

Climate Data and Information: Issues and Uncertainty

David Easterling

*NOAA/NESDIS/National Climatic Data Center
Asheville, North Carolina, U.S.A.*



NOAA'S National Climatic Data Center



Discussion Topics

- ✓ **Climate Data sets**
 - What do we have besides 2 Petabytes and growing of data?
- ✓ **Issues with data**
 - availability
 - quality
 - homogeneity
- ✓ **Issues interpreting the data**
 - Uncertainty
- ✓ **Model-data integration**
 - Making the leap from data to model simulations.
 - How do we (climate scientists) provide model information in a form useful to decision makers?



NOAA'S National Climatic Data Center



Climate Data

- ✓ **In situ data: we have data observed at weather and climate stations from all around the world, but:**
 - Availability varies by country, both in time and space.
 - Quality varies
 - Metadata may or may not be available.
 - Updates are problematic for many countries.
- ✓ **Remotely sensed data: we also have satellite and wx radar data, hundreds and hundreds of terabytes.**
 - Satellite data provides true global coverage
 - But period of record is short (earliest is mid-1960s, most useful is only since 1979).
 - Making the data useful, let alone useful for decision-makers is an issue.



NOAA'S National Climatic Data Center



Global Historical Climatology Network (GHCN) Daily Dataset

- ✓ **Global daily *in situ* dataset derived from multiple sources (~10 sources for the U.S. including Forts data and CRN)**
- ✓ **~25,000 temperature stations**
- ✓ **~44,000 precipitation stations**
- ✓ **~25,000 snowfall or snowdepth stations**
- ✓ **Currently >1.6 billion daily observations**
 - Earliest value from January 2, 1833
 - Latest value from yesterday



NOAA'S National Climatic Data Center

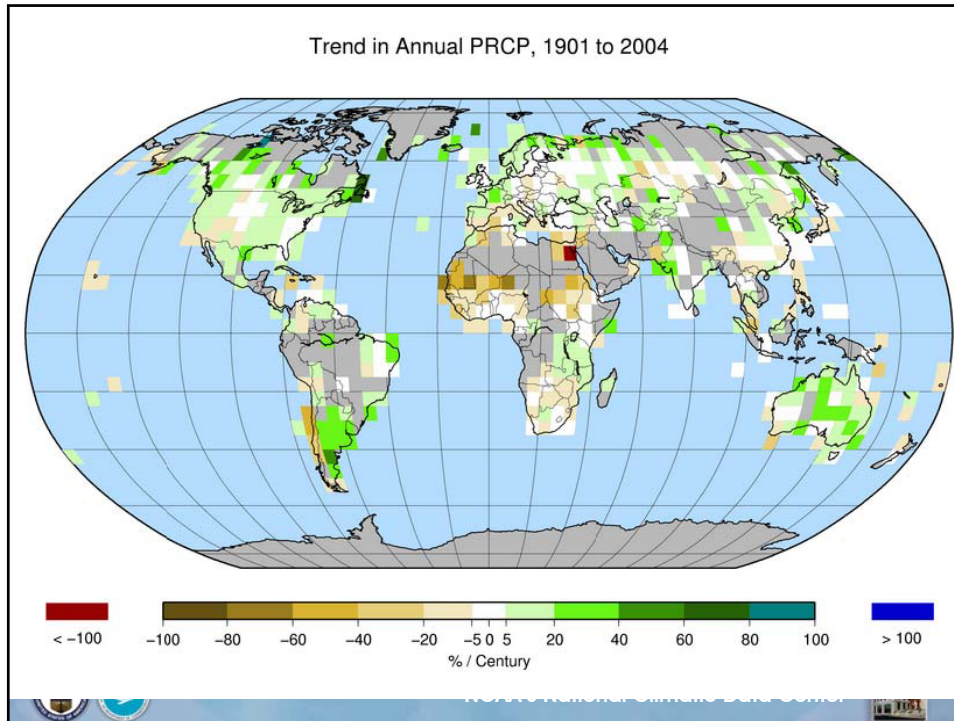




Issues with Data

- ✓ Availability
- ✓ Quality
- ✓ Homogeneity





Quality Control

- **Gross error checks**
 - Duplicate stations
 - Mislocated stations
 - Identical years
 - Changes in mean or variance
 - Change in temperature scale
 - Decimal shifts
 - Runs of identical data

- **Outlier checks**
 - Climatology
 - Nearest neighbors

Bilma, Niger

Homogeneity Issues

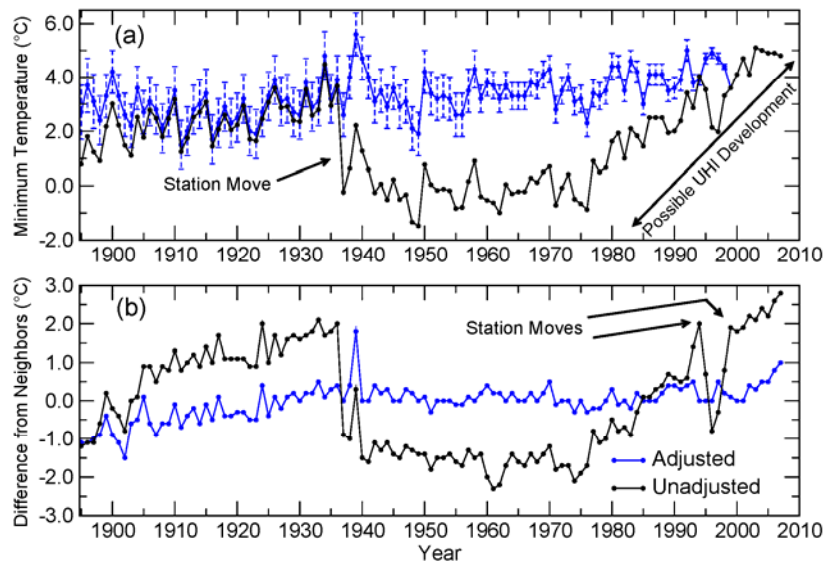
- ✓ A homogeneous climate time series is one where all variations and trends are due solely to climate.
- ✓ What causes an inhomogeneous time series?
 - Station moves
 - Instrument changes
 - Land use changes
 - Data processing changes
- ✓ What is done to homogenize the time series?
 - Discontinuity detection and adjustment by comparing with surrounding stations.



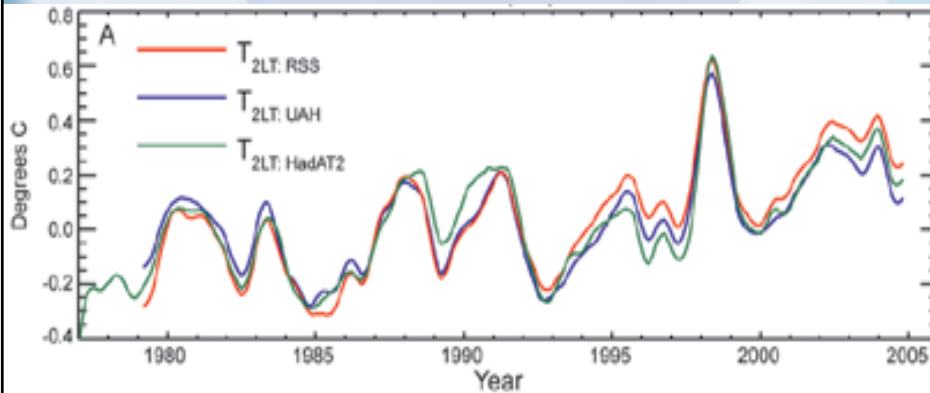
NOAA'S National Climatic Data Center



Reno: Mean Annual Air Temperature



Microwave Sounding Unit (Satellite) Lower Troposphere Temperature

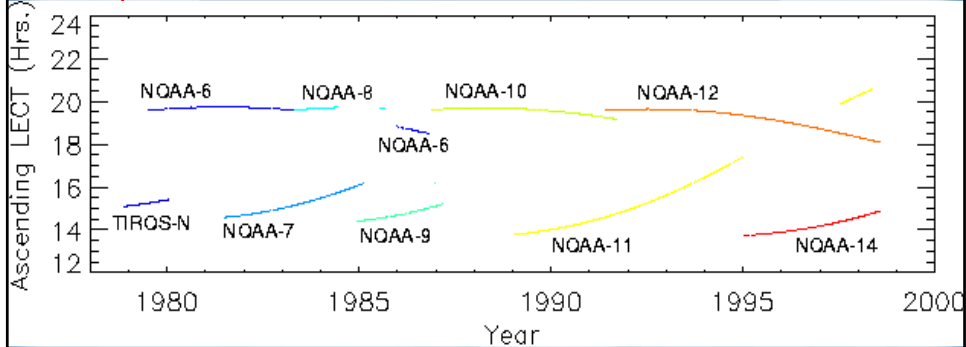


NOAA'S National Climatic Data Center



Why a Difference in Results?

✓ Impact of Satellite Orbital Drift



Changes in orbits, equator crossing times and altitude of satellites are aliased onto the diurnal cycle, requiring "corrections".

One research group was incorrectly applying "corrections" with resulting time series showing little warming.



NOAA'S National Climatic Data Center



All these issues increase uncertainty in results.



NOAA'S National Climatic Data Center



Uncertainty

✓ Statistical Uncertainty

- Uncertainty due to random “noise” in climate data
- Or the uncertainty due to how well the data are described by a particular model (e.g. linear trend, Gamma distribution, Principal Components, etc.) in an analysis of the data.

✓ Construction Uncertainty

- Uncertainty due to non-climatic factors impacting the data.



NOAA'S National Climatic Data Center



Construction Uncertainty

- ✓ **Structural Uncertainty**
 - Uncertainty due to choices or assumptions made in how to adjust data, or in how to formulate a climate model etc.
- ✓ **Parametric Uncertainty**
 - Once the choices are made the statistical parameters of the methods or model still have to be estimated from a sample, leading to more uncertainty.
- ✓ **True in both data homogeneity adjustments and in climate models.**



NOAA'S National Climatic Data Center



Model-Data Integration: what do we have planned for the next few years?

- ✓ **Model simulations are a rich source of information for decision and policy makers dealing with impacts of climate change.**
- ✓ **But how do we provide that information in a form usable and understandable by this user base?**
- ✓ **Decision support tools:**
 - Downscaling of model simulations.
 - Climate expertise to go along with it.



NOAA'S National Climatic Data Center



Optimal and Dynamic Normals

- ✓ **Normal: 30 year average based on defined period (e.g. 1961-1990 or 1971-2000)**
- ✓ **Two Issues:**
 - With trends 30 year period represents middle of period.
 - Users want options for different period definitions (e.g. the last five years).
 - Users also want better information on statistical behavior of the Normal using advanced statistical techniques (e.g. hinge fit).
 - How do you make a Normal useful for decision makers planning for, say, the next 40 years?
 - Integrate ensemble model simulations with the Normals to produce a “projected” Normal for some point in the future.
 - E.G. Normal July 25 High Temp in AVL projected to be 38C +-2C in 2050.



NOAA'S National Climatic Data Center



Questions?

David.Easterling@noaa.gov

www.ncdc.noaa.gov



NOAA'S National Climatic Data Center

