# History of Global Temperature Estimation – Köppen to Satellites

Phil Jones
CRU, ENV, UEA
Norwich

# Summary

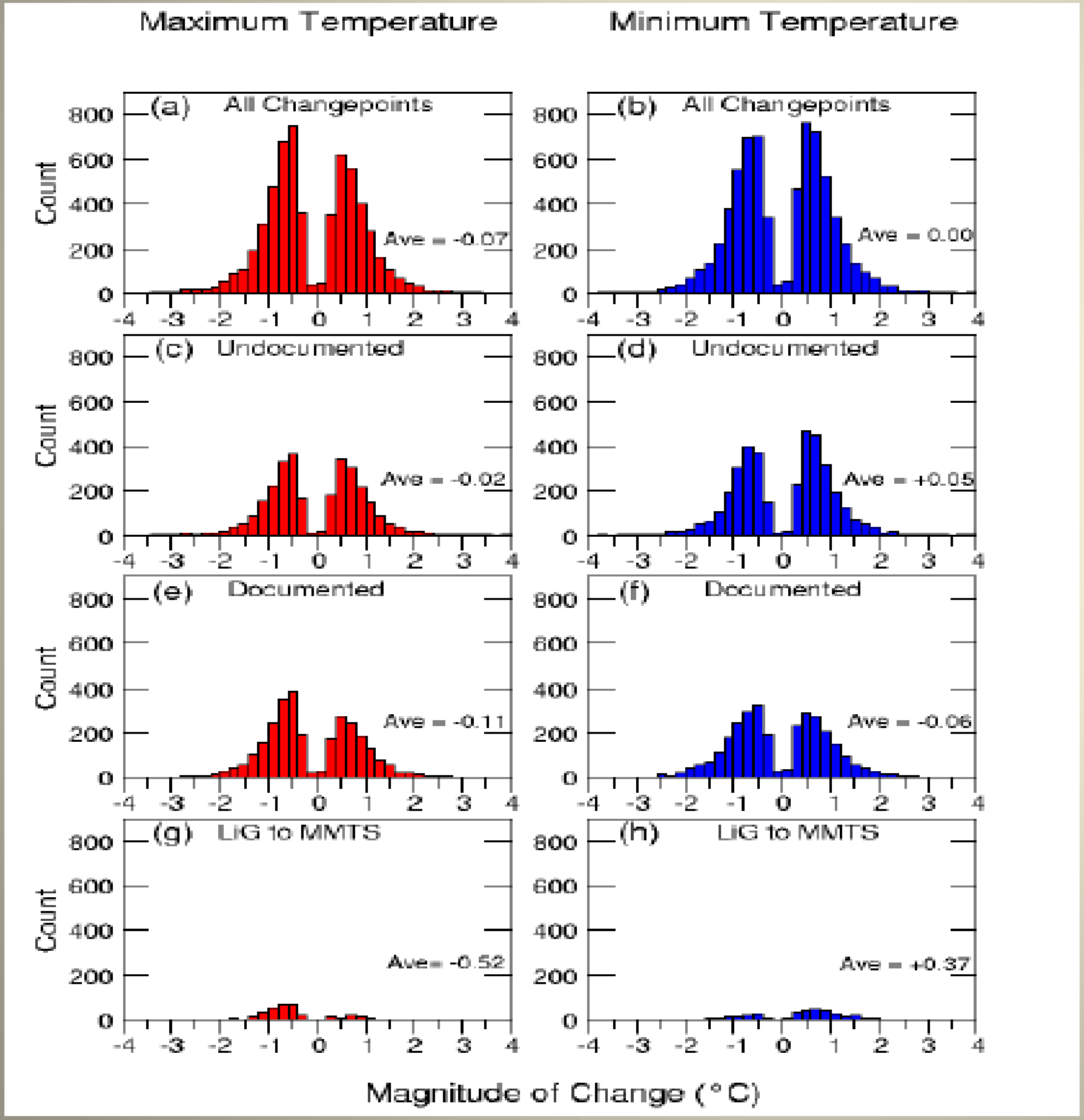
- Land Station Homogeneity
- Interpolation to Regular Grids
- Early compilations – all land only, until the 1980s
- Effect of omitting large regions
- Biases - Exposure issues before Stevenson Screens and Urban Heat Islands
- Marine data
- Combinations of land and marine data
- Satellite Estimates
- Reanalysis Estimates

# Station Homogeneity/Bias

- Homogeneity defined as a change in a temperature time series that is not due to the vagaries of the weather and climate
- Important to distinguish an issue that affects a single station (now more generally referred to as an Inhomogeneity) from one that affects many or all stations in a region (now referred to as a Bias – later we'll look at urban effects and exposure of the thermometers)
- Inhomogeneity tested by a number of relative and absolute homogeneity tests (important comparison paper by Venema et al. 2011 which resulted from an EU-COST Action called HOME)
- Whatever method is used it is key to have good Station Histories (Metadata), but a number of extensive studies indicate that even where this is good, only about half the station inhomogeneities can be related with a cause
- Many national assessments now being conducted – in a few countries every year or two, using new techniques with additional series recently digitised
- A number of studies using unadjusted and adjusted station series indicate that station homogeneity is not that important for large-scale averages, but the issue is more important at the local-to-regional scale


- Venema, V.K. et al., 2011: Benchmarking homogenization algorithms for monthly data. *Climates of the Past*, **8**: 89-115.

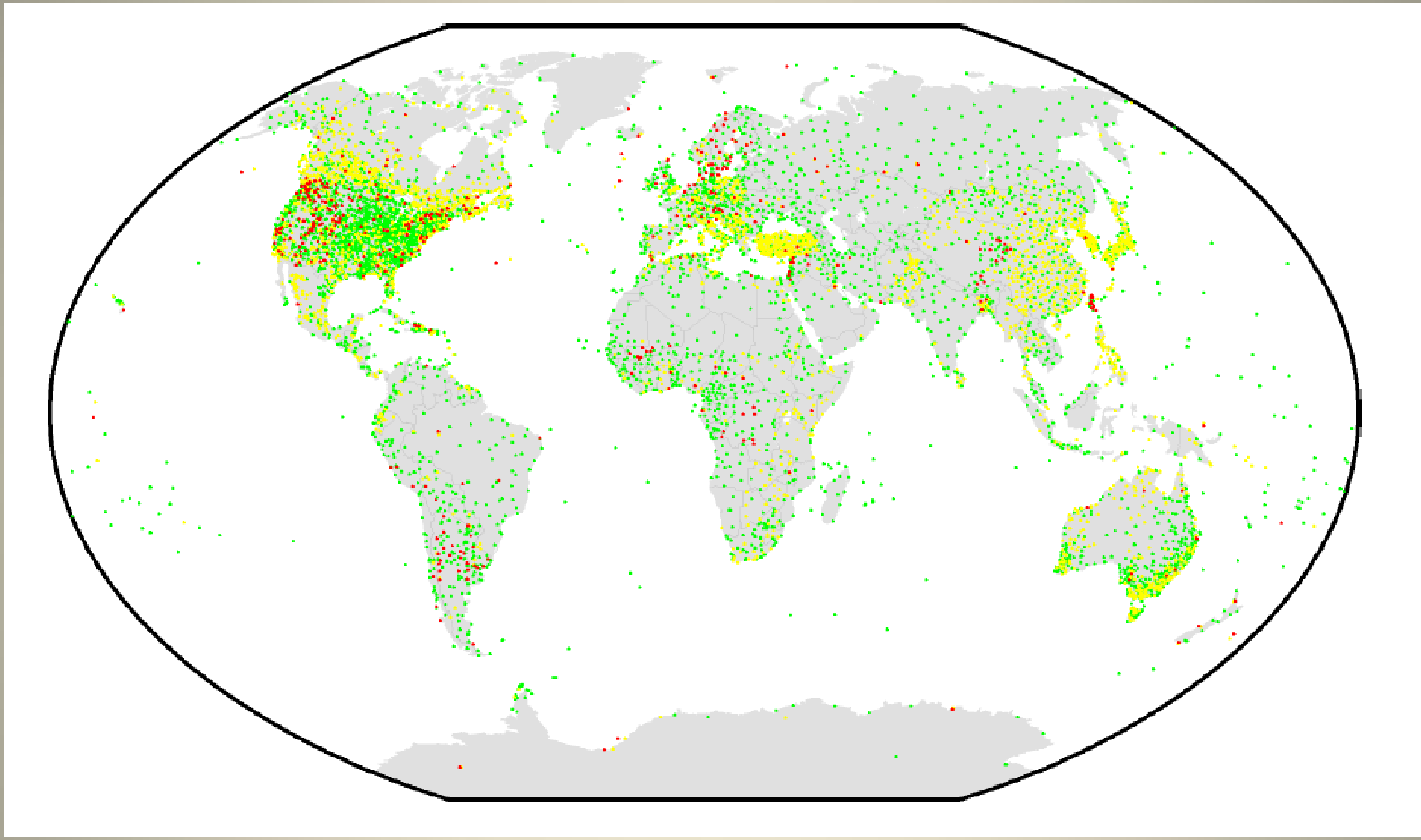
# Bimodal distribution of adjustments in a recent paper on the US HCN




Menne et al (2009) in BAMS

Matthew J. Menne, Claude N. Williams Jr., and Russell S. Vose, 2009: The U.S. Historical Climatology Network Monthly Temperature Data, Version 2, *BAMS*, **90**, 993-1007

# Station Distribution



Green – stations in CRUTEM3 and GHCN, Yellow just in GHCN and Red just in CRUTEM3

CRUTEM3 has ~5000 stations, GHCN ~7000, but BEST claim to have ~39000, but this is just more data where it is currently well sampled. Don't be deceived by station number counts.

Although GHCN and also BEST have more, it is mostly 'more' where we already have many!

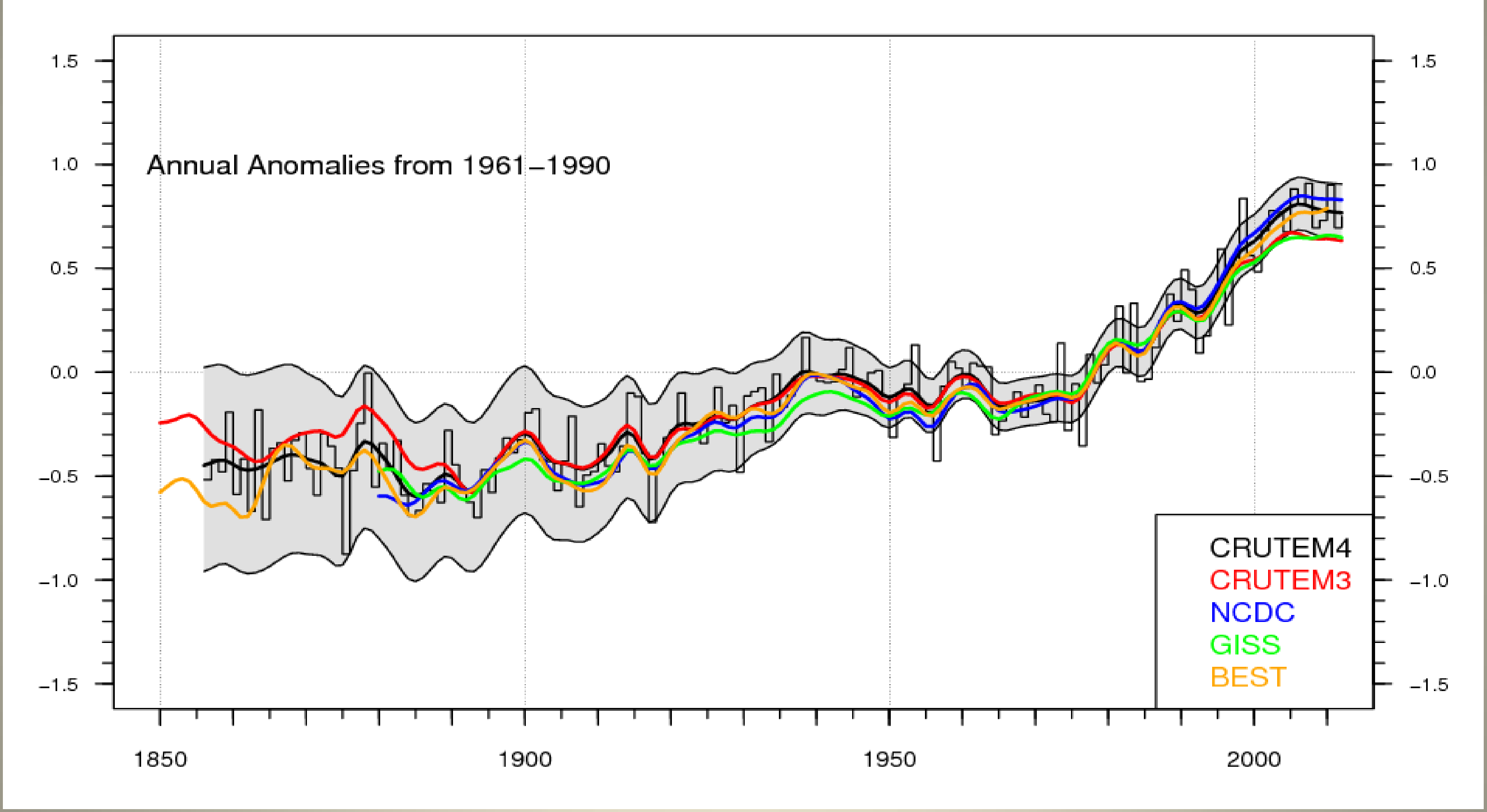
# Gridding the land station temperatures

- CRU approach is to associate each homogenous station with the 5° by 5° (latitude/longitude) grid box within which it is located
- Average for that box is then the simple average of all the station anomalies (from 1961-90) available
- Variance of the time series will be affected by changing station numbers (both over time at an individual grid box and between adjacent grid boxes)
- This can be allowed for (discussed in Brohan *et al*., 2006, based on earlier work of Jones *et al*., 1997) by adjusting all grid-box series to the infinitely sampled grid box
- Large-scale averages calculated as the weighted sum of all grid boxes in the domain – with the weights being the cosine of the central latitude of each box

- US groups undertake a variety of approaches (NCDC – PC Techniques, GISS – 1200km spheres of influence, BEST – Kriging). All three approaches are, in effect, spatially infilling. CRU doesn't do this in CRUTEM3/4, so if there are regions without station data, we get missing grid boxes
- CRU does spatial infilling, but it is a different dataset (CRU TS 3.10 and is discussed in Harris et al., 2013)

- BEST – Berkeley Earth Surface Temperature (Land Only)
- GISS – Goddard Institute of Space Sciences (part of NASA)
- NCDC – National Climatic Data Center (part of NOAA)

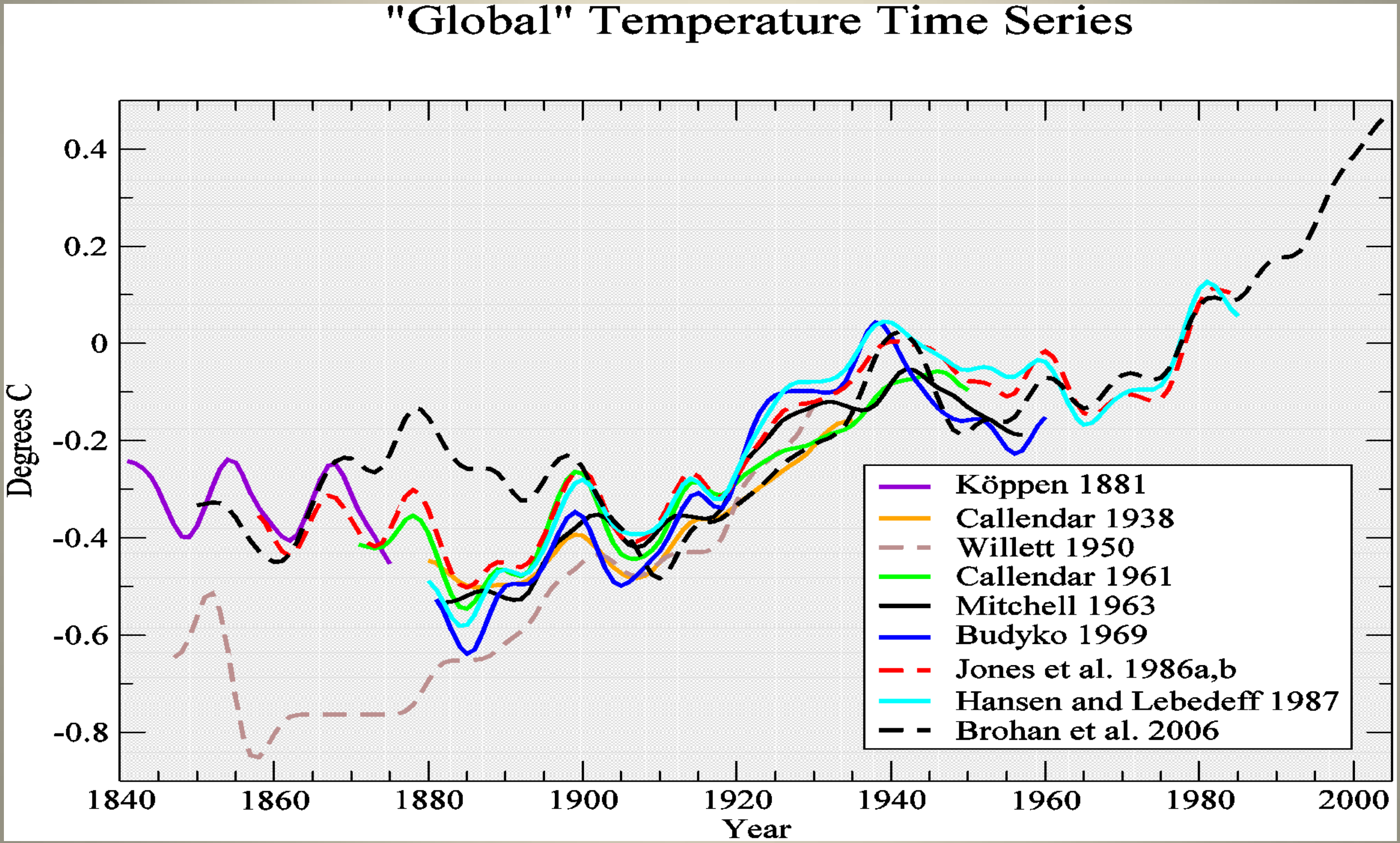
# Land-based Global Temperatures




Berkeley Earth (BEST) group reanalysed much more temperature data and got the same result. The above shows error ranges. BEST produced much smaller error ranges but they don't fully understand some components.

CRUTEM4 shows more warming than CRUTEM3, but this is due to our use of more station data in CRUTEM4 encompassing more regions. CRUTEM3/4 doesn't do any spatial infilling – all other groups do.

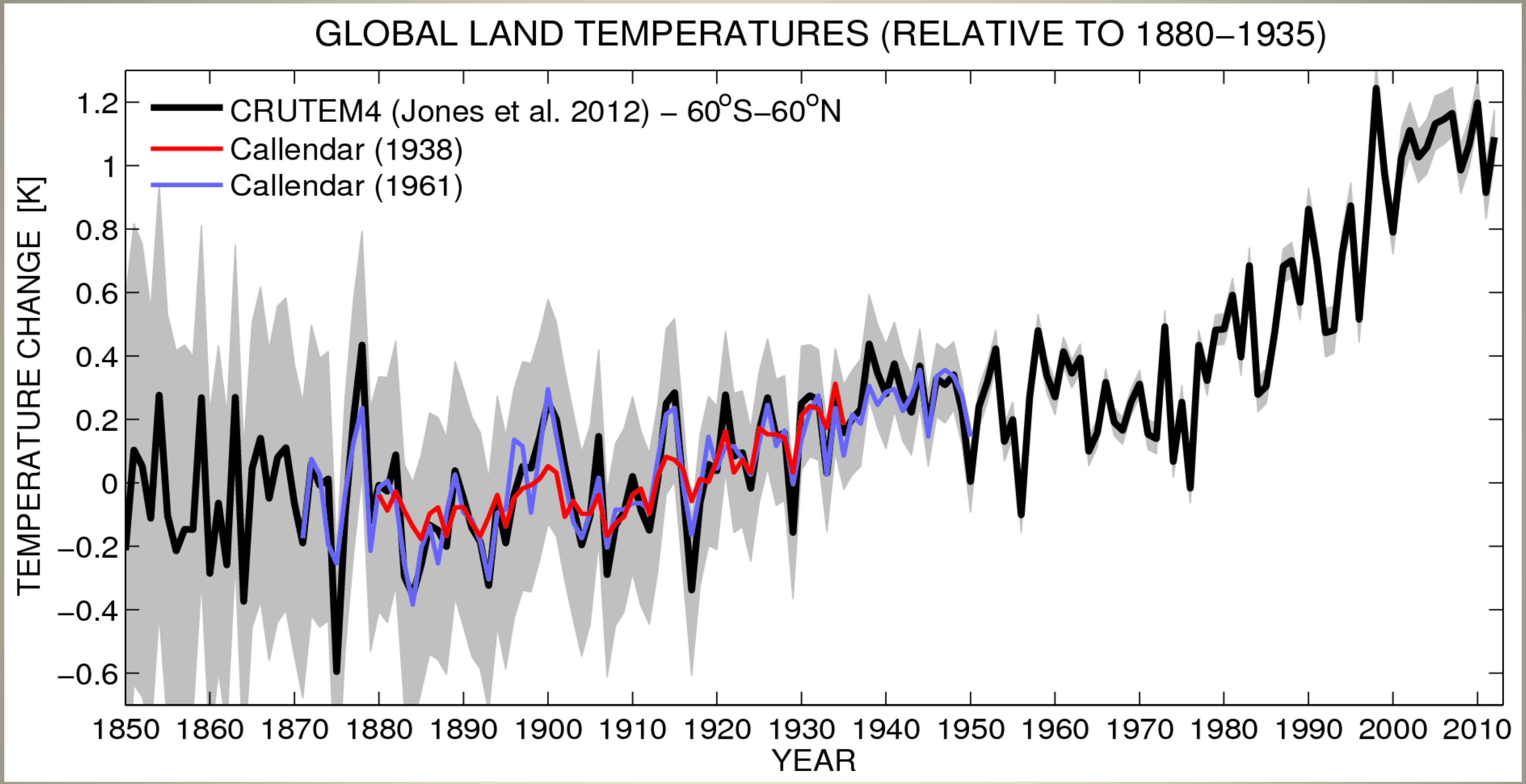
Pre-CRU land temperature series, each adjusted to have Brohan *et al* average (HadCRUT3) over their last 30 years of overlap (from Ch1 of AR4: zero line is 1961-90)




AR4 shouldn't have compared land only series with HadCRUT3 (which is land and marine)!

# Comparison of CRUTEM4 with papers by Callendar (1938, 1961)

Includes the error estimate ranges for CRUTEM4 developed by Morice et al (2012)

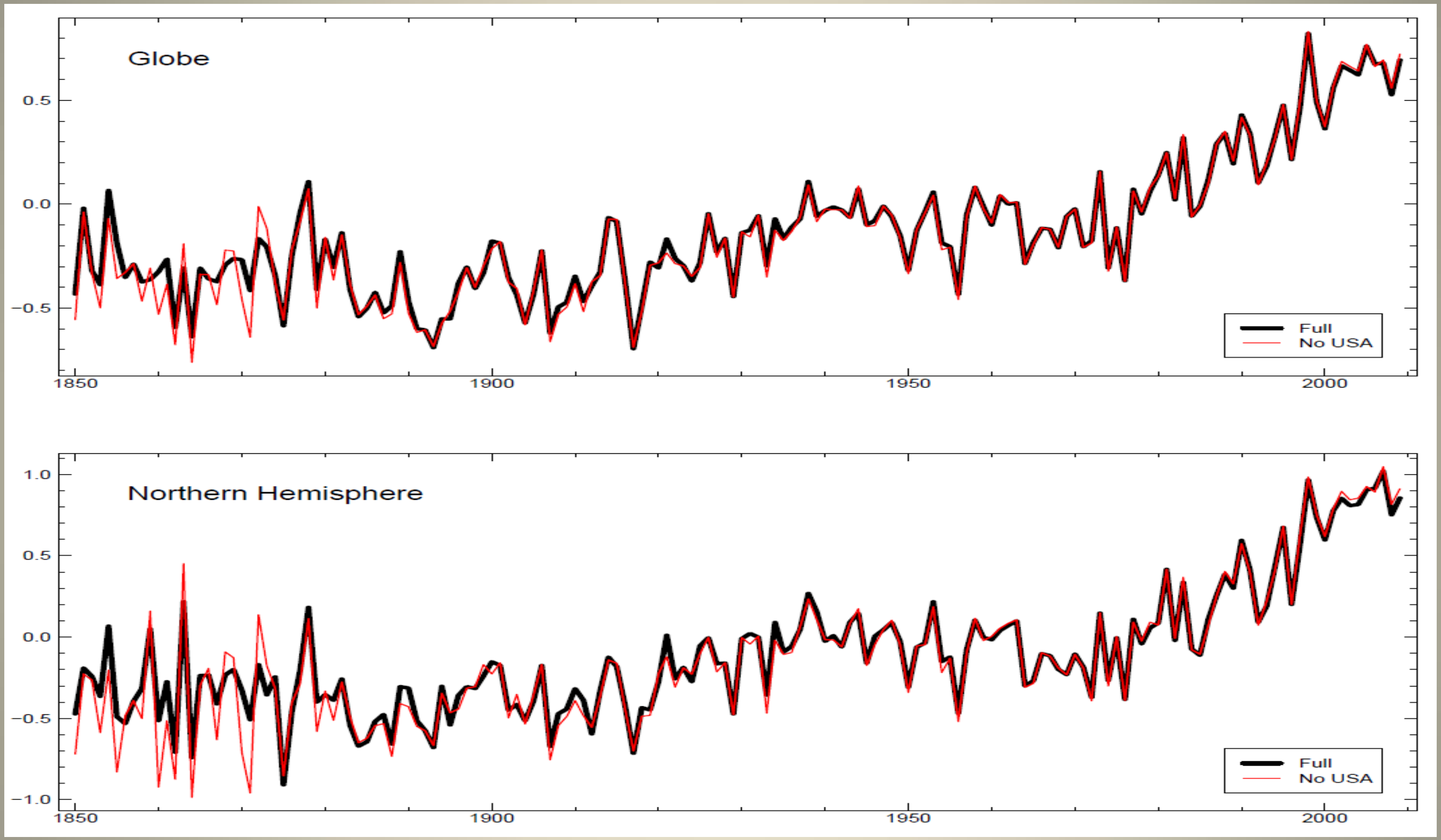



Hawkins, E. and Jones, P.D., 2013: On increasing global temperatures: 75 years after Callendar. *Q. J. Royal Meteorol. Soc.*, **139**, DOI:10.1002/qj.2178.

RMS are having a Callendar meeting in autumn 2014, to mark 50 years since his death.

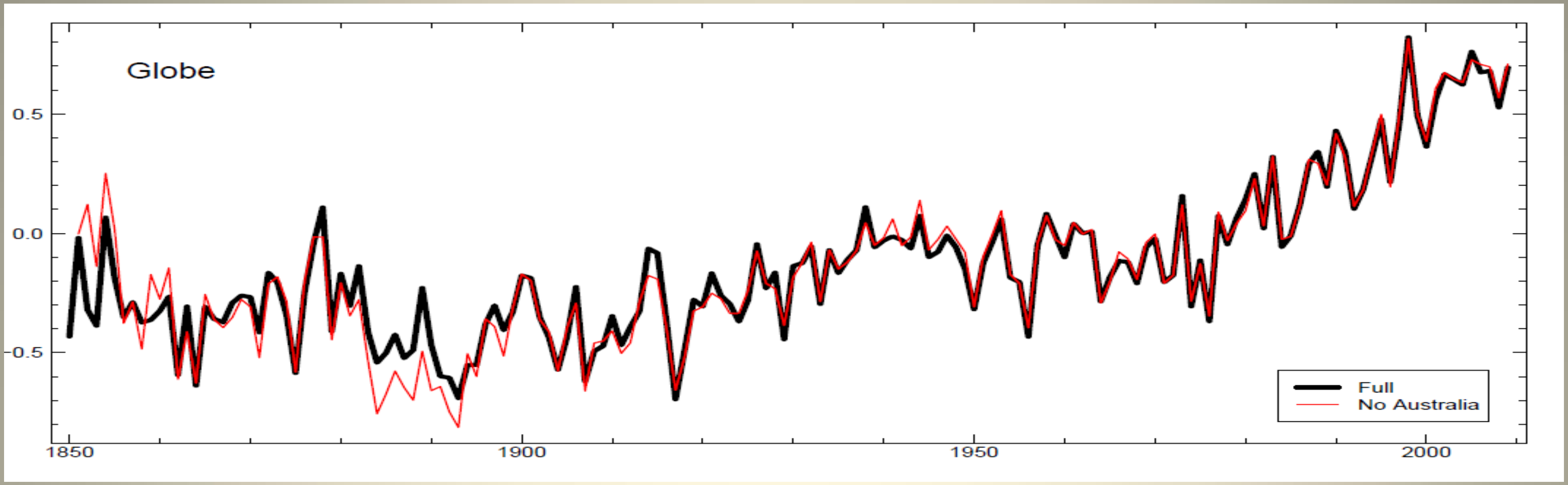
Will additionally discuss his assessment of early CO2 measurements – shown to be right by ice cores

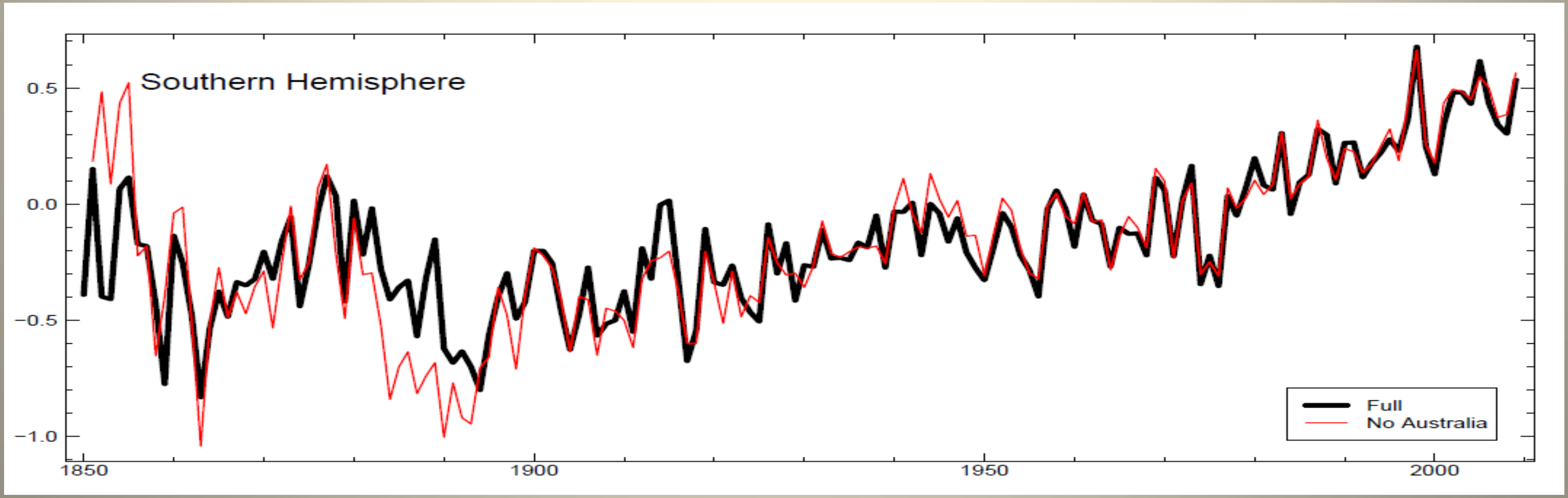
# Robustness of the global temperature record (1)
# Removal of all stations in the contiguous United States




Based on CRUTEM4

# Robustness of the global temperature record (2)
# Removal of Australia, using CRUTEM4



# Biases

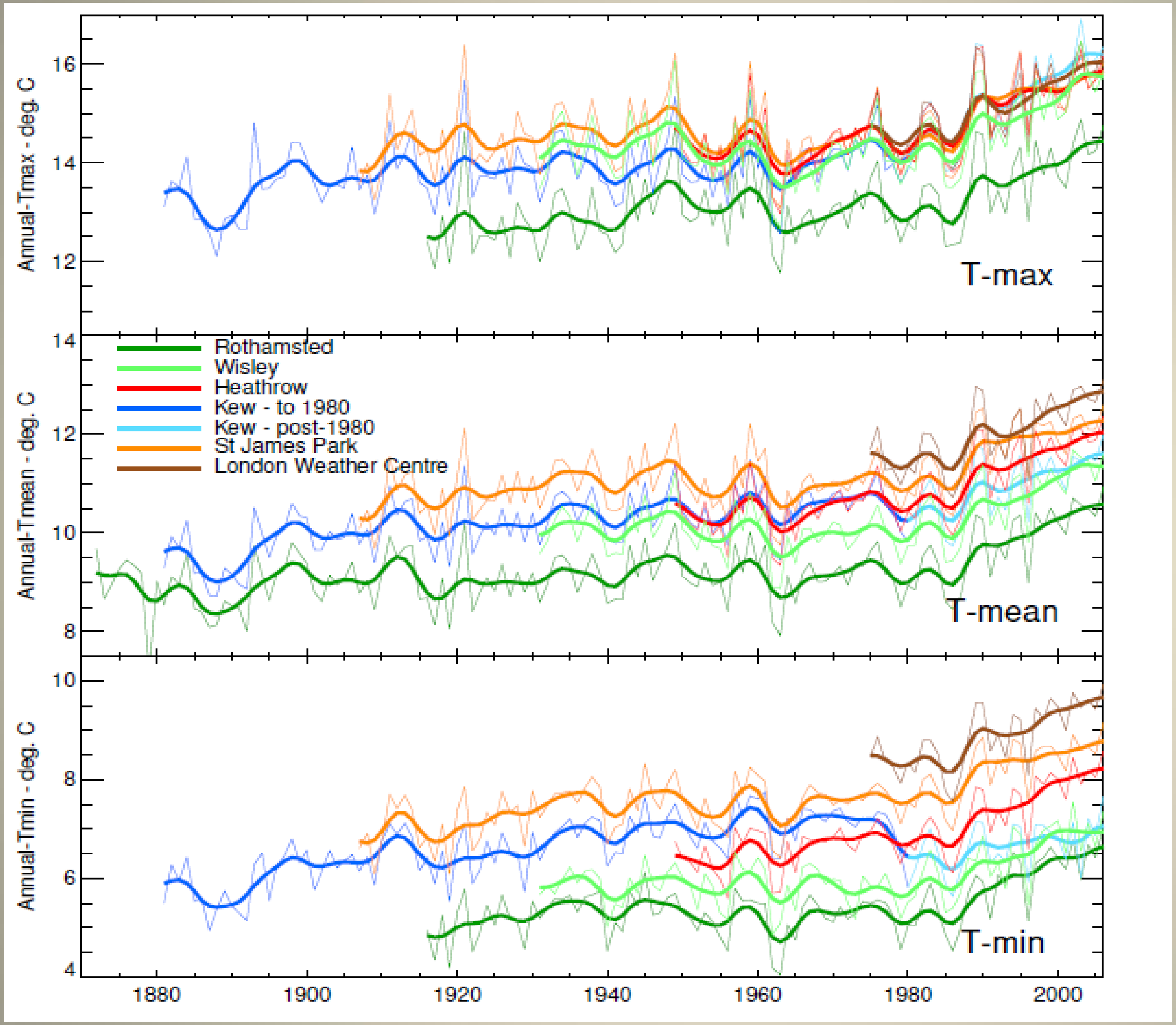
- Bucket and engine intake measurements for SSTs (most important of all, will be discussed later)
- Exposure of thermometers (Böhm et al., 2010, Brunet et al., 2010) in the pre-Stevenson screen era (only really effects the record pre-1880, but this is important in some regions later, such as Australia before 1910)
- Homogeneity techniques won't put this right, as most sites affected and also the effects are limited to the summer (May –September in the NH) season
- Urbanization


- Böhm, R., Jones, P.D., Hiebl, J., Frank, D., Brunetti, M. and Maugeri, M:, 2010: The early instrumental warm-bias: a solution for long Central European temperature series, 1760-2007. *Climatic Change* **101**, 41-67
- Brunet, M., Asin, J., Sigró, J., Bañon, M., García, F., Aguilar, E., Palenzuela, J.E., Peterson, T.C. and Jones, P.D., 2010: The minimization of the screen bias from ancient Western Mediterranean air temperature records: an exploratory statistical analysis. *Int. J. Climatol*. **31**, 1879-1895, DOI: 10.1002/joc.2192.

# Early exposure issues

- Europe affected, before the development of Stevenson screens
- Solution has come about from modern parallel measurements (in Austria and Spain, with the old screens). Needs to be looked at elsewhere
- Effect is annually ~0.3°C, with most series too warm by up to 0.7°C in June
- Surprisingly (for Austria), the effect is much smaller using the (Tx+Tn)/2 method of calculating averages than using the fixed hours method used in Austria
- Issue important as it is the summers that calibrate most natural and documentary proxies

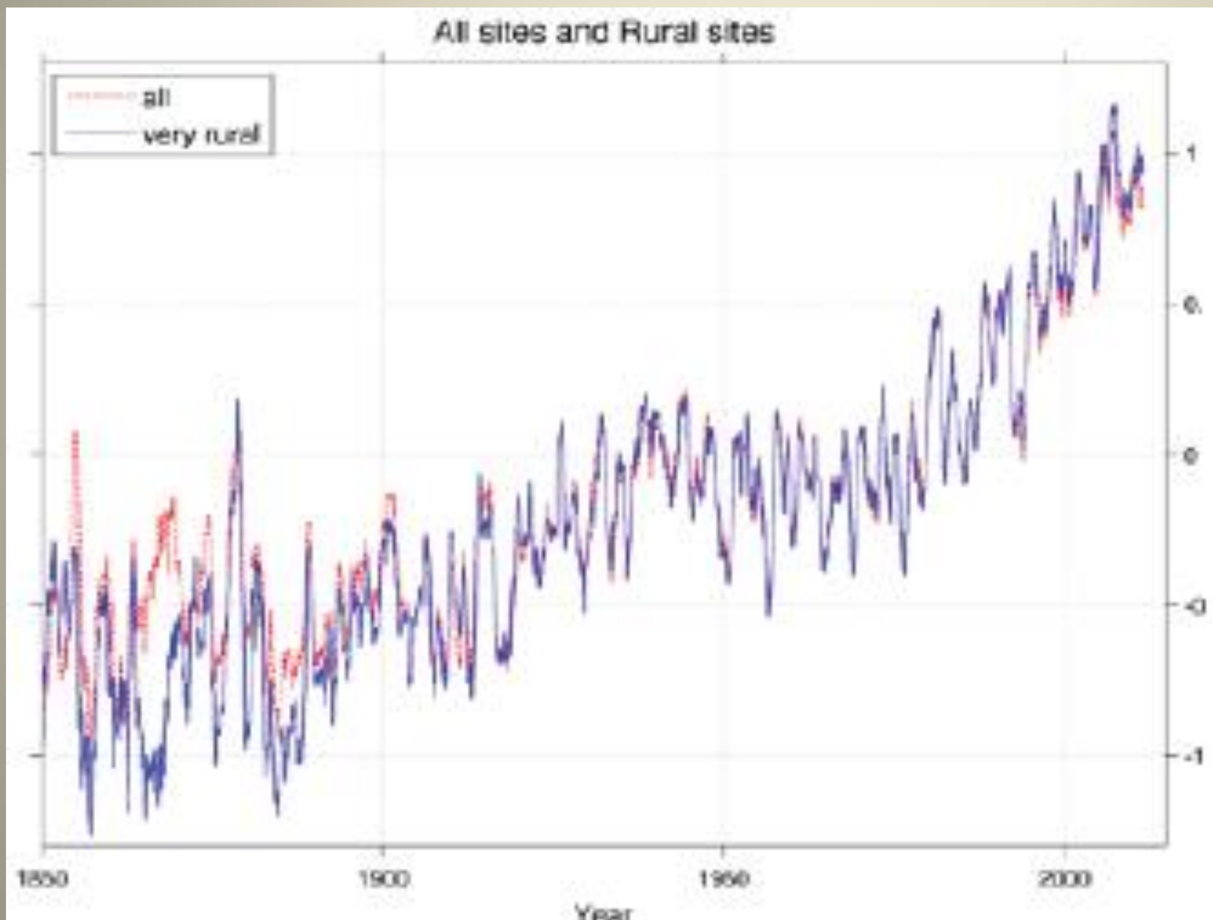
# Urban Heat Island - London




UHI greater for Tn than Tx. Central London sites always warmest at night, but warmer during day west of London

London has an Urban Heat Island (UHI), but no urban-related warming since at least 1900. In other words, the centre got warmer earlier.

# Large-scale urbanization influence is negligible

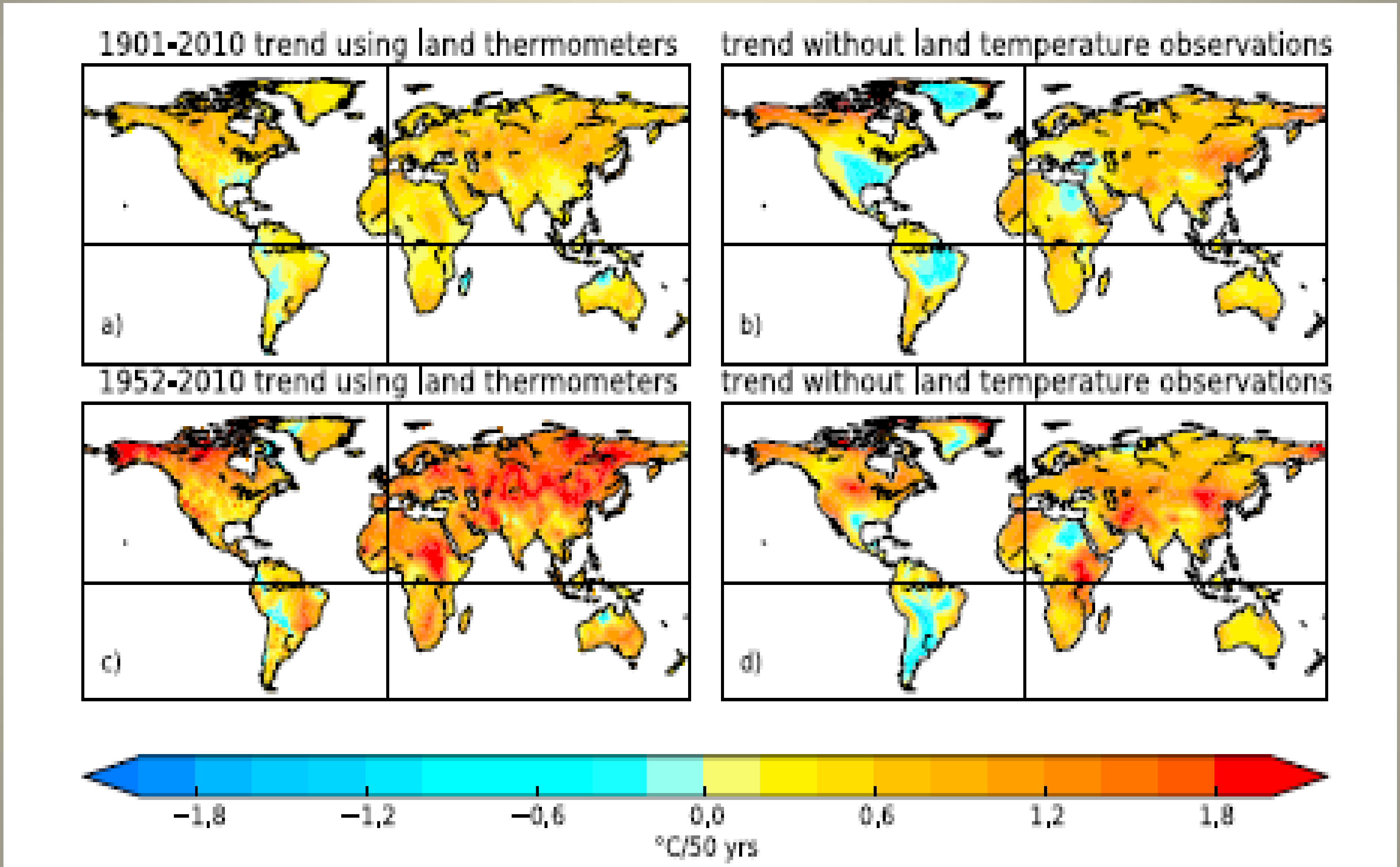



In this recent analysis by the BEST team, the very rural sites warm slightly more than the urban sites

Site category determined using satellite (MODIS) data

Wickham C, Rohde R, Muller RA, Wurtele J, Curry J, et al. (2013) Influence of Urban Heating on the Global Temperature Land Average using Rural Sites Identified from MODIS Classifications. Geoinfor Geostat: An Overview 1:2. doi:10.4172/gigs.1000104

# Using Reanalysis to Assess Urban Effects
# 20th Century Reanalysis (20CR)




CRU TS 3.10 20CR

Compo, G. et al 2013: Independent confirmation of global land warming without the use of station temperatures. *Geophys. Res. Letts*. **40**, 3170-3174, doi:10.1002/grl.50425

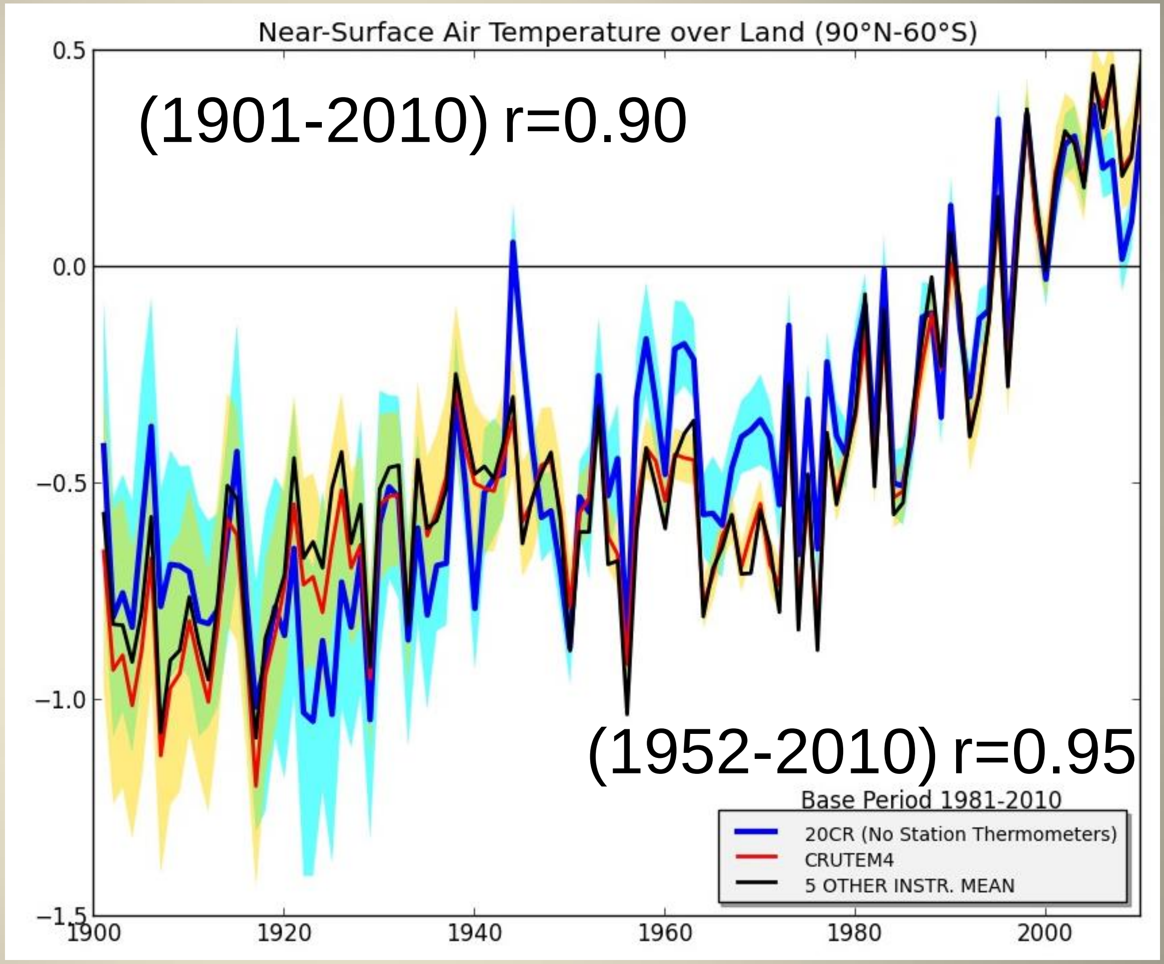
# *Annual anomalies of $TL_{2m}$ from 20CR, CRUTEM4, average of 5 other instrumental datasets (1901-2010)*

Other $TL_{2m}$ <u>datasets</u>

*1901-2010:*
r=0.84 to 0.92

*1952-2010:*
r=0.95 to 0.96

Shading:
95% confidence interval



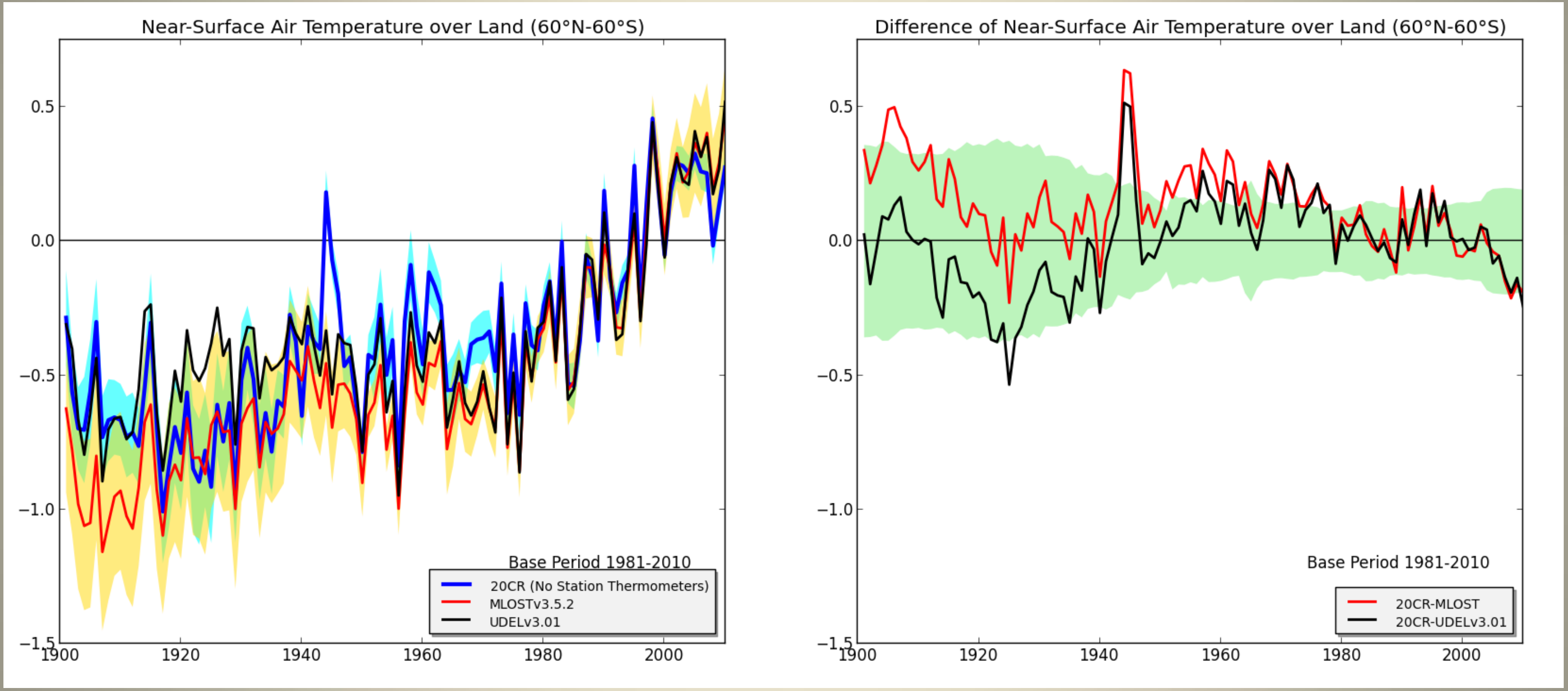

1901 2010

$TL_{2m}$ from stations and 20CR has consistent large-scale annual, decadal, and centennial $TL_{2m}$ variations

Annual anomalies of near-global $TL_{2m}$ from 20CR MLOST UDEL (1901-2010)

TL2m

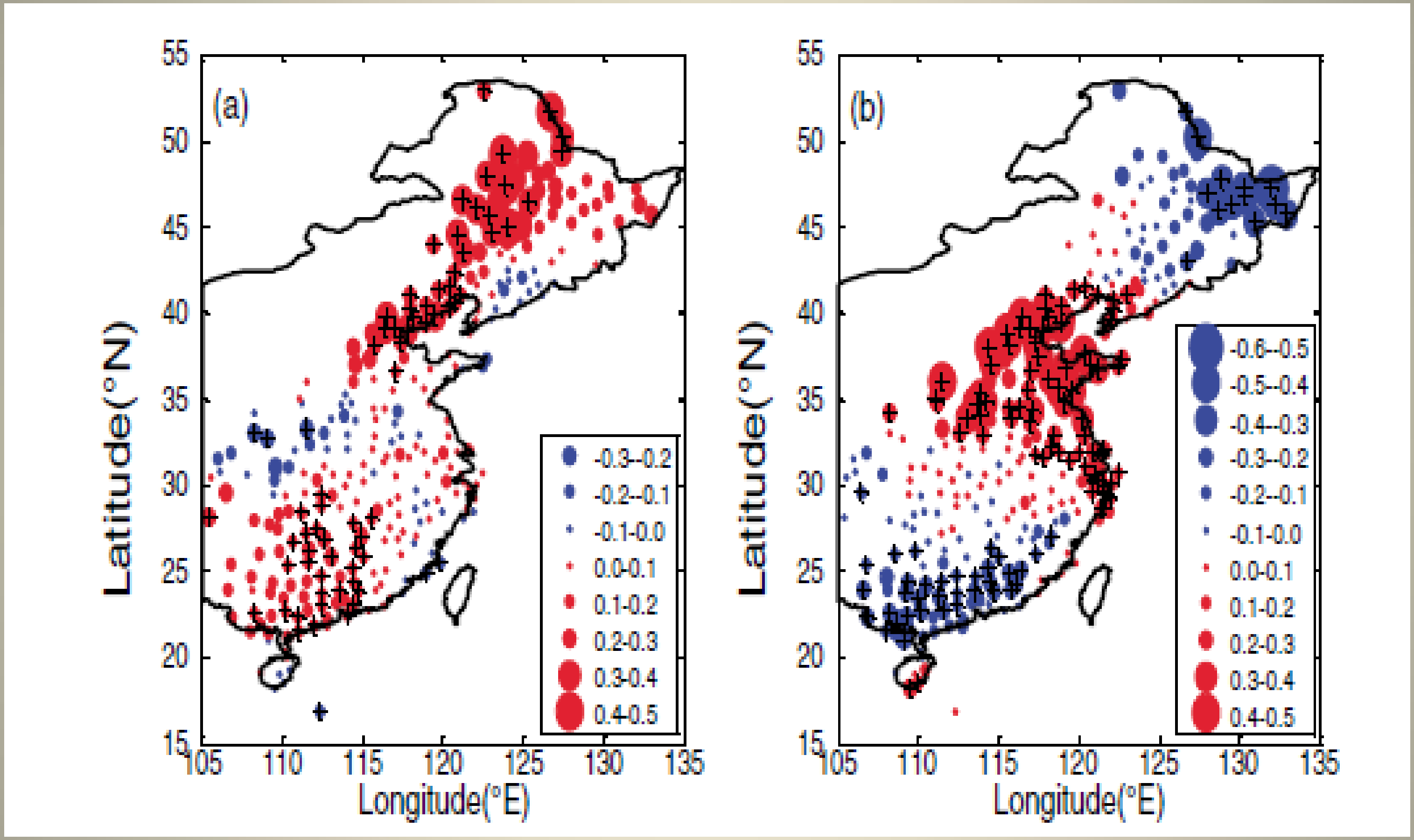
Difference (20CR – MLOST)




1901 2010 1901 2010

Uncertainty estimates are largely consistent.
Differences are not "urban warming".

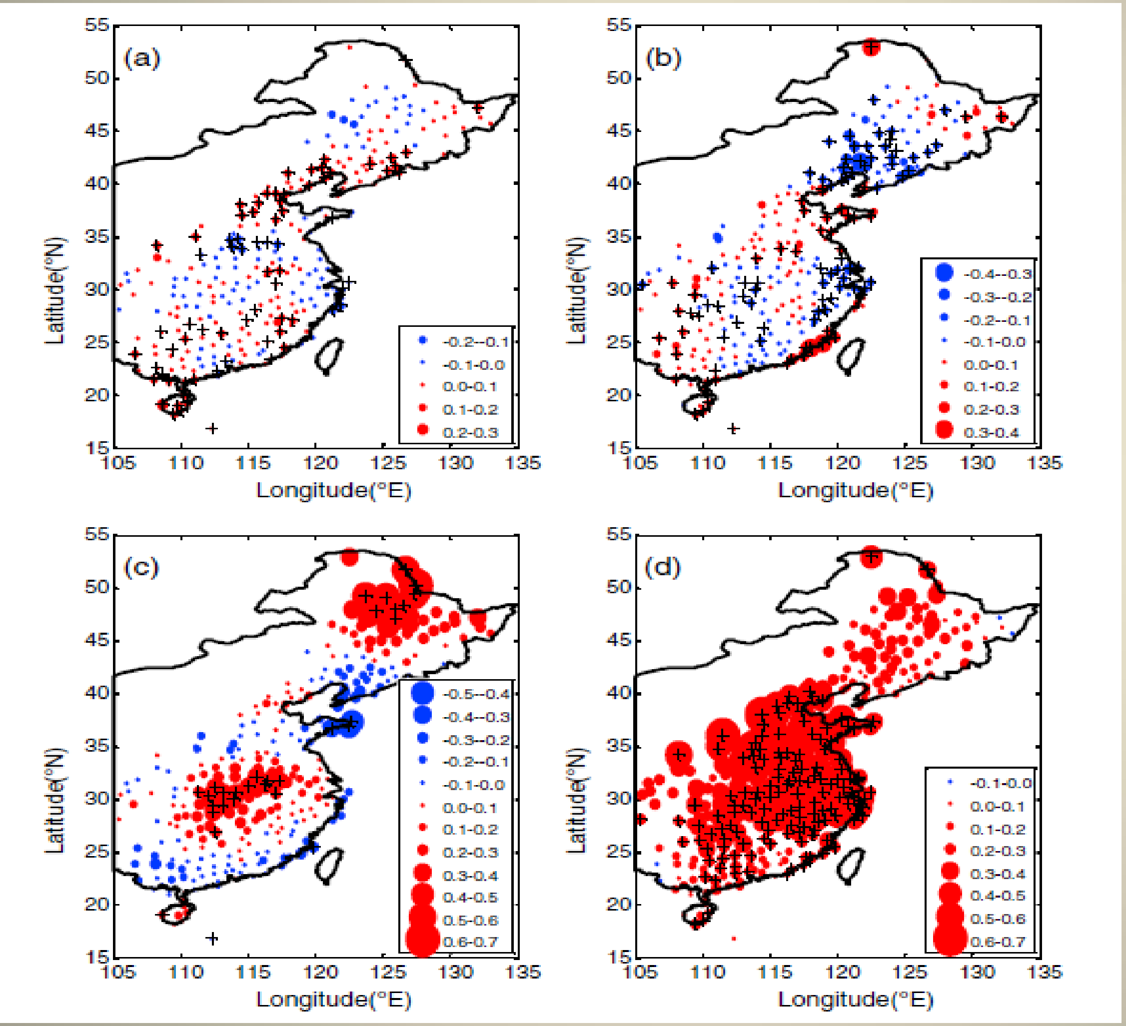
# Using Reanalyses to assess Urban Influence (2): OMR (Observations minus Reanalysis)




Chinese surface temperatures minus NCEP Reanalyses, Left: 1979-1998 and Right: 1989-2008

Wang, J., Yan, Z, Jones, P.D. and Xia, J., 2013: On 'Observation minus Reanalysis' method: A view from Multi-Decadal Variability. *J. Geophys. Res*. **118**, doi:10.1002/jgrd.50574.

# Same Approach: Different Reanalyses




ERA-Interim

Left:1979-1998

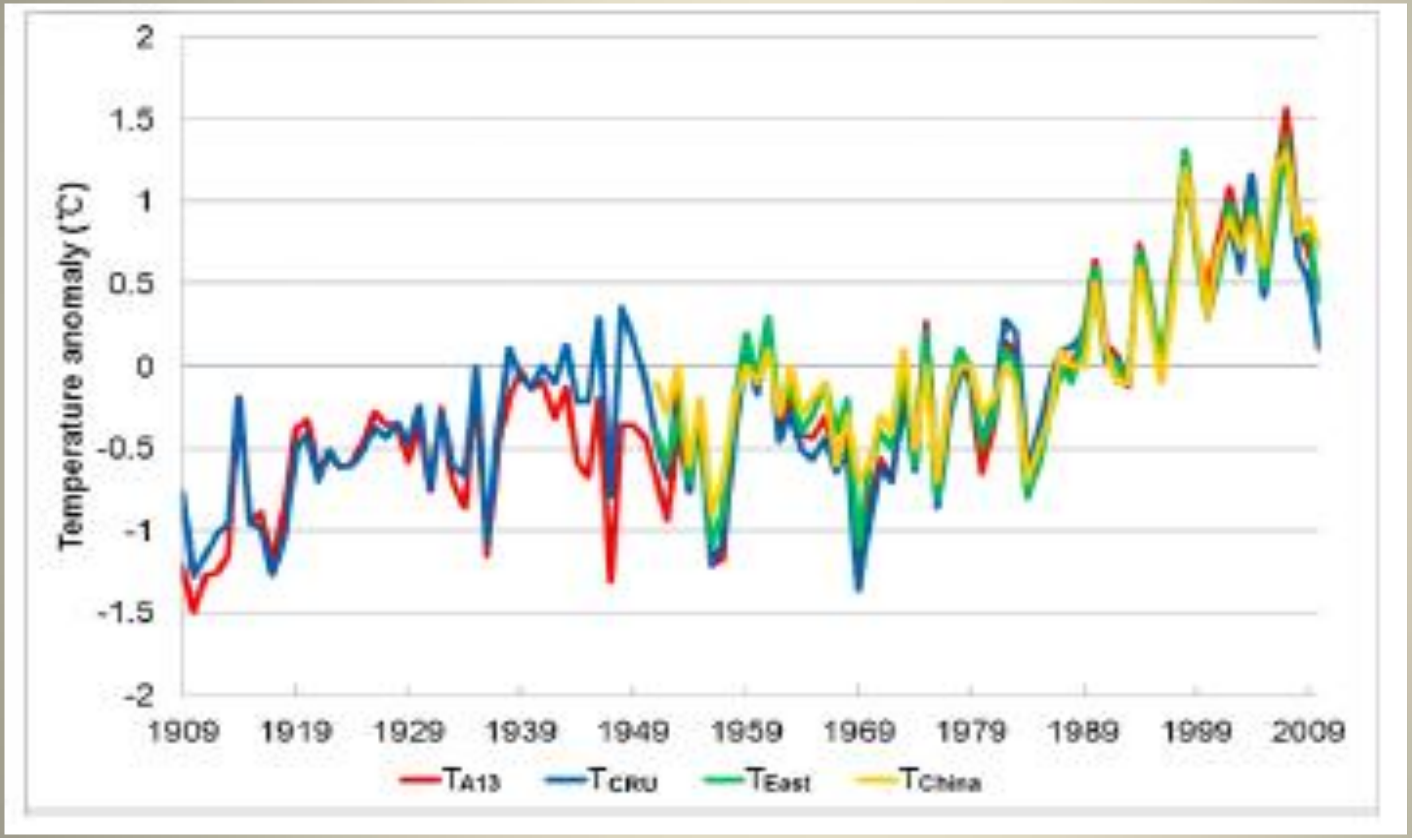
Right: 1989-2008

20CR

Left: 1979-1998

Right: 1989-2008

# Chinese average temperatures




Yellow – All China (500+)

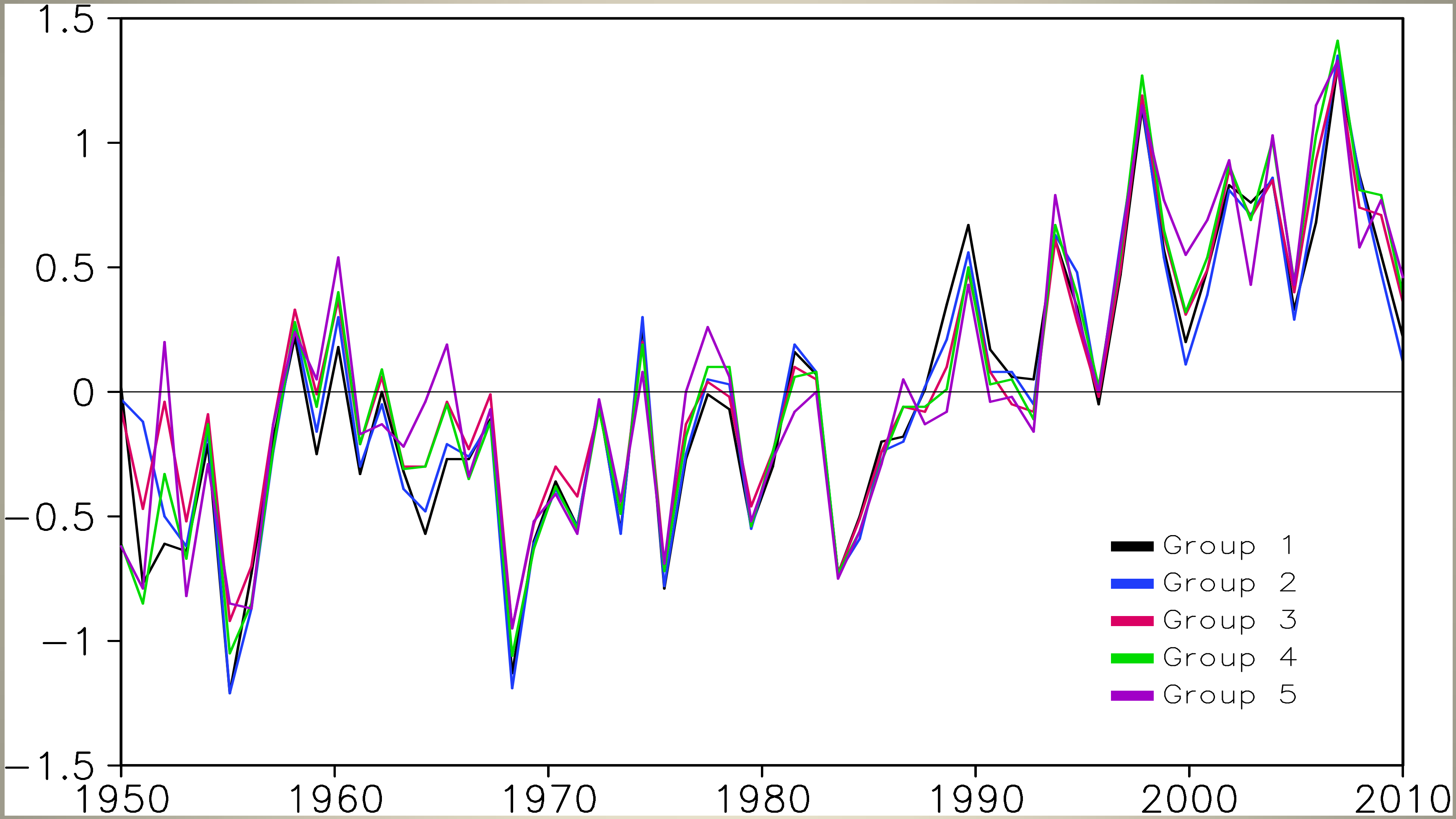
Red – 13 sites from Eastern third of China

Green – Eastern third of the all China series

Blue - CRU

Cao, L., Zhao, P., Yan, Z., Jones, P.D., Zhu, Y., Yu, Y. and Tang, G., 2013: Instrumental temperature series in eastern and central China back to the 19th century. *J. Geophys. Res.* **118**, doi/10.1002/jgrd.50615.

# Chinese Temperatures and Urbanization

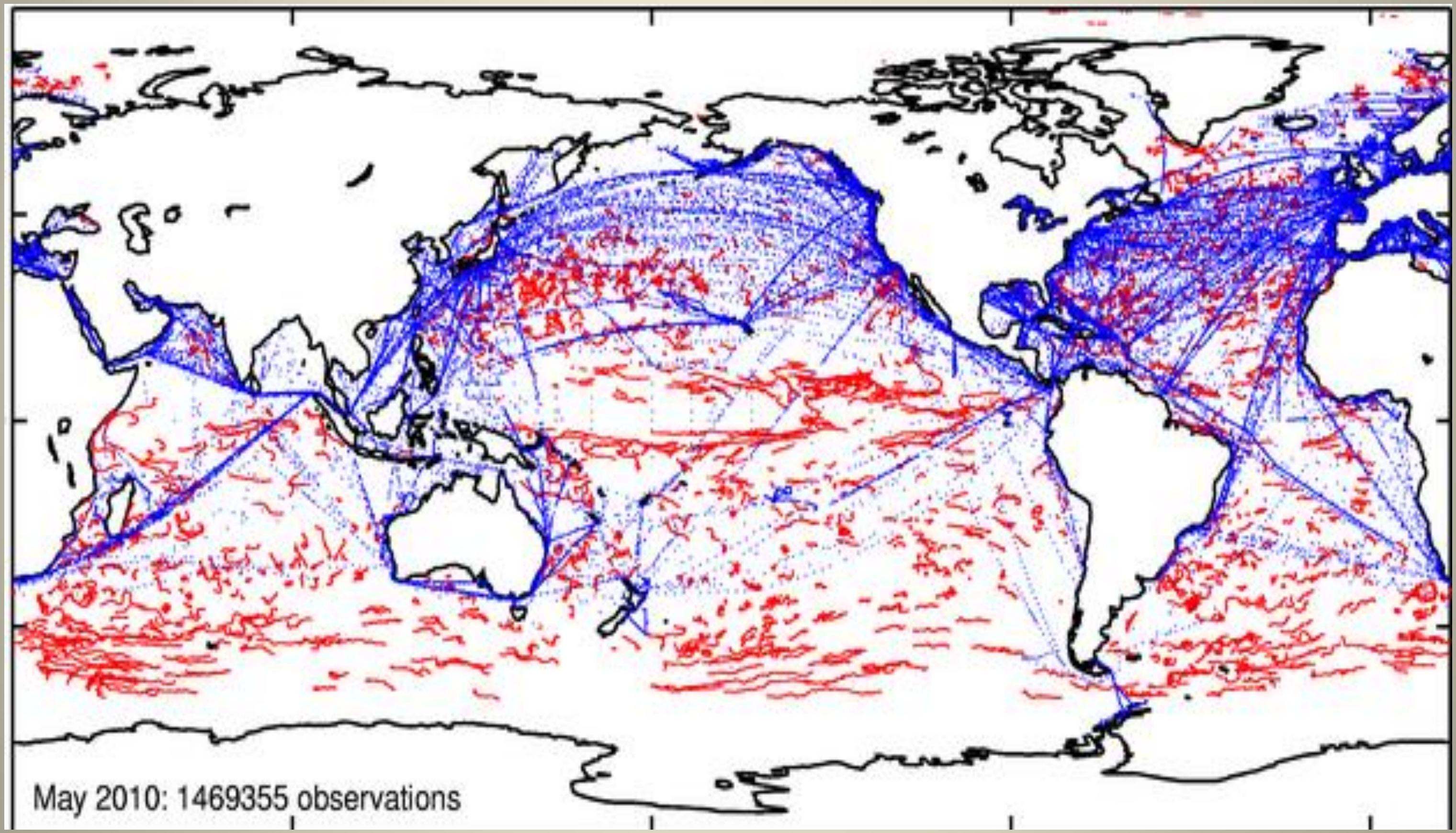



Eastern half of China, with 245 stations split into 5 groups based on population size

1: Lowest population group (<300K), 5: Highest population group (>900K)

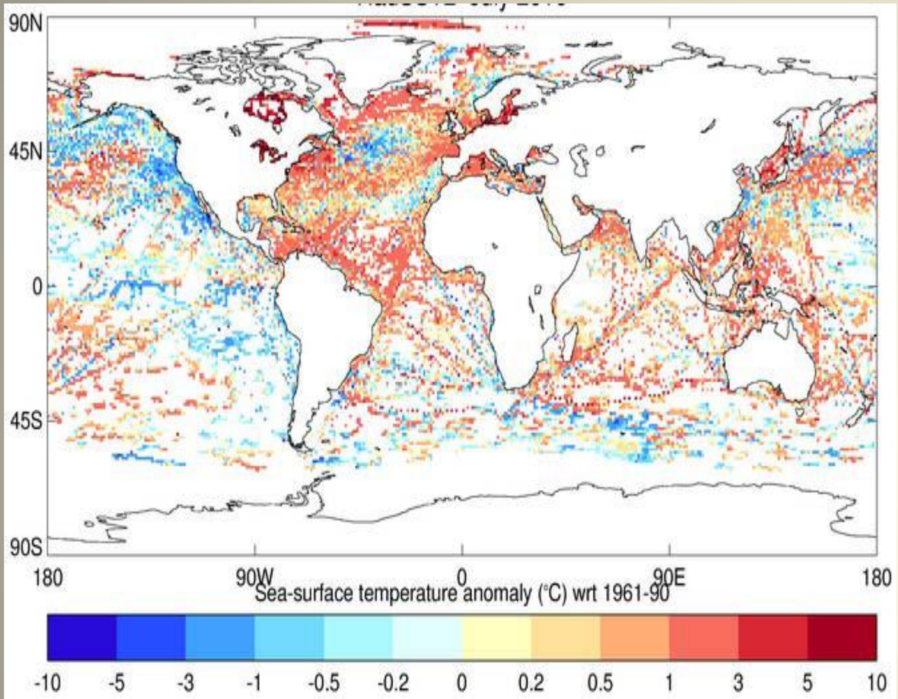
No stations in truly rural areas, like National Parks

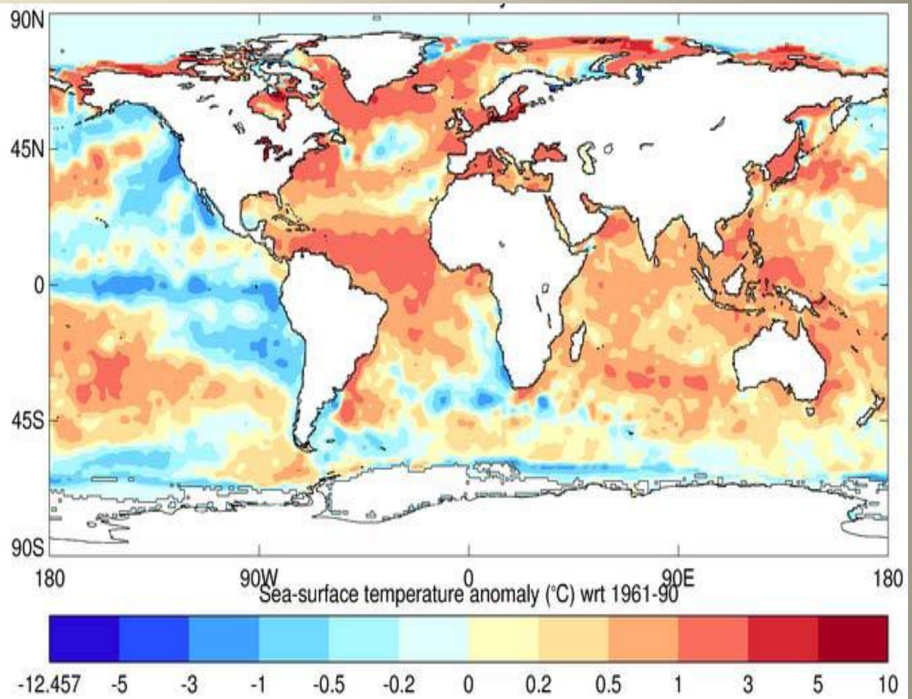
# SST Observations – May 2010




Blue – ships; Red – drifting buoys; Grey – fixed buoys (mainly TOGA-TAO array)

# SST Interpolation




Kennedy J.J., Rayner, N.A., Smith, R.O., Saunby, M. and Parker, D.E. (2011a). Reassessing biases and other uncertainties in sea-surface temperature observations since 1850 part 1: measurement and sampling errors. J. Geophys. Res.116, D14103, doi:10.1029/2010JD015218.

Kennedy J.J., Rayner, N.A., Smith, R.O., Saunby, M. and Parker, D.E. (2011b). Reassessing biases and other uncertainties in sea-surface temperature observations since 1850 part 2: biases and homogenisation. J. Geophys Res. 116, D14104, doi:10.1029/2010JD015220,

Rayner, N. A., P. Brohan, D. E. Parker, C. K. Folland, J. J. Kennedy, M. Vanicek, T. J. Ansell, and S. F. B. Tett (2006), Improved analyses of changes and uncertainties in sea-surface temperature measured in-situ since the mid-nineteenth century, **J. Clim.**, *19*, 446– 469.

# SST issues

- Principal problem is the changeover to engine intake measurements from buckets
- Countries and shipping (merchant and naval) fleets did this at different times
- Bucket design also varied between different shipping fleets
- The way the SST measurement was made was not put with the data until the early 1970s
- Dates and bucket types have only been discovered by looking at old books of instructions to marine observers
- ERI – Engine Room Intakes
- VOS – Voluntary Observing Ships
- Modern SST data come in with ship call signs and locations – problem is that the shipping fleets are becoming more reluctant to take the data – for security and trade/economic issues (e.g. they don't want others to know where they are – fishing fleets)

- SSTs are vital to many other areas of atmospheric sciences. They are necessary as the boundary values for weather forecasts and also Reanalyses.


- Thompson, D.W.J., Kennedy, J.J., Wallace, J.M. and Jones, P.D., 2008: A large discontinuity in the mid-twentieth century in observed global-mean surface temperature. *Nature* **453**, 646-649.

# HadSST3 versus raw observations (red)

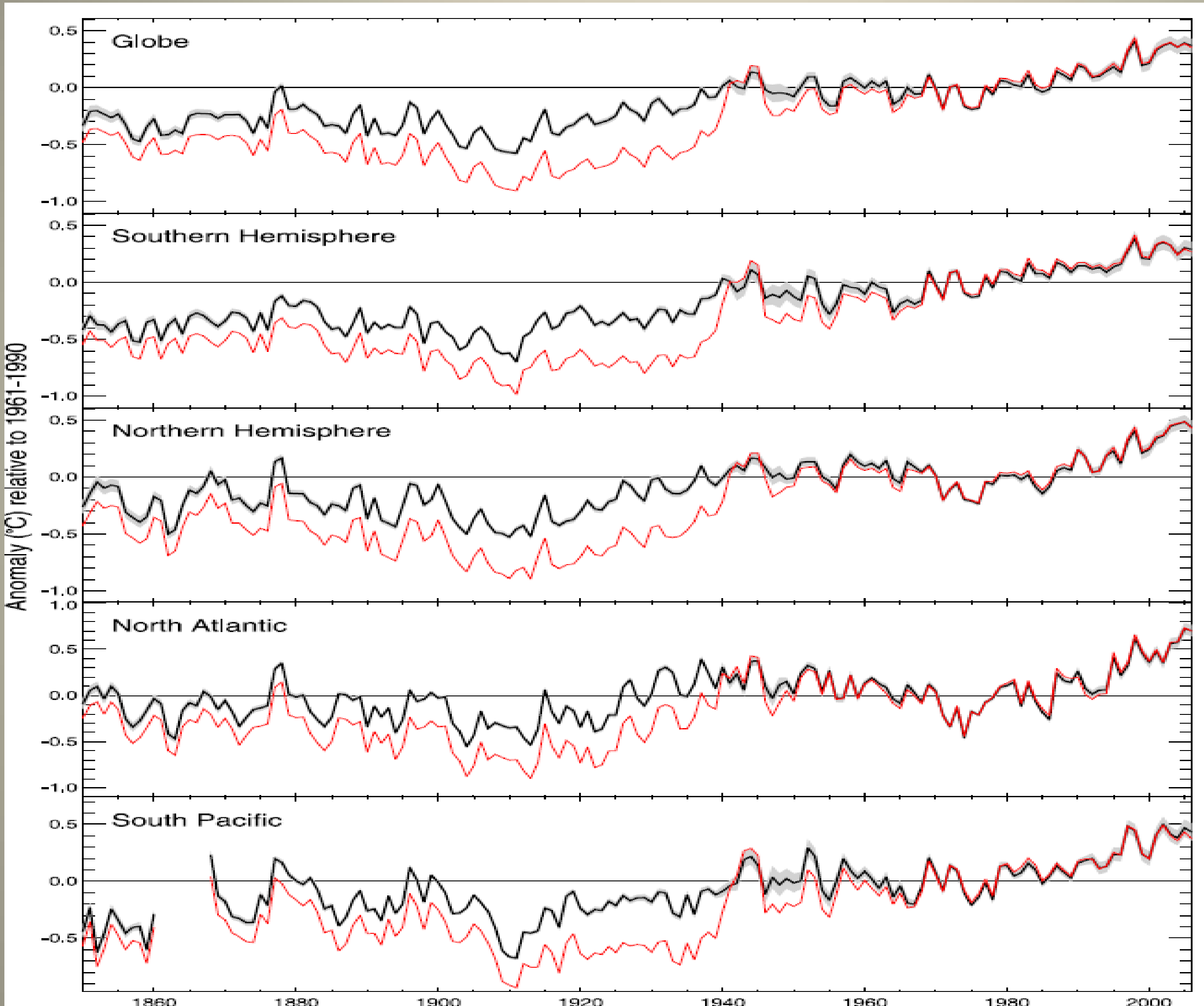



**Figure 4.** Regional annual average SST anomalies 1850–2006 (relative to 1961–1990) for 100 realizations of HadSST3 (black line plus 2 standard deviations grey area) and unadjusted gridded data (red line).

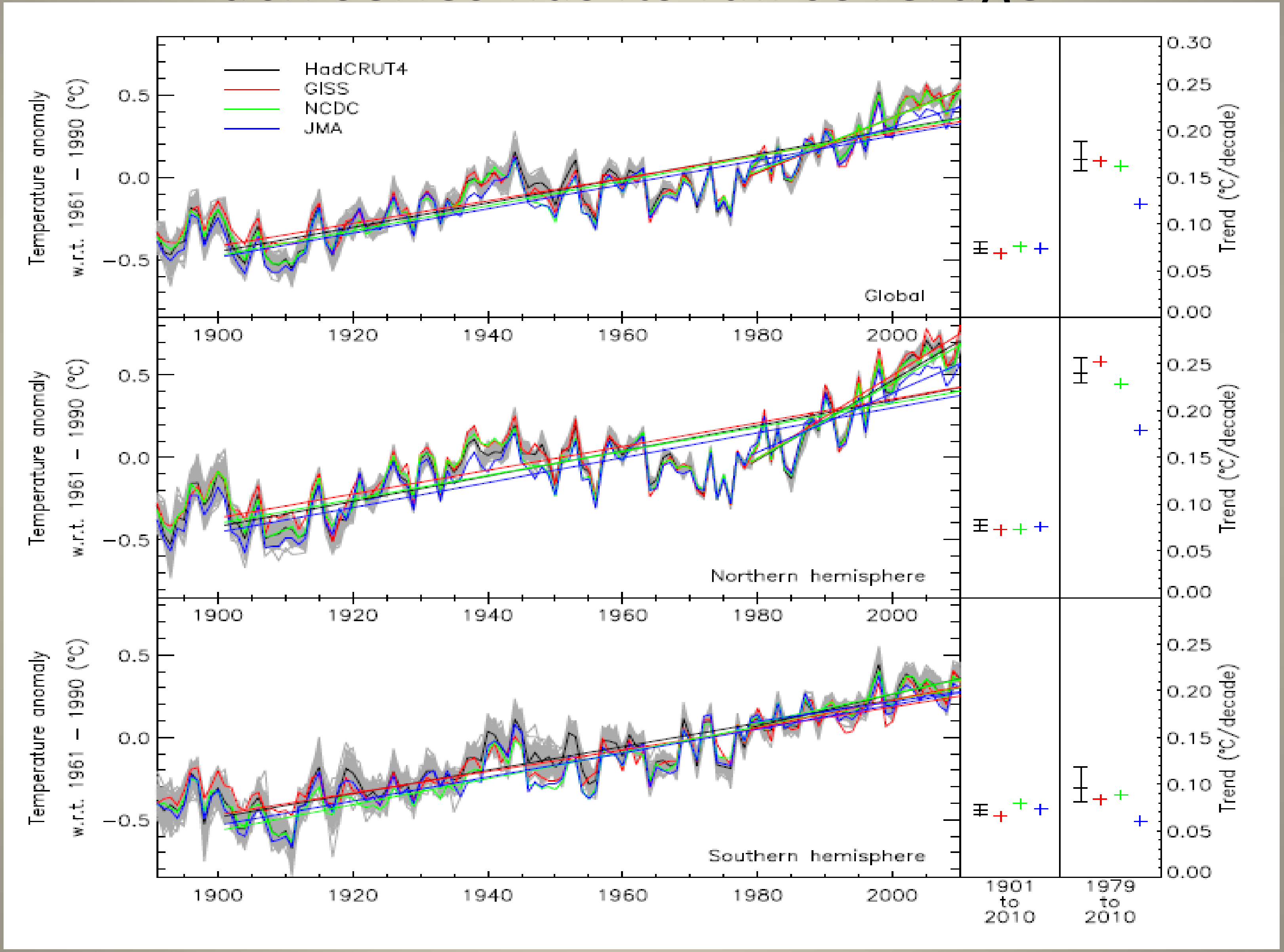
Grey band is error associated with the assumptions made

Definitions of the regions given in Kennedy *et al* (2011b).

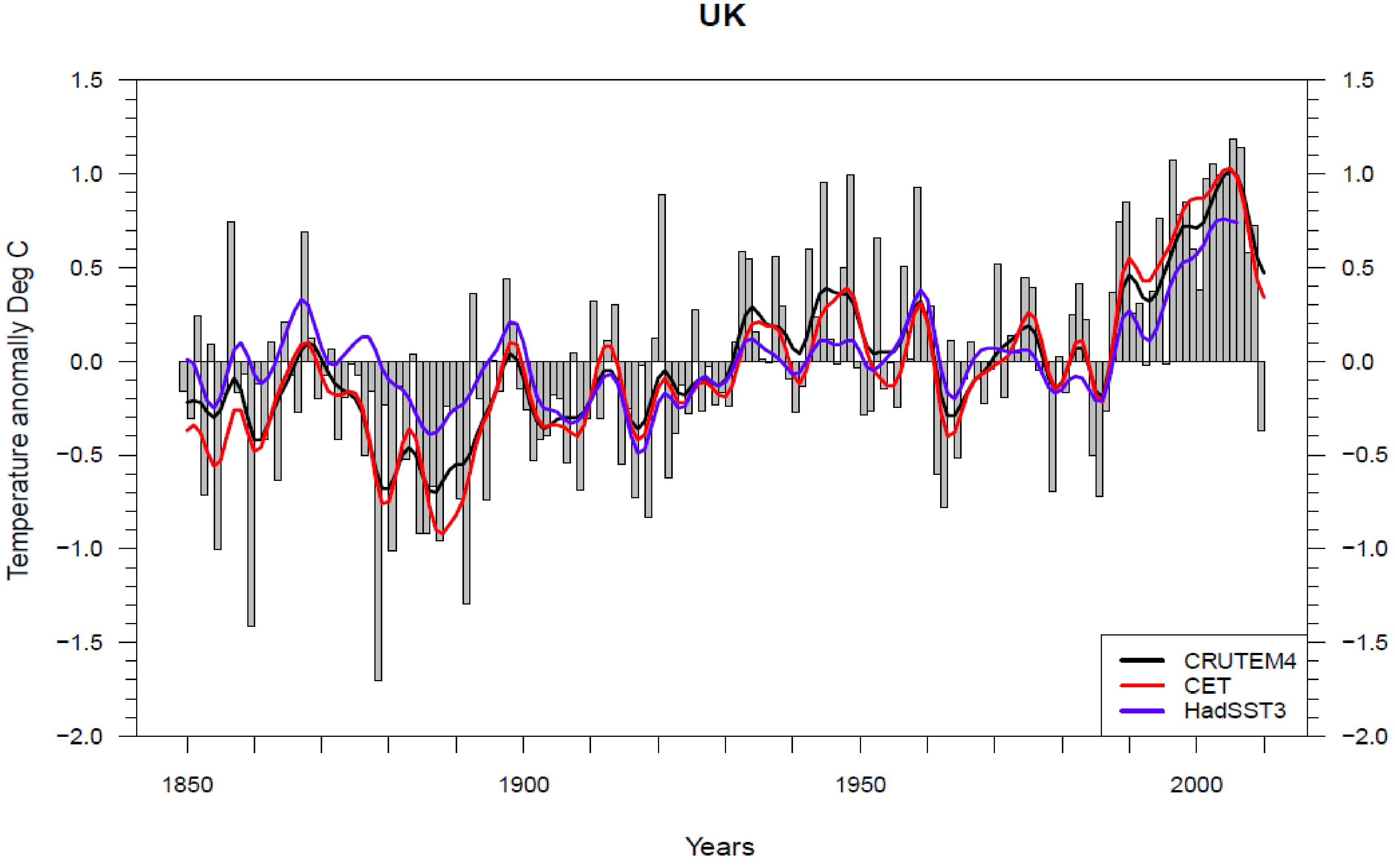
Major change from HadSST2 is in the period 1945-55

Canvas Buckets (1900-1941)
Wooden Buckets (19th Century)

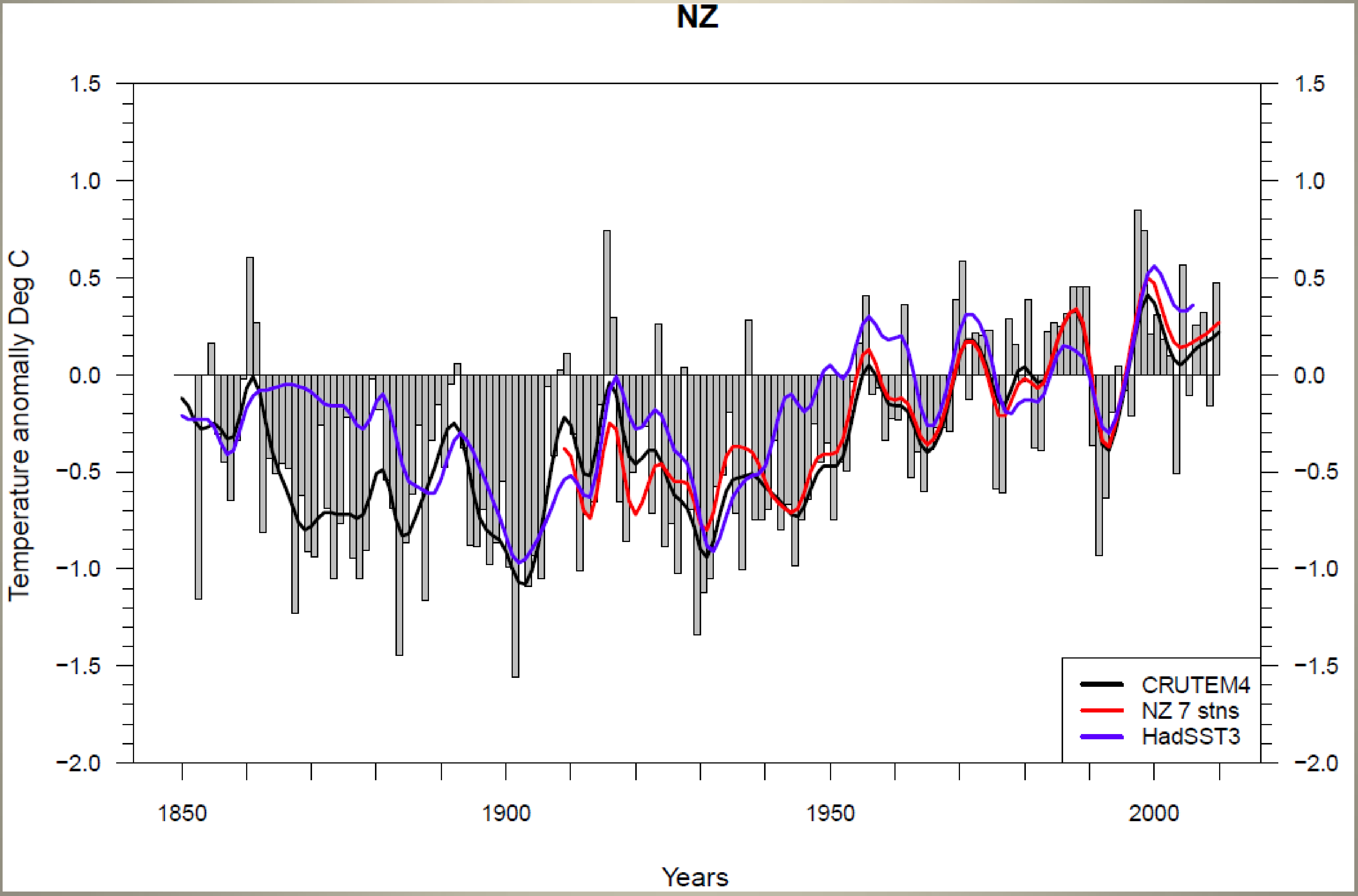
# HadCRUT4 vs other groups
# Each series has its full coverage



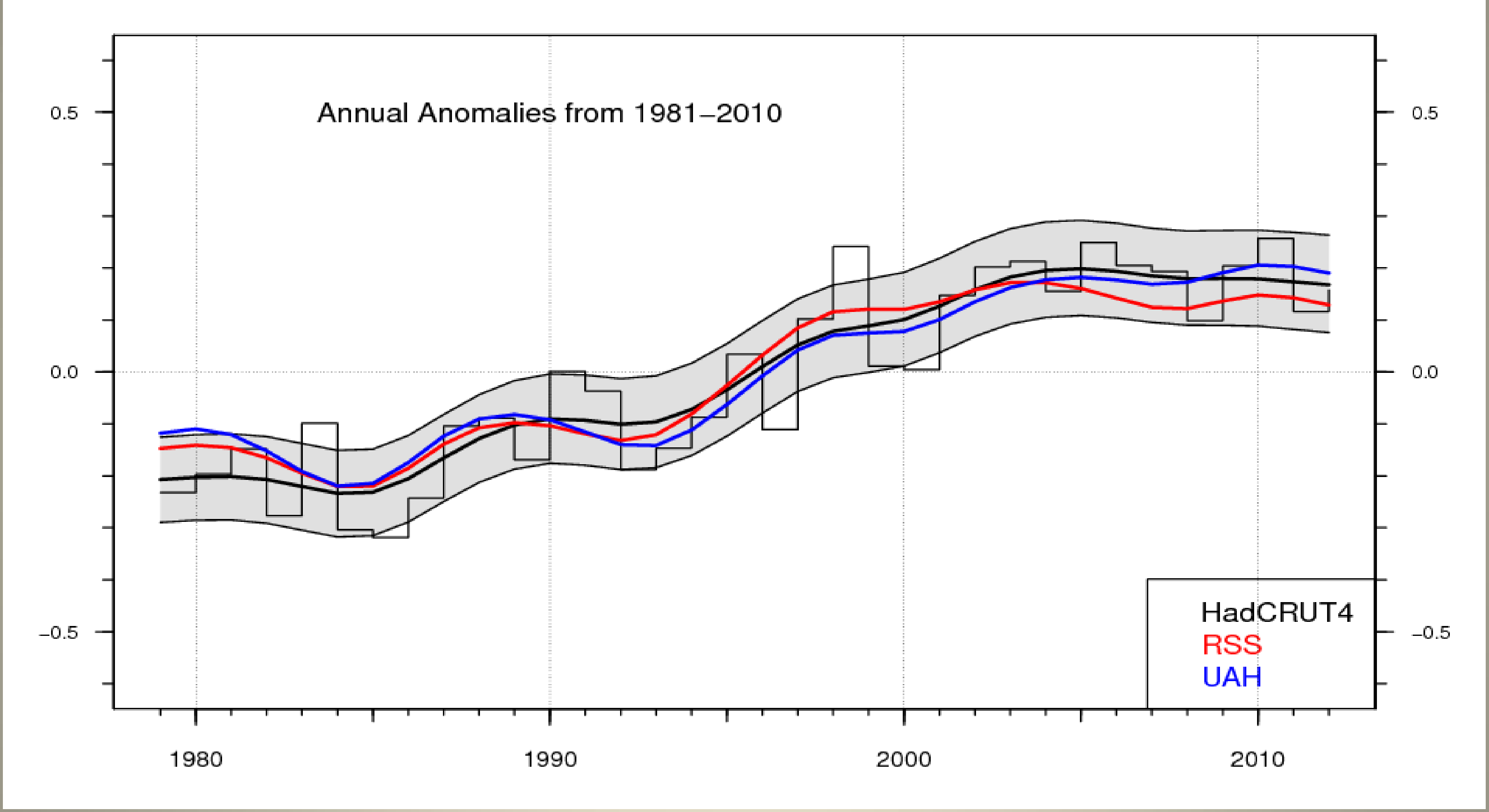
# British Isles (50-60°N, 0-10°W) - annual



# New Zealand (165-180°E, 35-50°S) - annual



# HadCRUT4 compared to Satellite Estimates

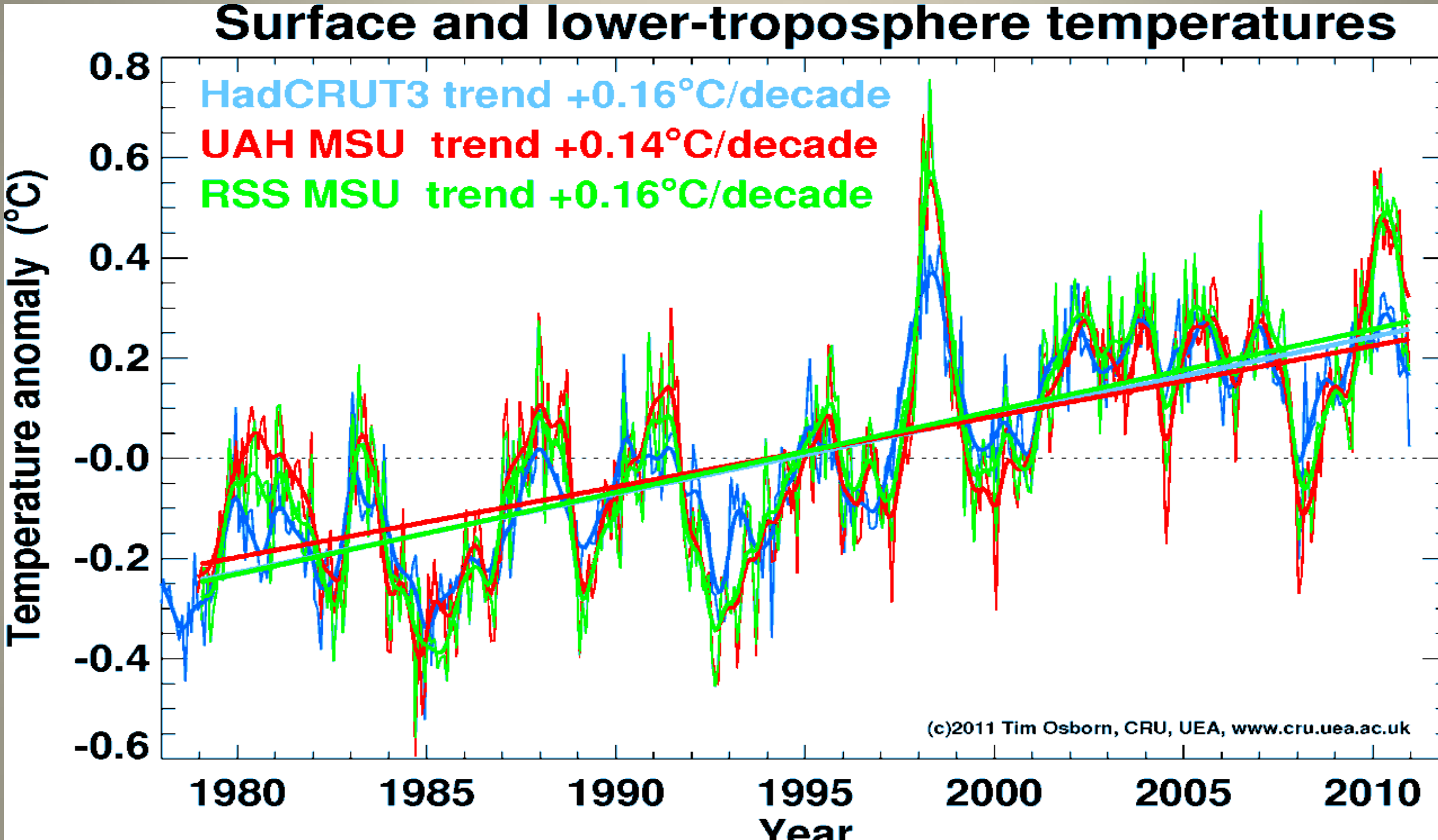



HadCRUT4 is the combined land/marine temperature curve

Surface and satellite estimates of temperature change agree over this 34-year period.

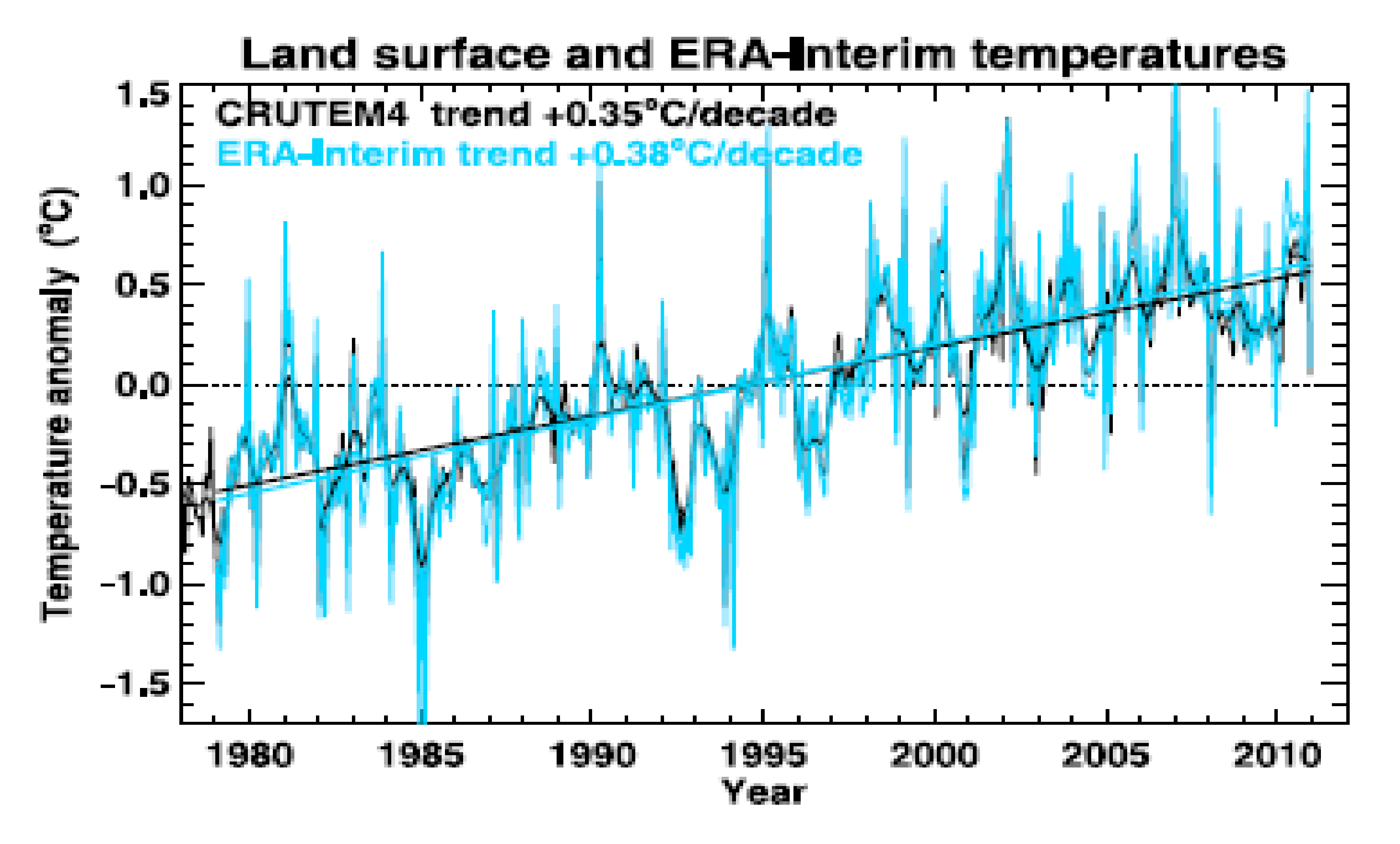
Satellite estimates are for the lower troposphere, centred at about 650hPa, or about 3-4km above the surface

# Surface vs Satellite



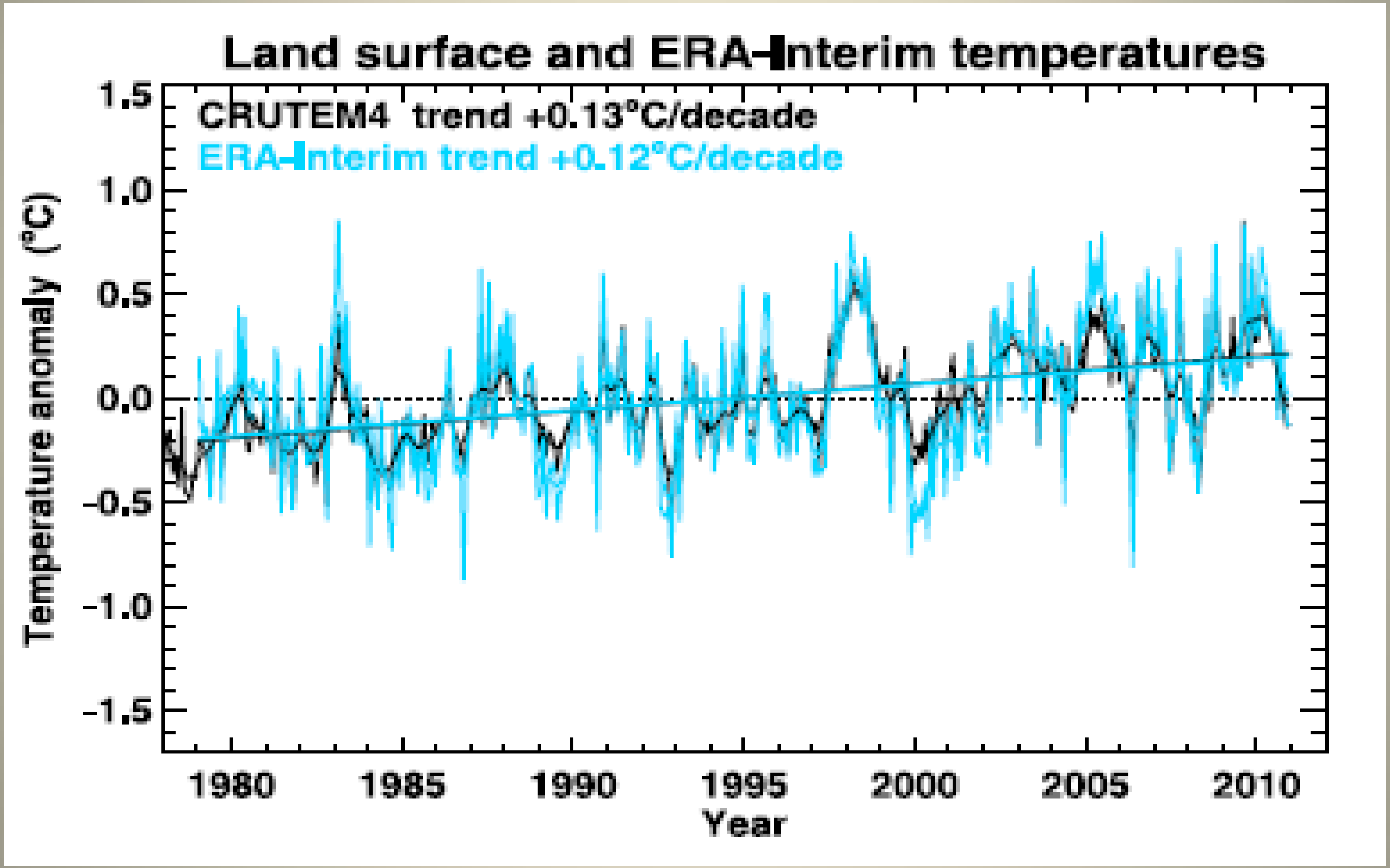

Surface and satellite records of temperature AGREE!! 2010 is exceptionally warm with the satellites too!

# Comparison with ERA-Interim (NH)
# Land only




ERA-Interim complete coverage for NH, so warms slightly more than CRUTEM4

# Comparison with ERA-Interim (SH 0-60S) Land only




ERA-Interim still quite poor compared to in situ observations across Antarctica

# Absolute Temperatures

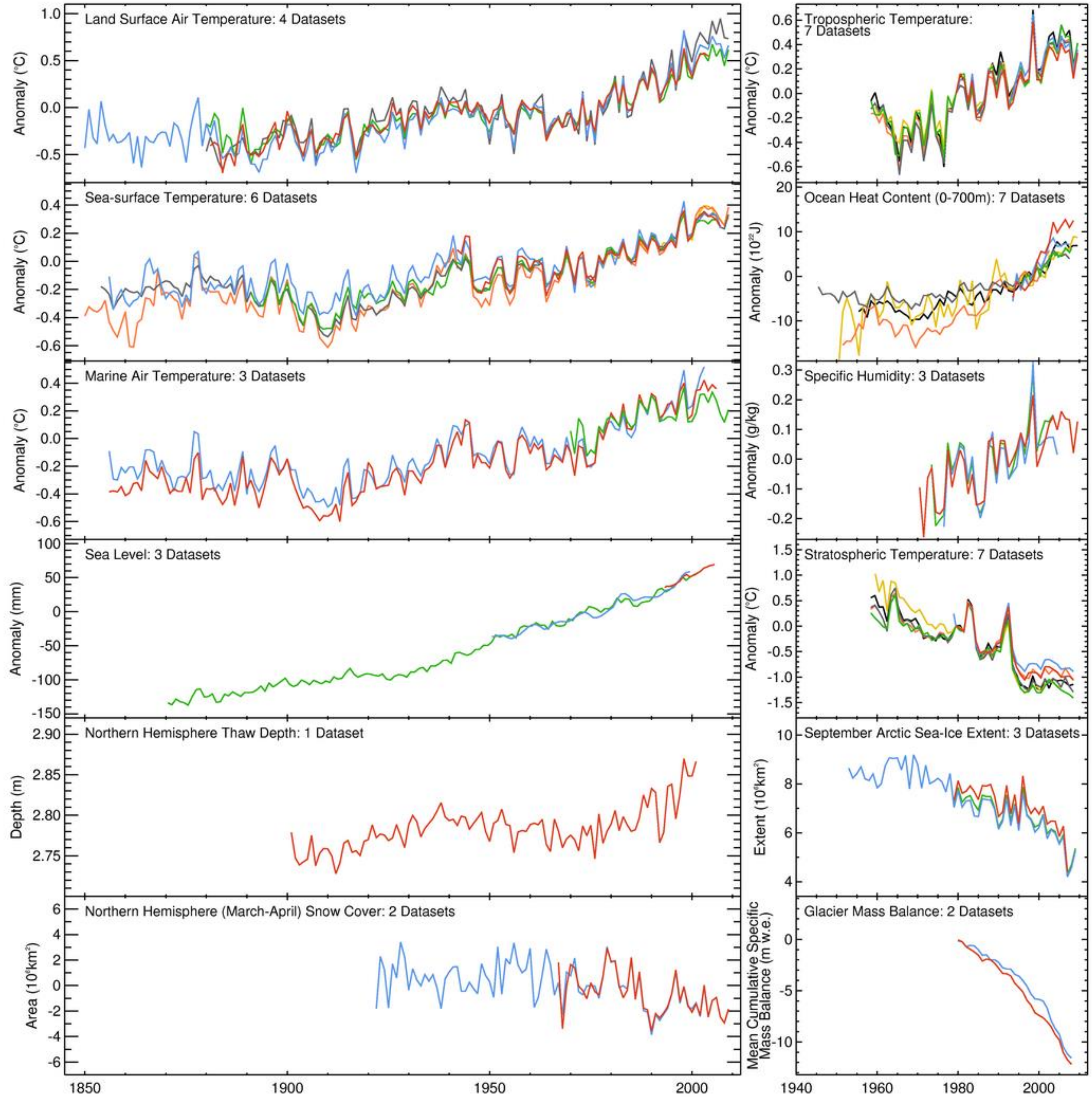
- Time series always shown as anomalies from a base period – often 1961-90
- I'm occasionally asked what the absolute average temperature of the world is
- It is ~14°C for the 1961-90 period (Jones et al. 1999)
- Recently compared with ERA-Interim (Jones and Harpham, 2013)
- Large differences over parts of the Antarctic and to a lesser extent over the Arctic, but ERA-Interim average would be about 0.2°C cooler (for 1961-90) based on comparisons over the 1979-2012 period
- So because 1981-2010 is warmer than 1961-90, ERA-Interim gives ~14°C, but there are large compensating differences with the limited surface network over Antarctica (ERA-Interim is warmer in the interior and cooler at lower latitudes, 65-75°S)


- Jones, P.D., New, M., Parker, D.E., Martin, S. and Rigor, I.G., 1999: Surface air temperature and its variations over the last 150 years. *Reviews of Geophysics* **37**, 173-199.
- Jones, P.D. and Harpham, C., 2013: Estimation of the absolute surface air temperature of the Earth, *J. Geophys. Res*. **118**, 3213-3217, doi:10.1002/jgrd.50359.

# Conclusions

- Biases generally much more important than individual station homogeneity issues
- Urbanization issues relatively unimportant at large-scales, but maybe issues at local scales
- Exposure issues pre-Stevenson screens an important issue in Europe before about 1880. Important for proxy climate calibration in Europe
- Largest bias is in SST data and related to changes in SST measurement from buckets to engine intakes – effect after adjustment is reduction in long-term warming
- Early assessments (e.g. Callendar in the 1930s) agree will with modern land-based estimates
- They were reasonable due to limited number of spatial degrees of freedom. We don’t need tens of thousands of stations to measure large-scale averages. We do need lots of stations to get local details
- Large-scale temperature estimates agree with satellite and reanalysis estimates

# Extra Slides




# OTHER TEMPERATURE – RELATED VARIABLES

12 variables

48 data sets

Figure taken from

*STATE OF THE CLIMATE IN 2009*
Special Supplement to the Bulletin of the American Meteorological Society
Vol. 91, No. 6, June 2010
D.S. Arndt, M.O. Baringer and M.R. Johnson, Eds.
Associate Eds. L.V. Alexander, H.J. Diamond, R.L. Fogt, J.M. Levy, J. Richter-Menge, P.W. Thorne, L.A. Vincent, A.B. Watkins and K.M. Willett

Based on some of these curves IPCC in 2007 said that the warming of the climate system was unequivocal