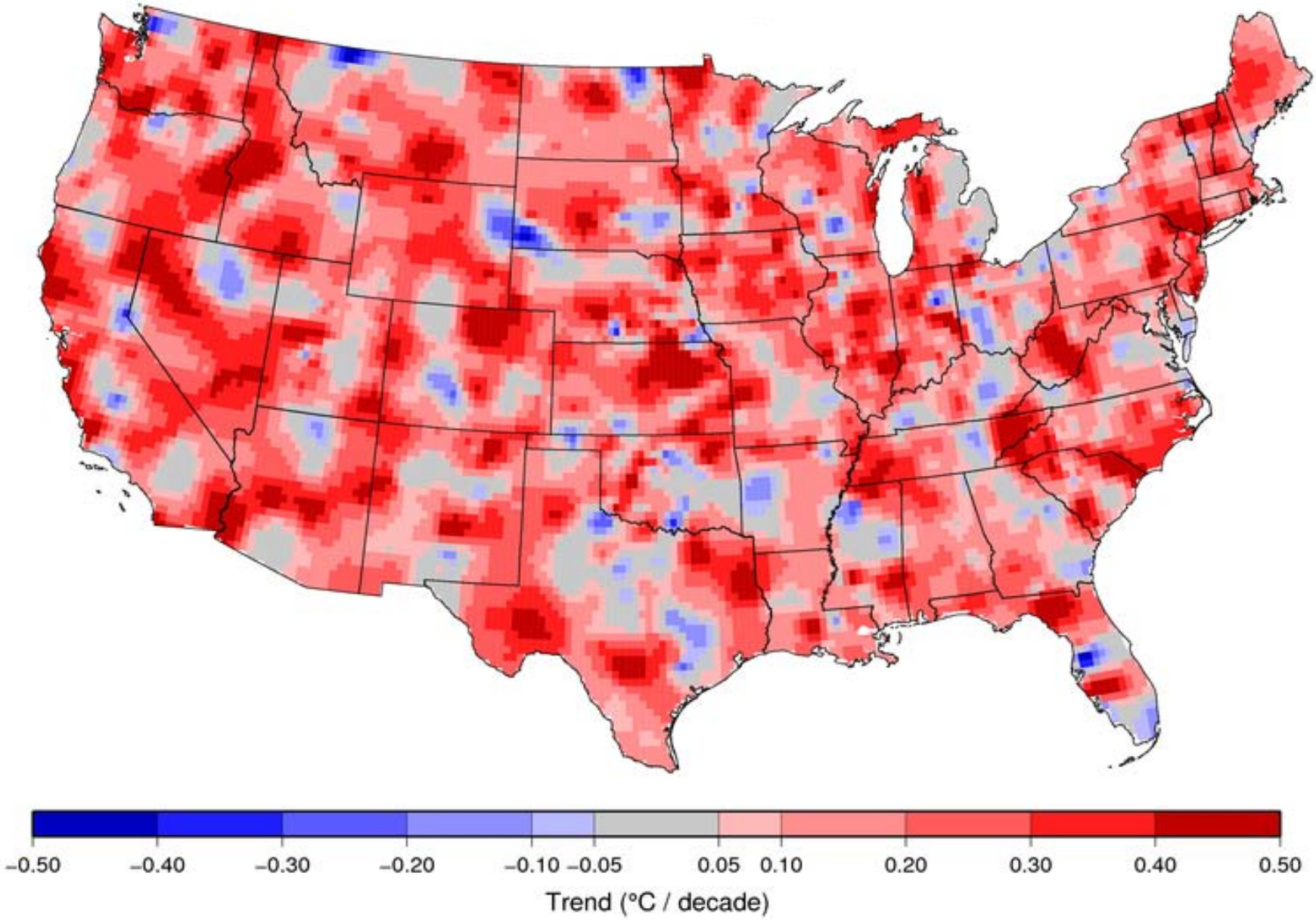


U.S. Climate Reference Network stations meet the highest standards for station siting

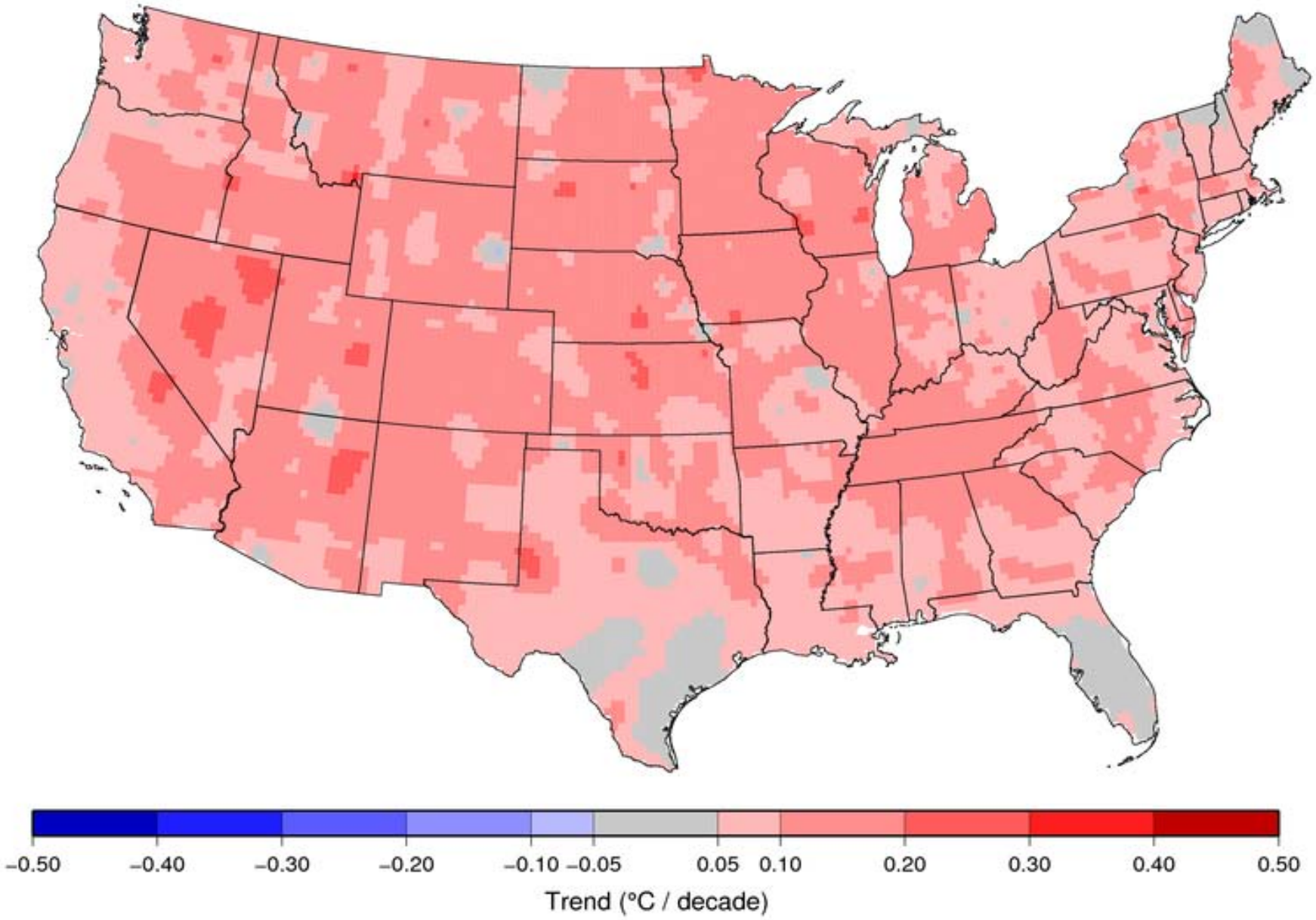
Annual Average Maximum and Minimum Temperature during the widespread instrument change in the USA (and possible degradation of siting conditions)



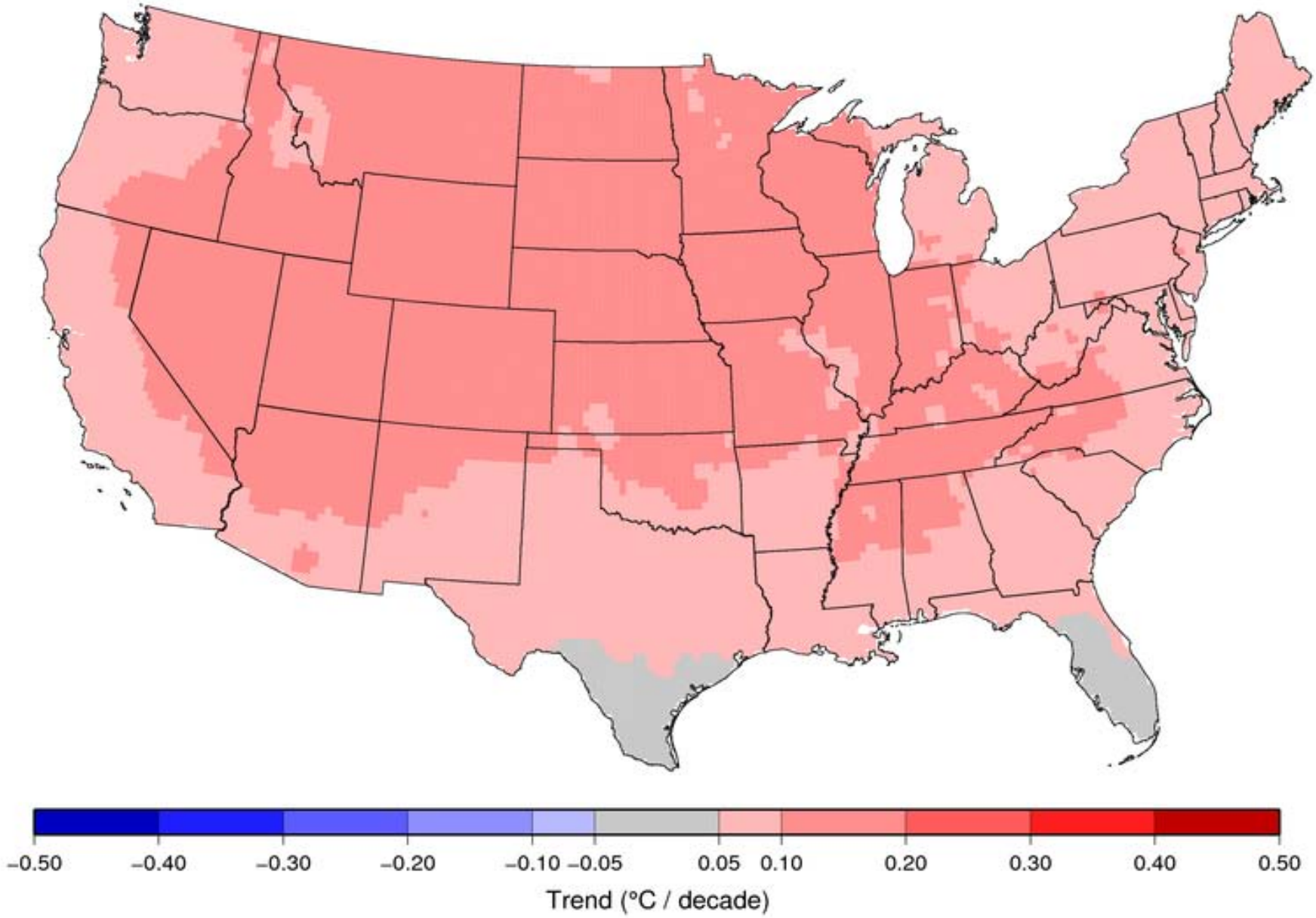
Temperature trends in a **benchmark** dataset simulated to have the same periods of record and spatial distribution of the U.S. Cooperative Observer Network station with an unknown (to us) distribution of shifts

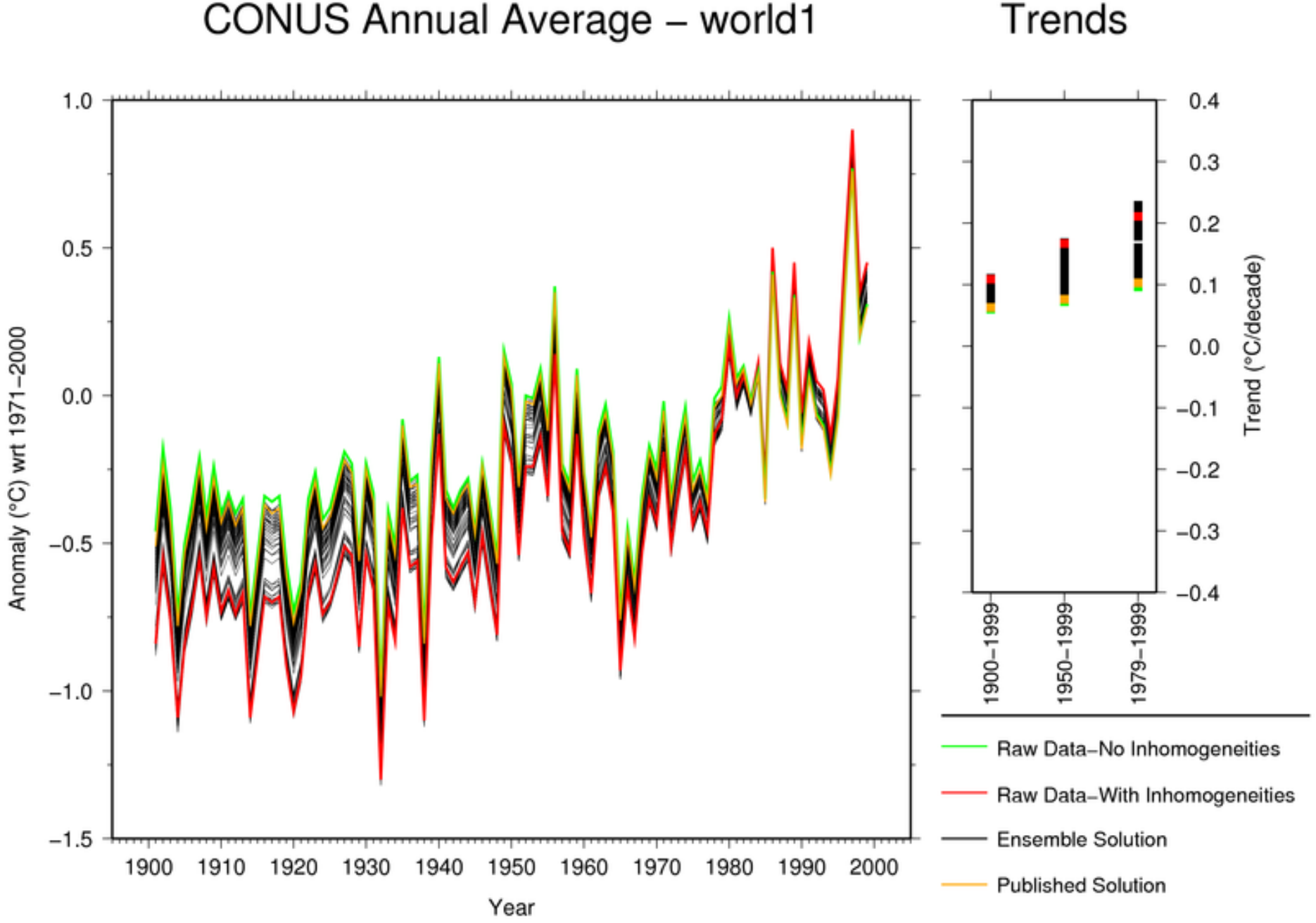


Temperature trends in the benchmark following homogenization



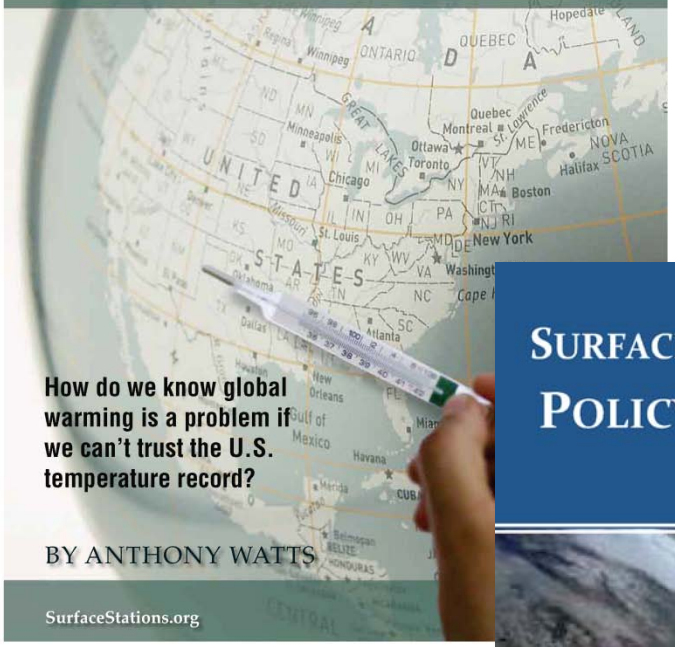
Temperature trends in the benchmark with no imposed shifts (i.e., the "truth")




CONUS Annual Average – world1
Trends
Anomaly (°C) wrt 1971–2000
Year
Trend (°C/decade)
1900–1999
1950–1999
1979–1999
Raw Data–No Inhomogeneities
Raw Data–With Inhomogeneities
Ensemble Solution
Published Solution

6) Critiques of surface temperature data and processing methods are increasingly coming from non traditional scientific sources (non peer reviewed) and the issue raised may be too numerous and too frequent for a small group of traditional scientists to address

Is the U.S. Surface Temperature Record Reliable?



How do we know global warming is a problem if we can't trust the U.S. temperature record?

BY ANTHONY WATTS

SurfaceStations.org

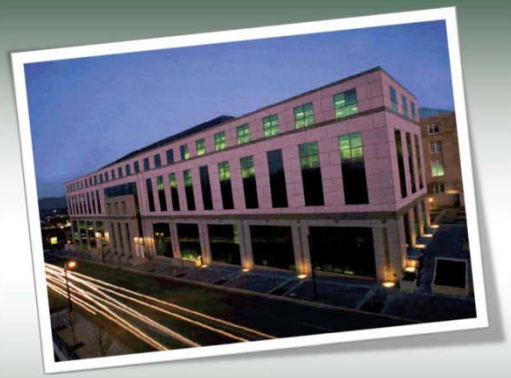
Surface Temperature Records: Policy-Driven Deception?

by Joseph D'Aleo and Anthony Watts

SPPI Original Paper ♦ Updated: Aug. 27, 2010

Contiguous U. S. Temperature Trends Using NCDC Raw and Adjusted Data for One-Per-State Rural and Urban Station Sets

by Edward R. Long, Ph.D.



SPPI Original Paper ♦ February 27, 2010

Climate Blogs

(Just to name a few…)

- Air Vent
- Bishop Hill
- Climate Progress
- Climate Science
- Dr. Roy Spencer
- Harmonic Oscillator
- Master Resource
- MoshTemp
- Niche Modeling
- Real Climate
- Sceptical Science
- SF Environmental Policy Examiner
- The Whiteboard
- Trees for the Forest
- Watt's Up With That
- Yale Forum on Climate Change
- A Few Things Ill Considered
- The Blackboard
- Climate Audit
- Climate Science Watch
- Climate.org
- ClimateWire
- DeSmogBlog
- Eco-Compass
- Greenfyre's
- Gristmill
- It's Getting Hot In Here
- Master Resource
- RomanM
- Solve Climate
- Steven Mosher
- Think Progress
- TreeHugger

Citizen scientists are a growing presence in the field

Examples of Postings by Zeke Hausfather on "The Blackboard"

- **A cooling bias due to MMTS?**
  - **8 April, 2010 (15:23) | Data Comparisons**
- **Comparing global land temperature reconstructions**
  - **29 March, 2010 (12:35) | Data Comparisons**
- **UHI in the U.S.A.**
  - **18 March, 2010 (11:44) | Data Comparisons**
- **In search of the UHI signal**
  - **9 March, 2010 (09:56) | Data Comparisons**
- **A detailed look at USHCN min/max temps**
  - **3 March, 2010 (13:33) | Data Comparisons**

7) There is a growing interest in the nature of surface temperature data (reaching up to the highest levels of government)

Congress of the United States
Washington, DC 20515

June 19, 2009

Todd J. Zinser, Inspector General
U.S. Department of Commerce
Office of Inspector General
1401 Constitution Avenue, NW
Room 7898C
Washington, DC 20230

Dear Mr. Zinser:

We are writing to express our deep concern that our national temperature data are not being measured and recorded accurately by the United States Historical Climate Network (USHCN). As you are aware, this temperature data is at the heart of not only national decision-making, but significant international efforts to understand our global climate systems. We must be certain that this data conforms to appropriate levels of accuracy and quality.

A recent report by climatologist Anthony Watts titled "Is the Surface Temperature Record Reliable?" states that 89% of the national weather stations in the USHCN "fail to meet the National Weather Service's own siting requirements that stations be 30 meters (about 100 feet) or more away from an artificial heating or radiating/reflecting heat source." The report further states that 69% of all stations produce unreliable data with an expected error greater than 2 degrees Celsius according to the NOAA quality rating system.

Based upon this and other studies, there is a significant likelihood that the data provided by the USHCN are being used by unsuspecting individuals without a real understanding of the quality, accuracy, and margins of error of the data.

These data sets, which are widely held up as the most accurate in the world, are sent to the NASA GISS (Goddard Institute for Space Studies) for their Surface Temperature Analysis. This Surface Temperature Analysis study is widely used throughout the world, including for the United Nations Intergovernmental Panel on Climate Change (UN IPCC). Therefore it is clear that the USHCN data is critical to our scientific work towards understanding climate systems.

REGTED

The USHCN is comprised of 1,221 stations specifically selected by the National Climatic Data Center (NCDC) and the Carbon Dioxide Information and Analysis Center (CDIAC) of Oak Ridge National Laboratory. No independent assessment of the quality of these stations has ever been performed, and no ongoing process to maintain quality exists. In fact, it appears that many of the stations in the USHCN don't even know that their data is used as a "high-quality" subset for national and international decision-making.

Even more disturbing than the errors introduced into the raw data is the fact that the USHCN database construction makes five separate adjustments to the raw data before it is considered usable. Each of these adjustments necessarily produces errors and decreased reliability in the data set. Four of the adjustments result in a significant increase in the "Final" data set values. The overall result of these adjustments has been an ever-increasing gap between the observed data and the reported data.

Even more disturbing than the errors introduced into the raw data is the fact that the USHCN database construction makes five separate adjustments to the raw data before it is considered usable. Each of these adjustments necessarily produces errors and decreased reliability in the data set. Four of the adjustments result in a significant increase in the "Final" data set values. The overall result of these adjustments has been an ever-increasing gap between the observed data and the reported data.

Therefore, we request that the Department of Commerce Office of Inspector General commence an investigation of the data collection processes carried out by the USHCN sites and data manipulation processes carried out by NOAA. In particular, we urge you to look at the USHCN and perform an independent assessment of: (1) the temperature station site guidelines; (2) the suitability of station site locations according to existing guidelines; (3) the quality, utility, objectivity, and integrity of the data, for both raw and adjusted data; (4) transparency with regard to the adjustments to the data including margins of error for each adjustment and for the final data set; and (5) specific known or recommended uses of the data and transparency of the margins of error as displayed by data users in their products.

A prompt, independent evaluation of this network is critical as Congress considers sweeping legislation that will have a tremendous impact on many aspects of our lives, based on data that may be substantially inaccurate. Thank you very much for your attention to this matter of national and international importance.

Sincerely,

ETC

Joe Barton R-TX
Joe Barton
Member of Congress

Dana Rohrabacher R-CA
Dana Rohrabacher
Member of Congress

R-MD
Roscoe Bartlett
Member of Congress

R-GA
John Linder
Member of Congress

R-CA
George Radanovich
Member of Congress

R-UT
Rob Bishop
Member of Congress

R-TN
Marsha Blackburn
Member of Congress

R-TX
Michael Conaway
Member of Congress

R-TX
Michael McCaul
Member of Congress

R-CA
John Campbell
Member of Congress

R-GA
Paul C. Broun
Member of Congress

ENERGY & COMMERCE

R-UT
Jason Chaffetz
Member of Congress

R-WY
Cynthia M. Lummis
Member of Congress

Take Aways from the Lessons Learned

Scientific and methodological context:

1) **Lesson 1**: Metadata records are helpful, but we must be prepared to have less than comprehensive station histories
2) **Lesson 2**: There is a continuous need for “sifting and winnowing” to discover the causes behind systematic shifts in bias, especially those that aggregate across regions and the globe
3) **Lesson 3**: It may be difficult to come to full agreement on the details and philosophy for data homogenization (e.g., how to handle gradual vs abrupt changes in bias), but our benchmark datasets should be as realistic and varied as possible to quantify the skill of any particular approach
4) **Lesson 4**: A small group of people can make big contributions to the field—especially with automated algorithms
5) **Lesson 5**: Numerous avenues exist to build confidence and assess uncertainties in dataset construction (many of which are underway)

“Societal” context

6) **Lesson 6**: Non-traditional climate scientists will likely play a significant role in advancing the field of climate dataset construction
7) **Lesson 7**: For now, the world is watching so let’s find a way to build momentum