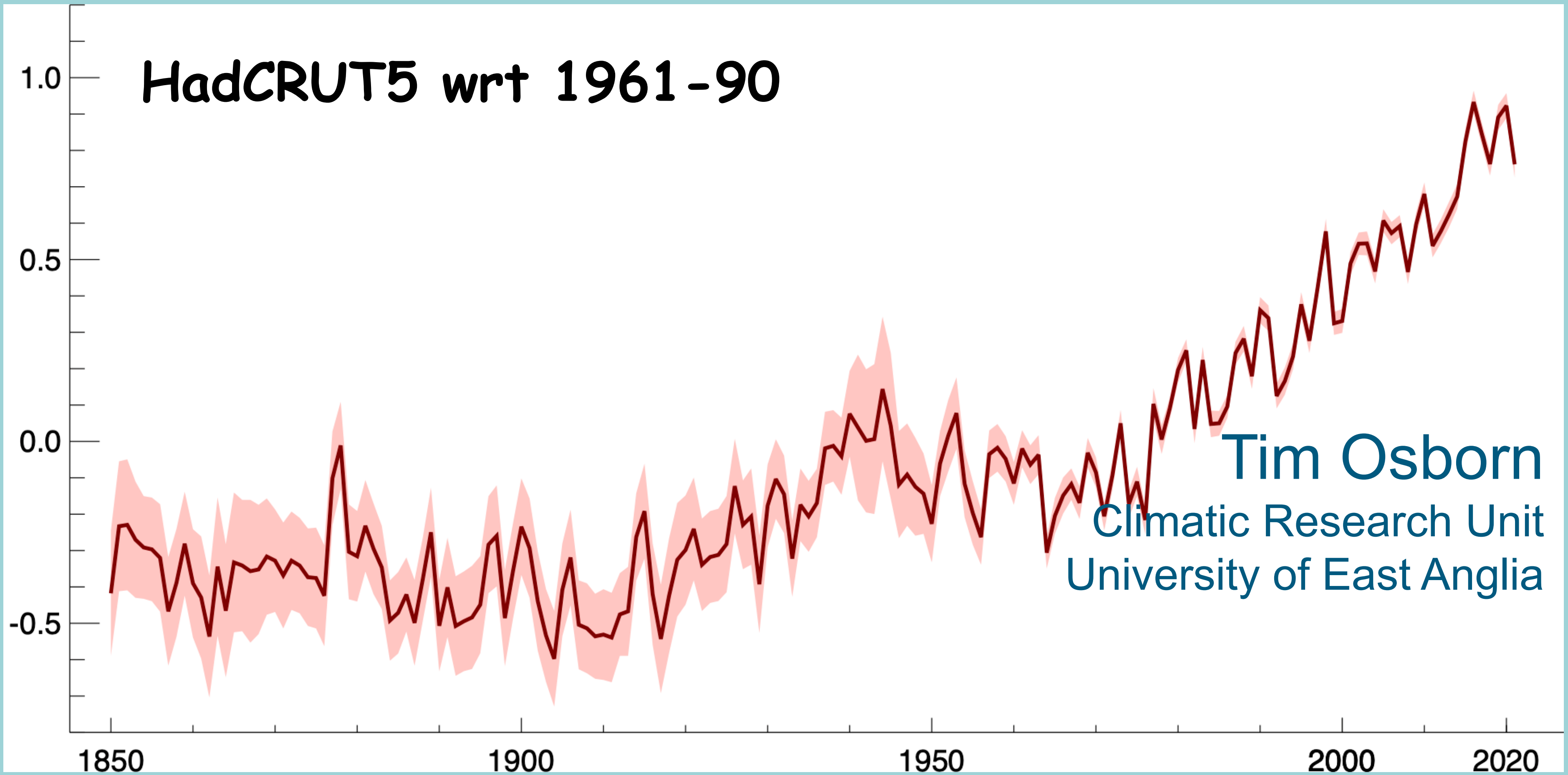
Global Temperature Series – History and the Important Issues

Phil Jones CRU, UEA

- Talk will discuss developments of these series from 1982 to the present
- Surface Air Temperature (SAT for Land Areas)
- Issues with SAT (Coverage, Homogeneity, Urbanization)
- Marine temperatures (Sea Surface Temperature, SST and Marine Air Temperature, MAT)
- Issues with SSTs (**These have the largest effects**)

- At the end, I'll show comparisons between CRU versions from the 1980s to the present, and then with US groups
- Very recent work – finding/using additional data

Development of global temperature datasets - the latest one, from 2021

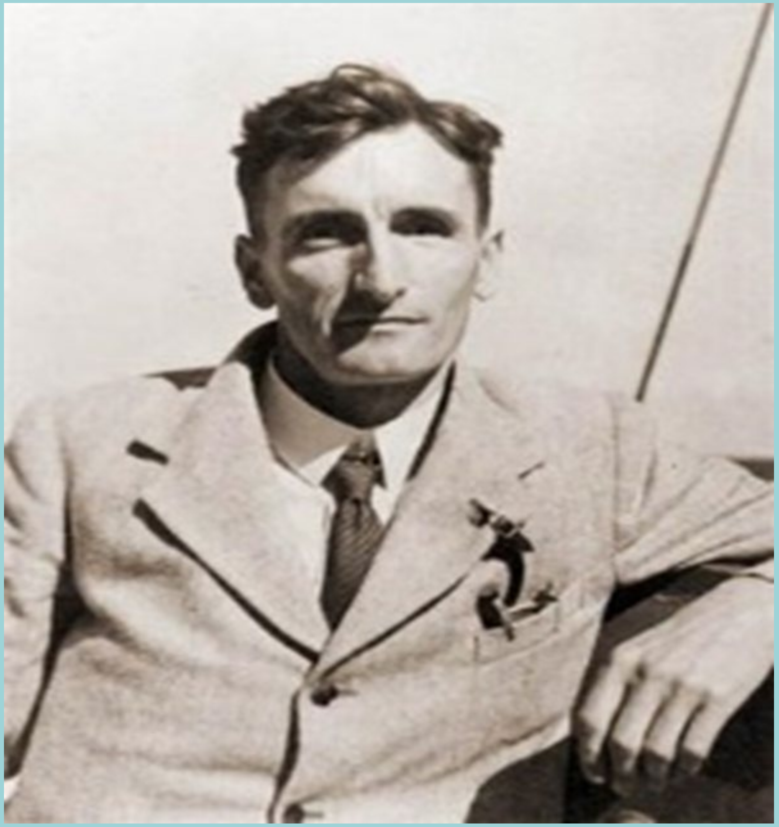


Thanks to Tim Osborn (CRU) and Liz Kent (NOC, Soton) for a few slides

The Early Global Temperature Series

- First was Wladimir Köppen in 1873 and 1881 (2 papers looking at links to solar output and effects of volcanic eruptions). Much better known for his climate classification which is still used today
- Later Guy Stewart Callendar in 1938, then updated in 1961 (In 1938 he also worked out the pre-industrial level of CO_2 value of 280 ppmv before this was confirmed by ice cores). Worth reading comments in QJRMS!! This paper and comments are on the RMS website.

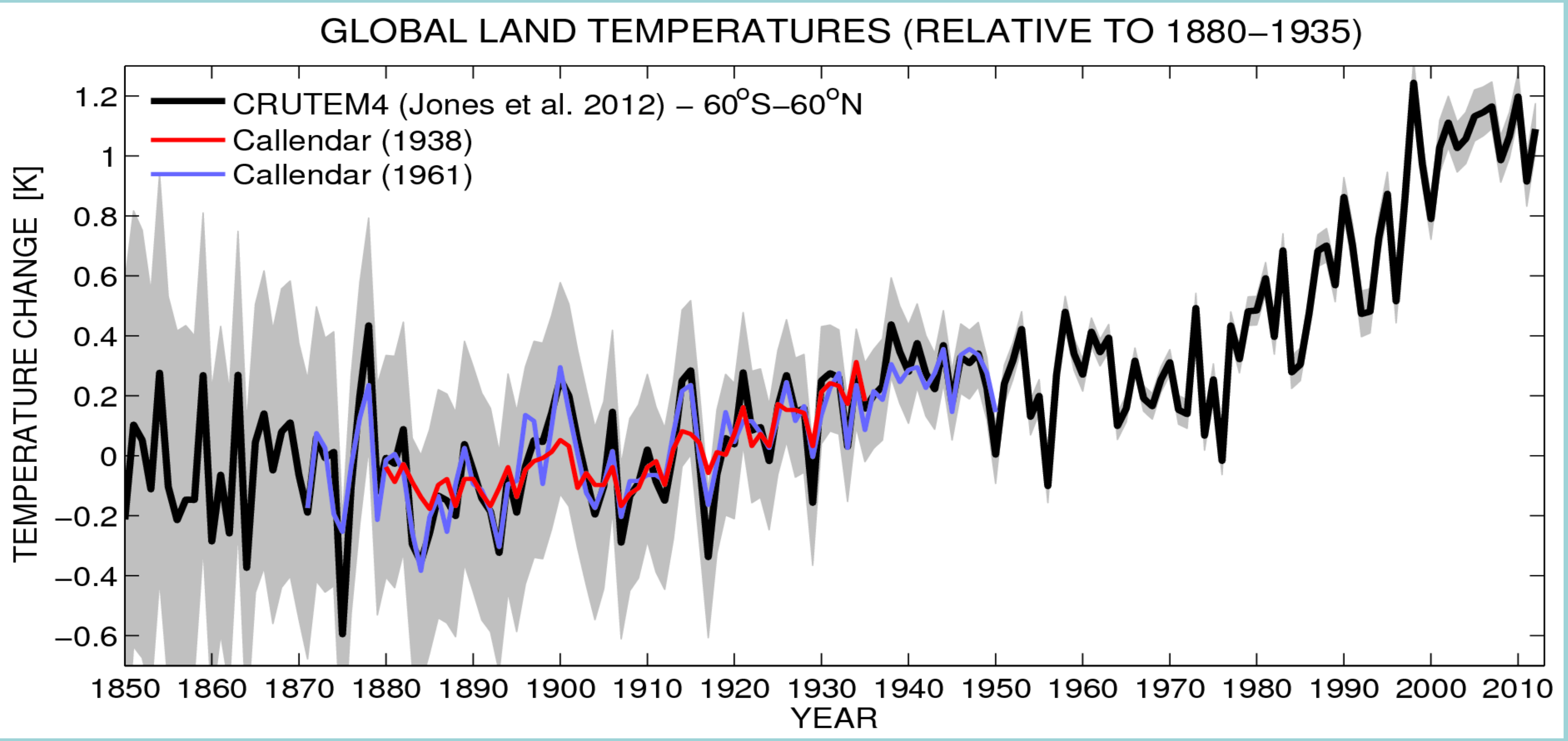
Wladimir Köppen (1846-1940) Picture: 1921



Guy Stewart Callendar (1898-1964). Picture: 1934

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

Includes the error estimate ranges for CRUTEM4 developed by Morice et al (2012)
Callendar had no Antarctic data, so also didn't use the few Arctic stations he had access to



Callendar used about 150 stations in 1938 and about 400 in 1961. All calculations by hand.

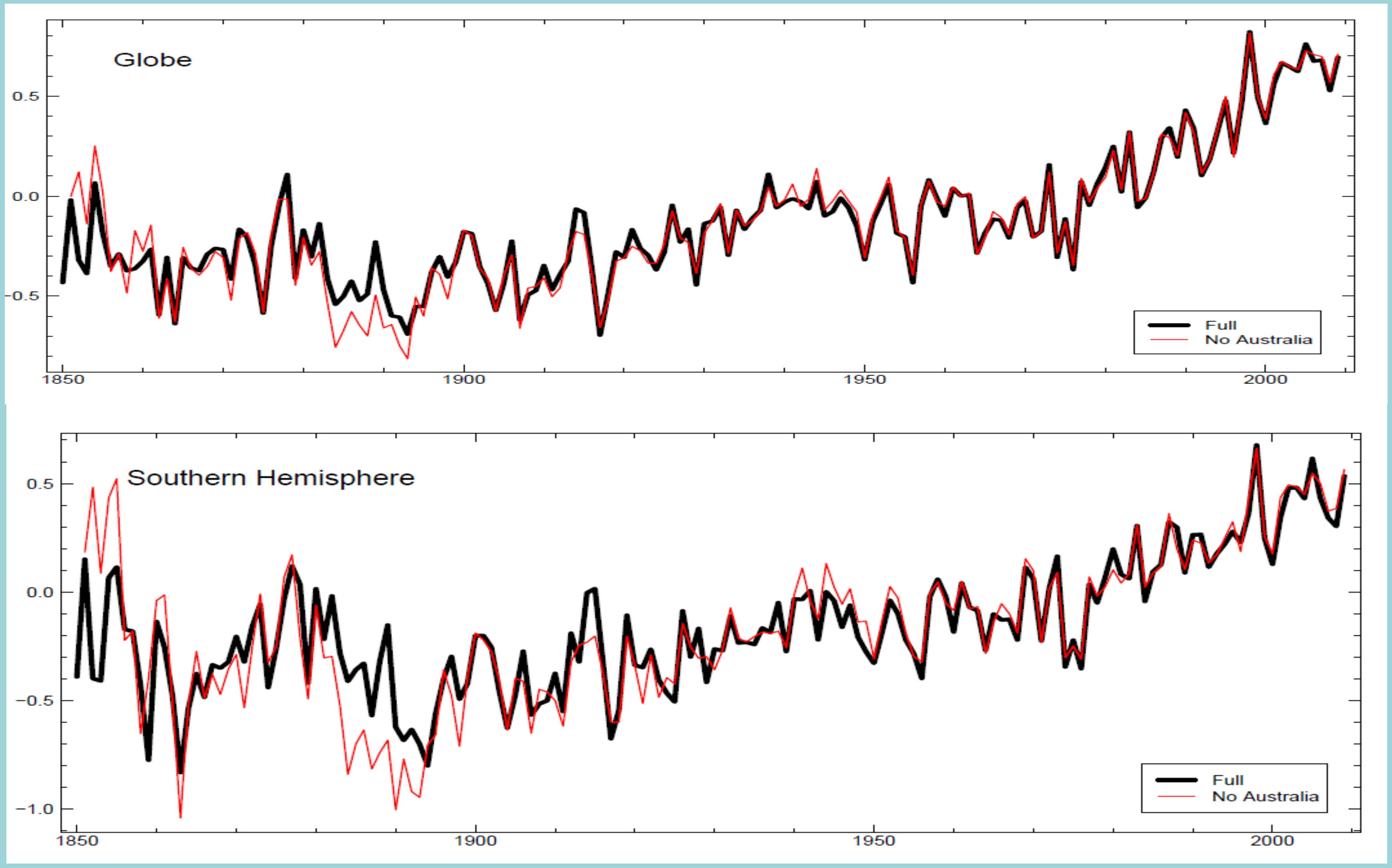
Hawkins, E. and Jones, P.D., 2013: On increasing global temperatures: 75 years after Callendar. Q. J. Royal Meteorol. Soc., 139, DOI:10.1002/qj.2178.

Coverage Change (Spatial Degrees of Freedom)

- Using Callendar's work with many fewer stations, it should be obvious that there must be a limiting set of well-spaced sites that will achieve the same Global land temperature average as the series based on thousands of sites
- This number is the effective number of spatial degrees of freedom
- For temperature at monthly timescales the number is about 100. It will be more on daily timescales, less on decadal and century. Much greater for another variable like precipitation
- The fact that the number is relatively small is the reason that temperatures prior to the instrumental series can be reconstructed from proxy data
- **For a gridded product, then use all you can access**
- Modern Reanalyses (like ERA5 and JRA55) can now be used to subsample and calculate the effective number

Jones, P.D., Osborn, T.J. and Briffa, K.R., 1997: Estimating sampling errors in large scale temperature averages. *J. Climate* **10**, 2548-2568.

Removal of Australia



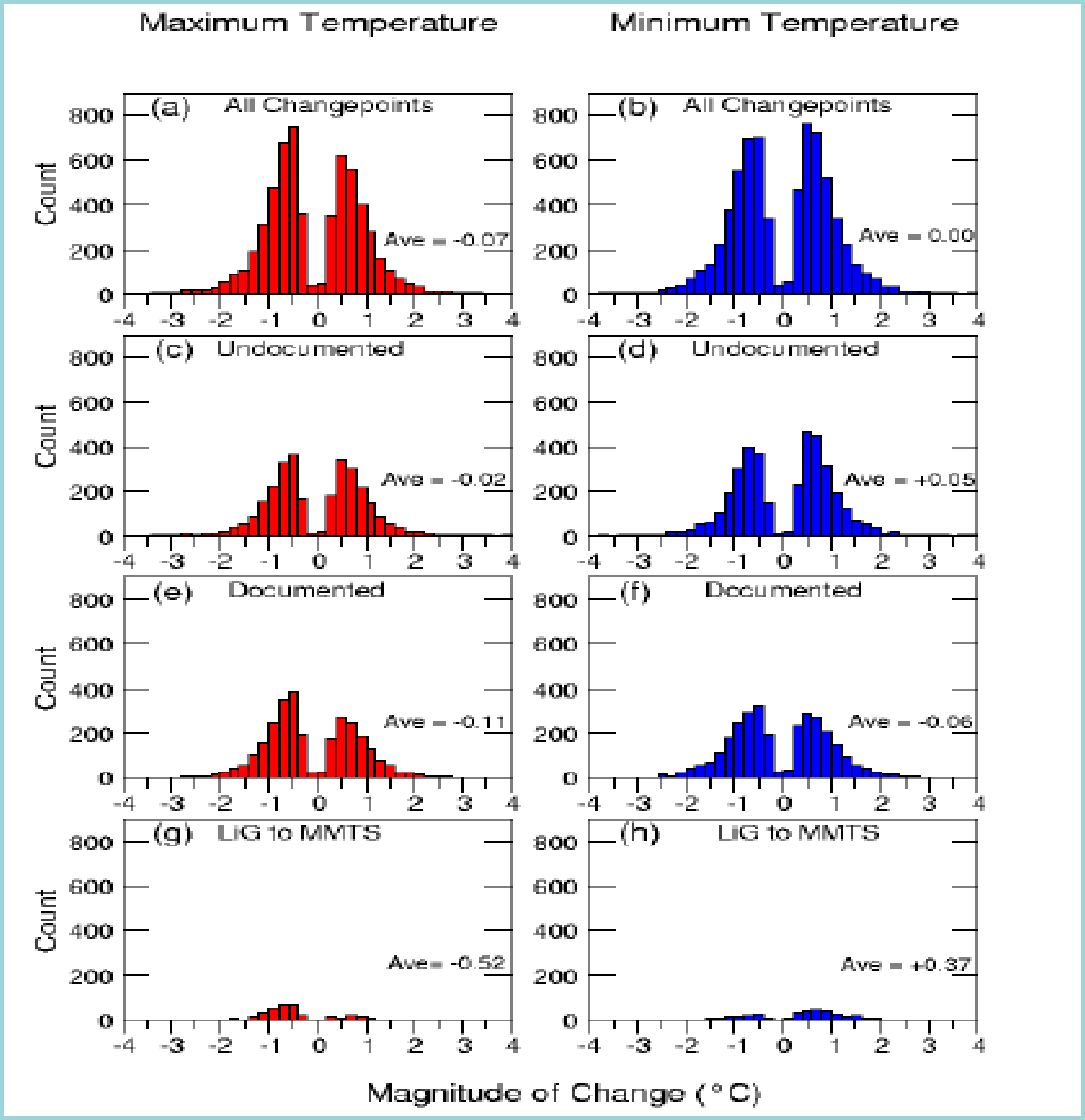
This is by far the most dramatic removal as it is over half of the stations in the SH
Smaller effects in the NH of the removal of stations in the US, fUSSR, China and Canada

Homogeneity (long-term consistency)

- Maybe not this word, but consistency of measurements was realized in the 17th and 18th centuries (and allowed for in the 19th century). How do we know?
- Defined by Conrad and Pollack (in the 1950s) to indicate that the series should only include the vagaries of the weather and climate, **so not the effects of our instruments, their exposure, our observation times, and changes around the environment of the site**
- Various homogeneity algorithms have been recently developed (assessed by Venema et al 2011, 2020)
- Essential to know the site's history (when changes happened, how the environment around the site has changed etc.). **Very unlikely to be the case – see next slide**

Venema, V.K. et al., 2011: Benchmarking homogenization algorithms for monthly data. *Climates of the Past*, **8**: 89-115
Venema, V.K., Trewin, B. and Wang, X.L., 2020: Guidelines on Homogenization. WMO No. 1245, 64pp (on WMO website https://library.wmo.int/index.php?lvl=coll_see&id=13#.Xz5Nu-hKg2w)

Bimodal distribution of adjustments using the USHCN (~1200 sites over the contiguous US, records back to 1890s)



Menne MJ et al., 2009: **The U.S. Historical Climatology Network Monthly** Temperature Data, Version 2, *BAMS*, **90**, 993-1007

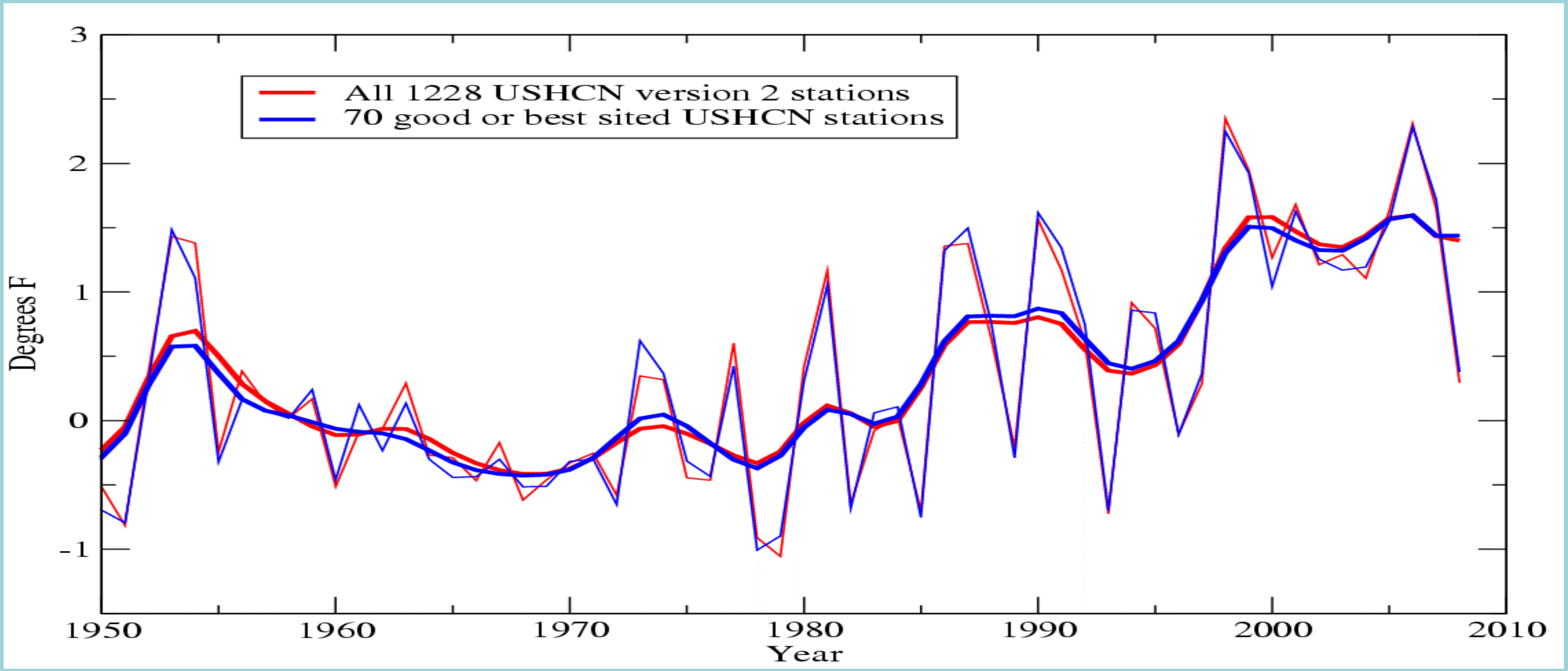
More discussion of the differences in the CRU approach with the US PHA approach and Berkeley Earth in Menne et al. (2018)

Menne, M. J. et al., 2018: The Global Historical Climatology Network Monthly Temperature Dataset, Version 4, J. Climate, 31, 9835–9854, 2018

PHA – Pairwise Homogeneity Algorithm

CRU approach – access NMS adjusted data

USHCN – all and 70 best (latter partly determined by surface stations.org from pictures of the sites)



The 70 obviously omit large parts of the contiguous US.

Shows you can't assess homogeneity by looking at pictures of sites. It is necessary to look at the data, compare with neighbours and if necessary make adjustments

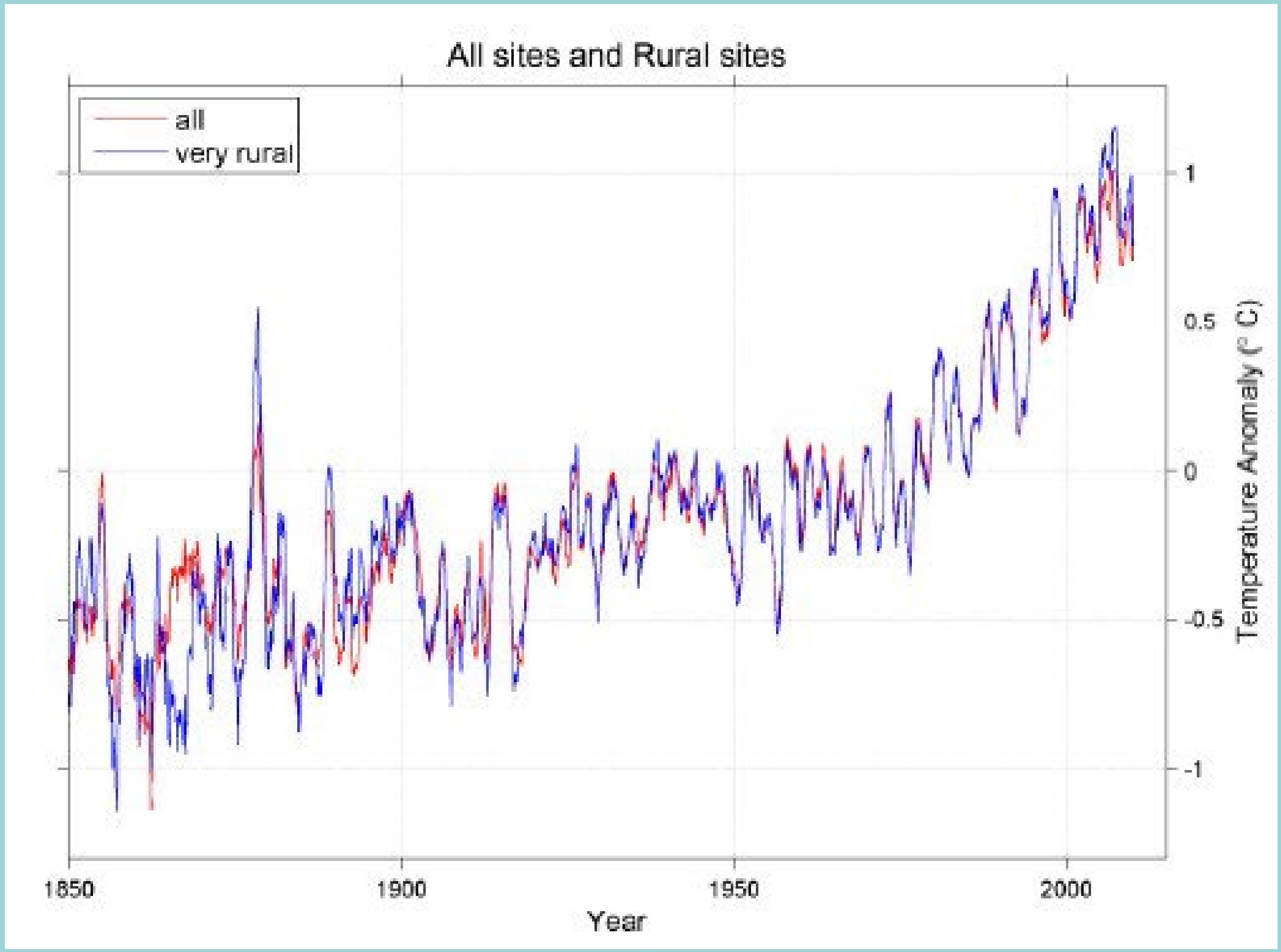
Also shows homogeneity algorithms work

Urbanization Influences

- Homogeneity testing may not remove all urban affected sites if all neighbouring sites are similarly affected by urban growth
- CRU approach has been to develop a dataset of rural-only stations.
- Grid the rural-only stations and then compare with the grid with all the stations
- There are numerous studies which look at urban effects, but these are mostly for individual sites and look principally at maximum (or extreme) effects on daily values
- **The issue we are assessing is the effect on monthly and annual average temperatures**

Parker, D. E. (2010), Urban heat island effects on estimates of observed climate change, *Wiley Interdisciplinary Reviews: Climate Change*, **1**, 123–133, doi:10.1002/wcc.21.

Large-scale urbanization influence is negligible



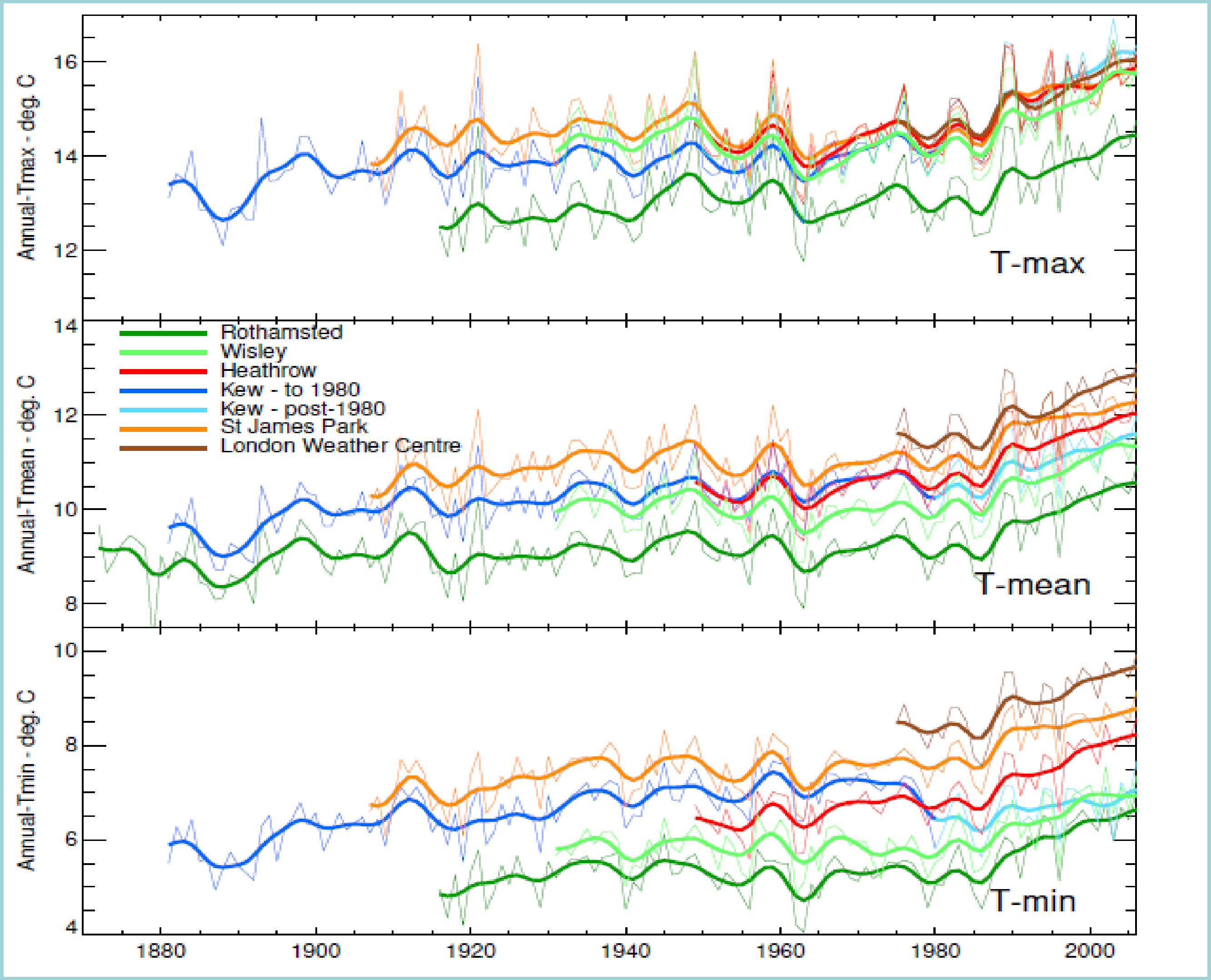
In this recent analysis by the Berkeley Earth team, the very rural sites warm slightly more than the urban sites

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

Need to look at global land areas: not at individual sites and not on extreme days

London



Jones, P.D. and Lister, D.H., 2009: The Urban Heat Island in Central London and urban-related warming trends in Central London since 1900. *Weather* **64**, 323-327.

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**

SJP warmer by about 1.1 deg C than if London wasn't there. **Anomalies from 1961-90 look like other sites**

A little more on London and its UHI

- The previous slide showed that SJP had a UHI of about 1.1 deg C
- Much earlier work by Luke Howard in the 1830s in his Climate of London (reprinted in 2007) calculated the UHI for central London to be 2 deg F (1.1 deg C)

- Luke Howard also developed the cloud naming schemes we use today

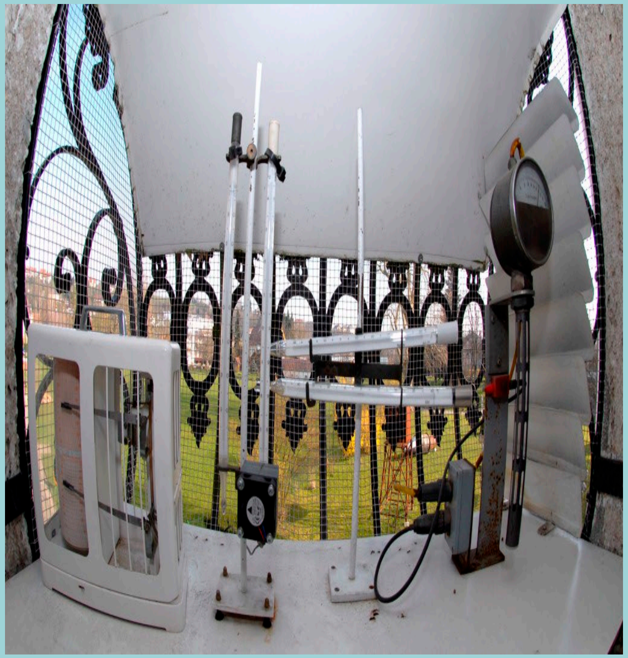
- Howard L. 2007: The Climate of London. IAUC edition available at www.lulu.com in two volumes.
- Mills, G. 2008: Luke Howard and The Climate of London. *Weather* **63**, 153-157.

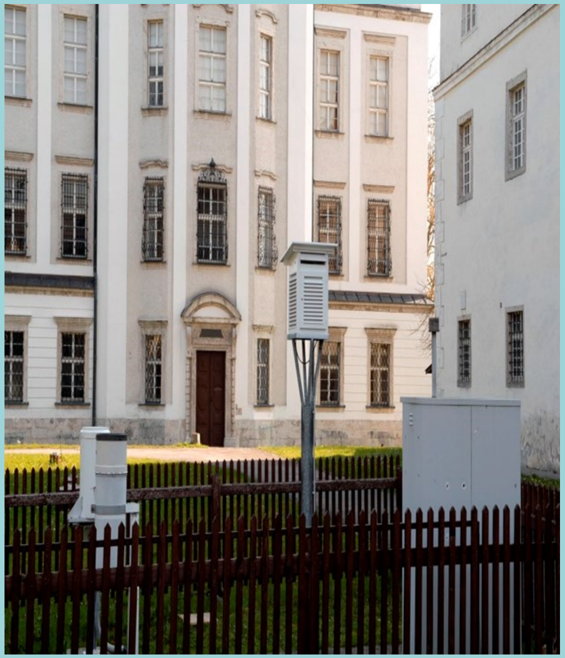
IAUC – International Association of Urban Climates

Early exposure issues

- Europe affected, before the development of Stevenson screens (~1850 in the UK)
- Solution has come about from modern parallel measurements (in Austria and Spain, with the old screens)
- Effect is annually ~0.4°C, with most series too warm by up to 0.7°C in June (before screens)
- **Issue important as it is the summers that calibrate many natural and documentary proxies**

Exposure Issues pre-screens Kremsmünster - Austria

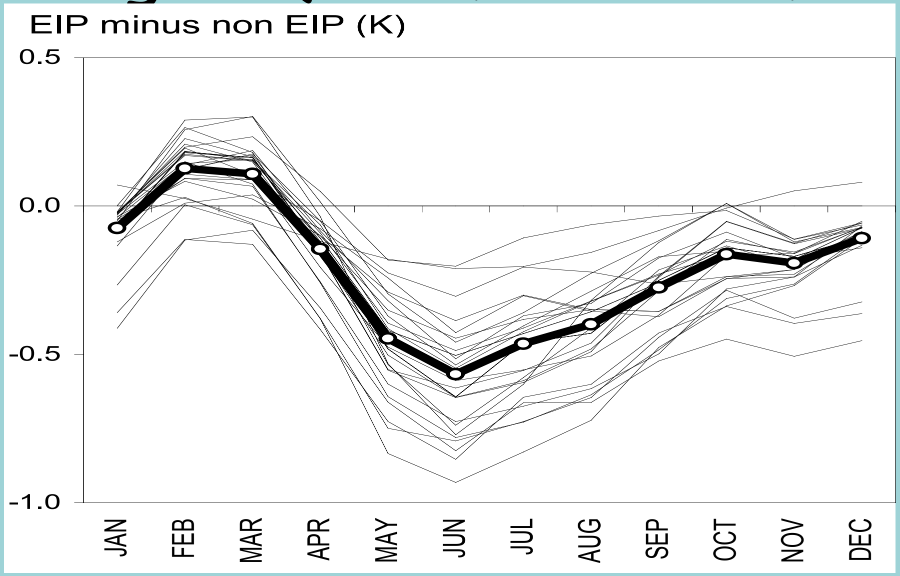




Böhm, R., Jones, P.D., Hiebl, J., Frank, D., Brunetti, M. and Maugeri, M:., 2010: The early instrumental warm-bias: a solution for long Central European temperature series, 1760-2007. *Climatic Change* **101**, 41-67.

Also papers from Spain and Australia

Adjustment across all 32 sites in the Greater Alpine Region (GAR, 43-49N, 4-19E)



EIP is pre-screen period

Early Instrumental Period (EIP) is pre screen. EIP temperatures in summer were warmer than they should be. Winters were OK

If uncorrected there will differences between the seasons. Needs correction for calibrating proxies as most respond to summer temperatures

More on Exposure Issues

- Initial measurements from the 17th century were from a variety of exposures
- The effect of the sun on early thermometers was well known. A 1725 recommendation (James Jurin) was for the thermometer to be in an unheated room. This didn't work, as winter temperatures at **Uppsala** from 1725 to 1739 were far too warm. The Austrian monastery window worked well
- Comparisons were likely made at the time, but few of these have survived.
- **Spanish approaches have developed adjustments by rebuilding old screens and making modern comparisons**
- **Longest overlap is at Adelaide – between Glaisher stand and Stevenson screen (1887-1947). A paper on this has just appeared in IJC (Ashcroft et al., 2021)**
- **Glaisher was a competitor to Stevenson, but comparisons at the time (1860s/1870s) favoured Stevenson**

Marine Temperatures

- Ships take two temperature measurements – SSTs and MATs
- Both require some form of adjustment
- Until recently more was known about SST (so their anomalies used for the oceans). More is now understood about MATs. A new project (GLOSAT, NERC funded) will soon combine MAT and SAT
- Both MAT and SST vary much less from day to day than SAT, and particularly for SST. So, few observations in a month provide good estimates. GLOSAT will allow daily MAT values to be used

- Coverage issues with SSTs are helped by having satellite estimates since 1979, so base periods to derive anomalies can be calculated

Marine Data Coverage

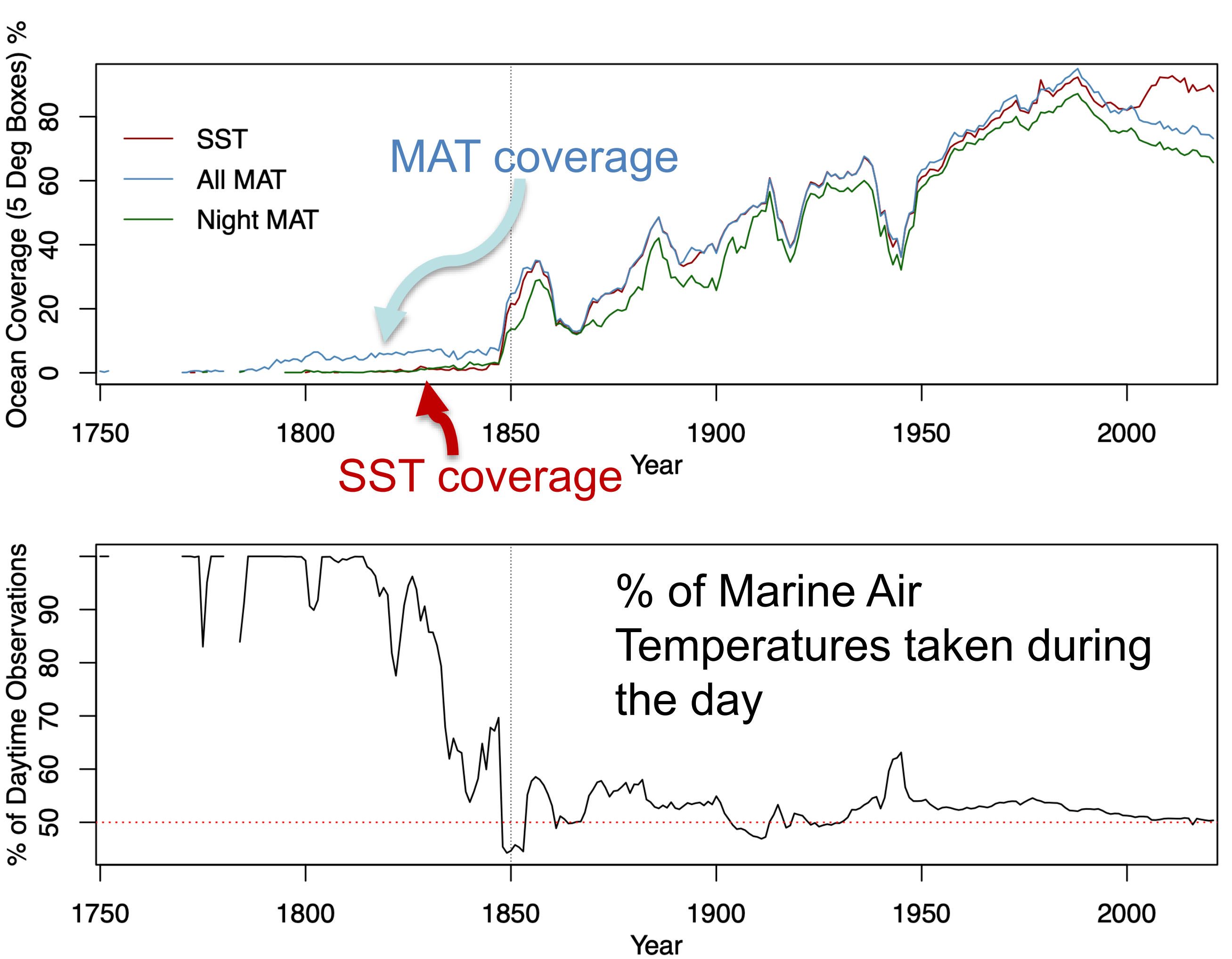


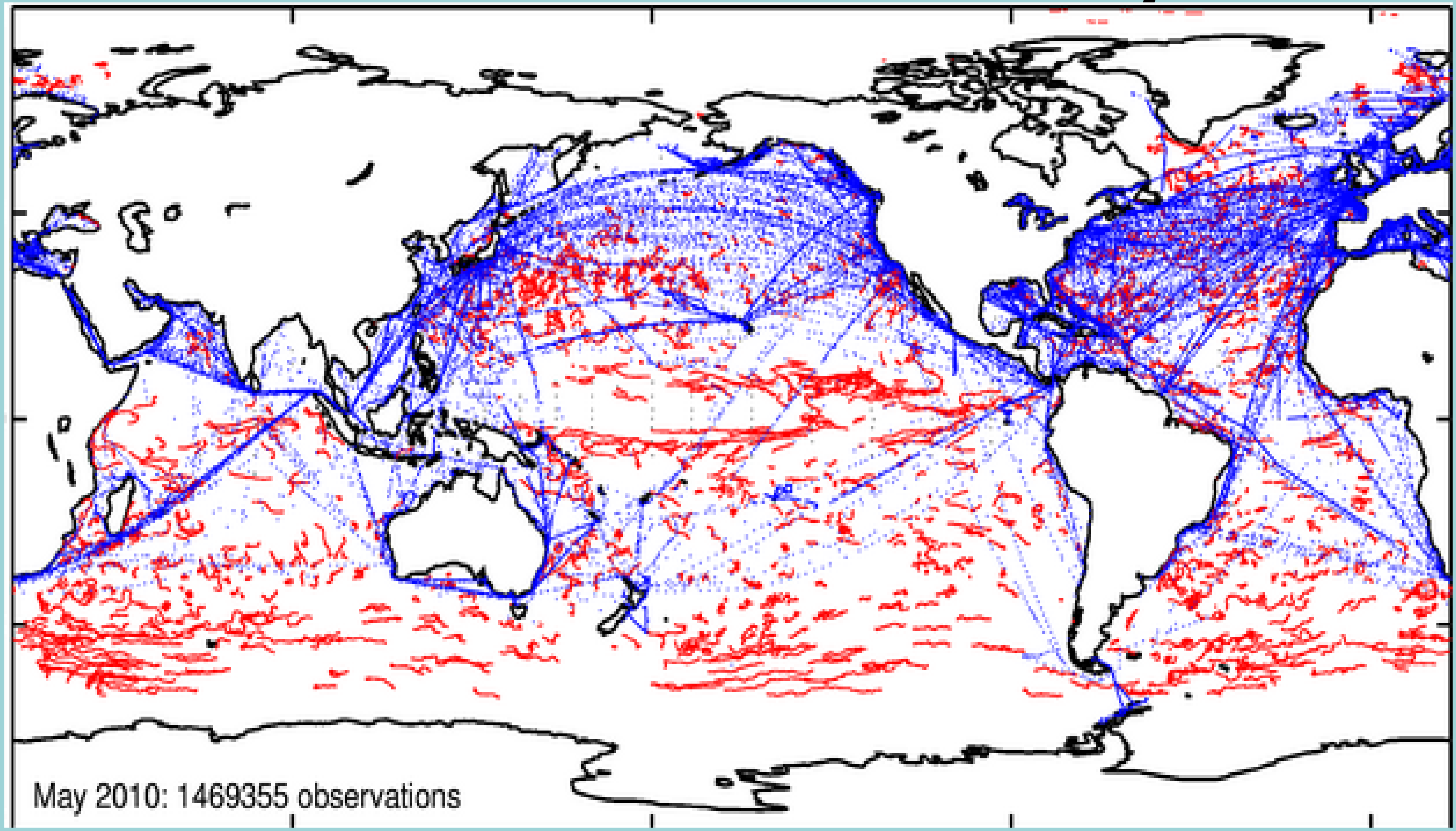
Figure from Tom Cropper, National Oceanography Centre

Before 1850, most observations are MAT not SST, and most MAT are daytime

GloSAT project (led by Liz Kent at National Oceanography Centre) is grasping this nettle so we can extend back pre-1850

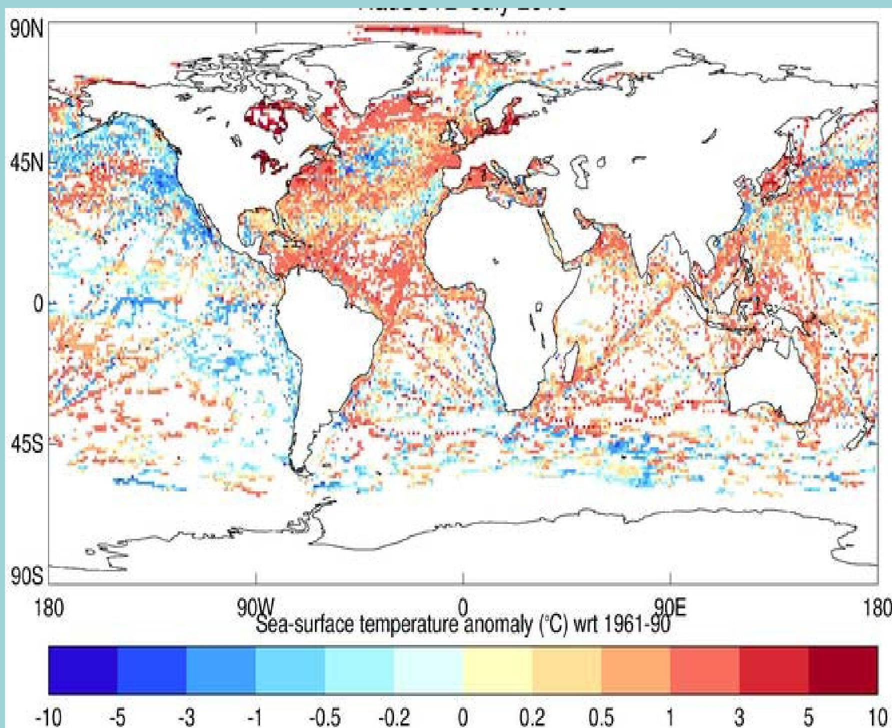
Explains why the global temperature series go back now to 1850

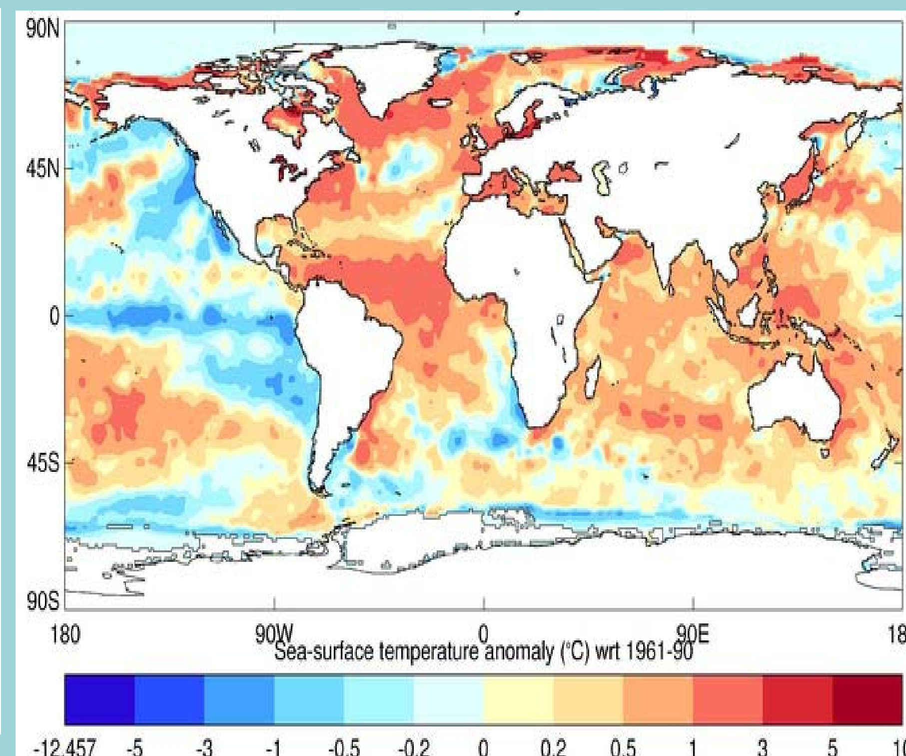
SST Observations – May 2010



Blue – ships; Red – drifting buoys; Grey – fixed buoys

SST Interpolation – May 2010





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.

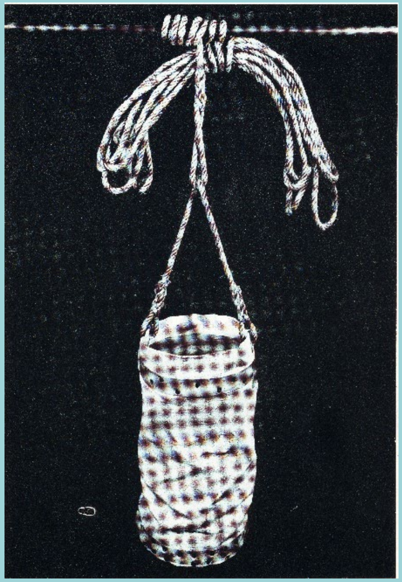
Base period is 1961-90. This is a ship-based estimate. If more recent buoy/drifter observations are slightly cooler then this needs to be allowed for. Need to have enough observational overlap to do this

The evolving *in situ* SST observing system: methods

From: Met Office

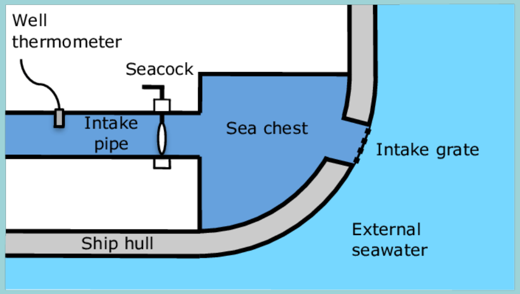
Wooden Bucket

Early obs up to 1920s



From: "Marine Observers Handbook" HMSO, 1969 Ed.

Canvas Bucket
From 1880s to 1940s



From: Matthews, 2013: Comparing historical and modern methods of SST measurement – Part 1, Ocean Sci., doi: 10.5194/os-9-683-2013, 2013.

Engine Intake
From 1930s to now

From: "http://www.metoceanservices.com

Drifter
From ~1990

Method used not recorded until the early 1970s

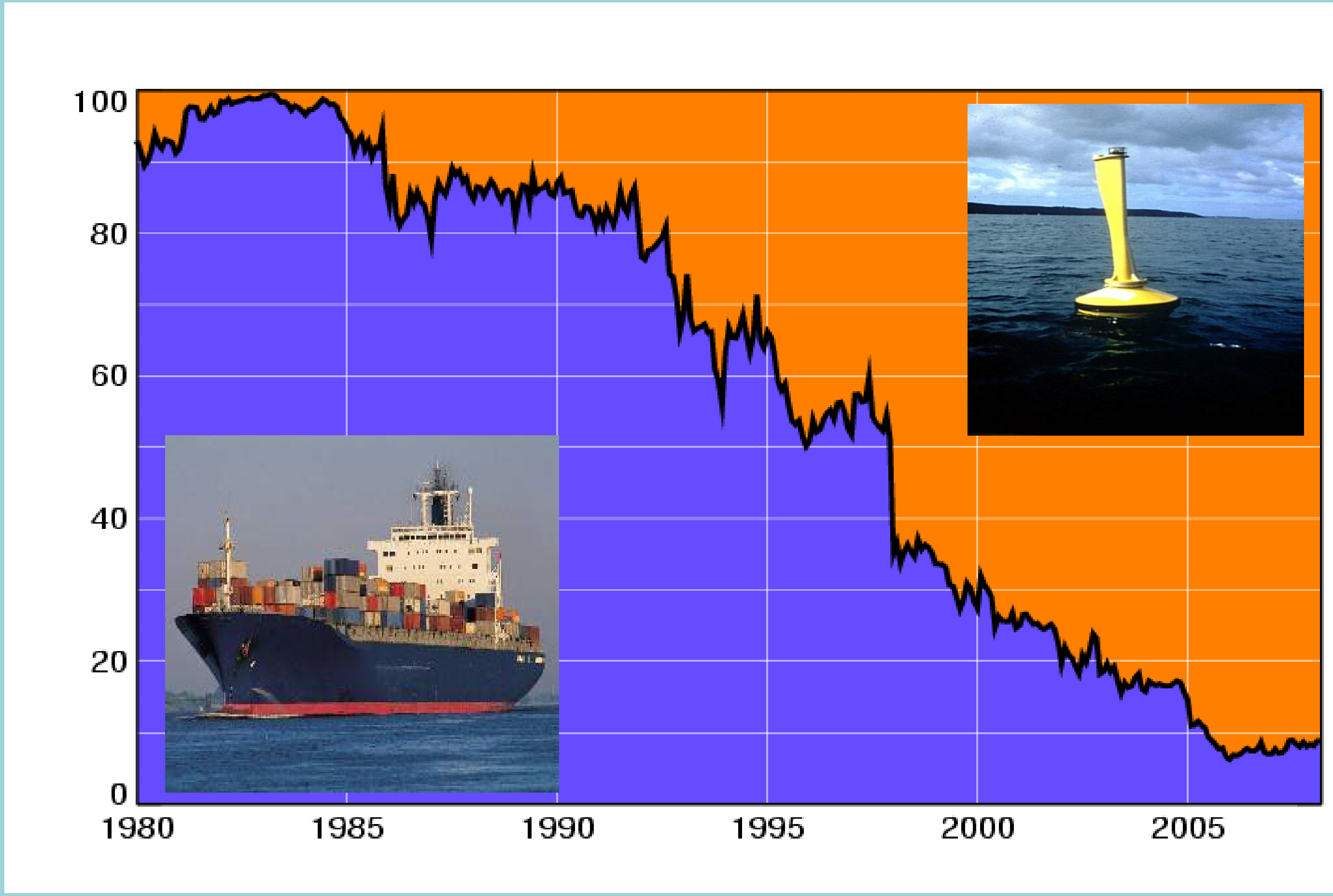
Huge change in marine observing network in the past 25 years

Percentage of observations coming from

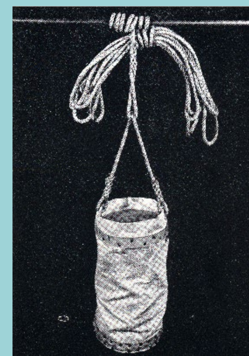
DRIFTERS

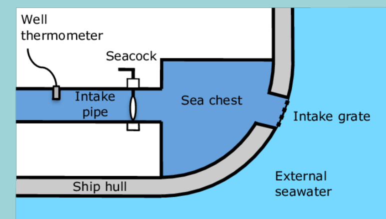
and

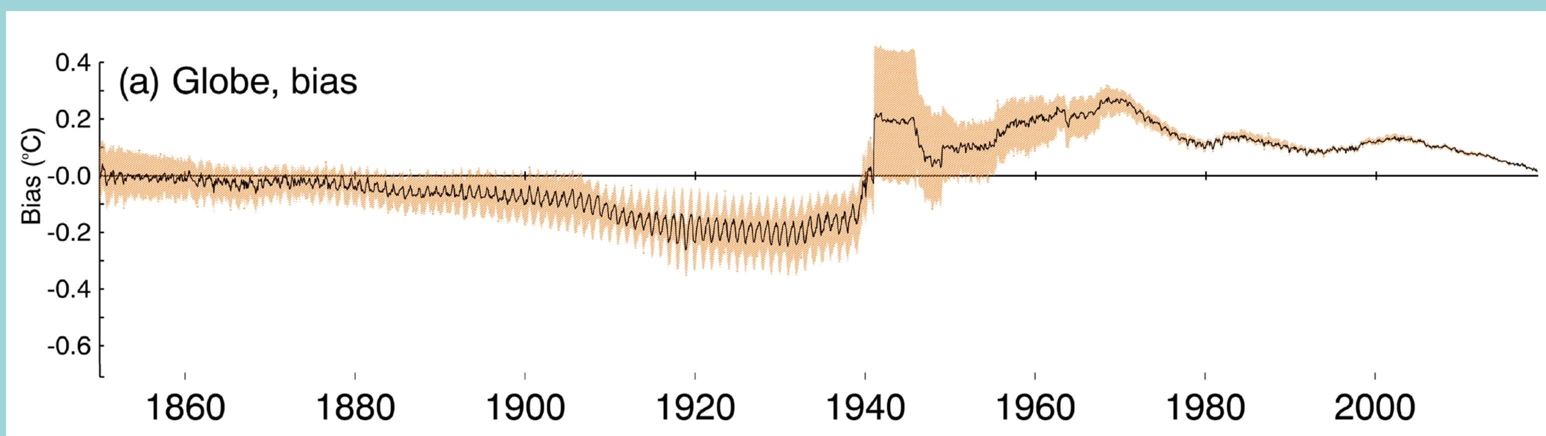
SHIPS



HadSST4 bias adjustments







JGR Atmospheres

RESEARCH ARTICLE

10.1029/2018JD029867

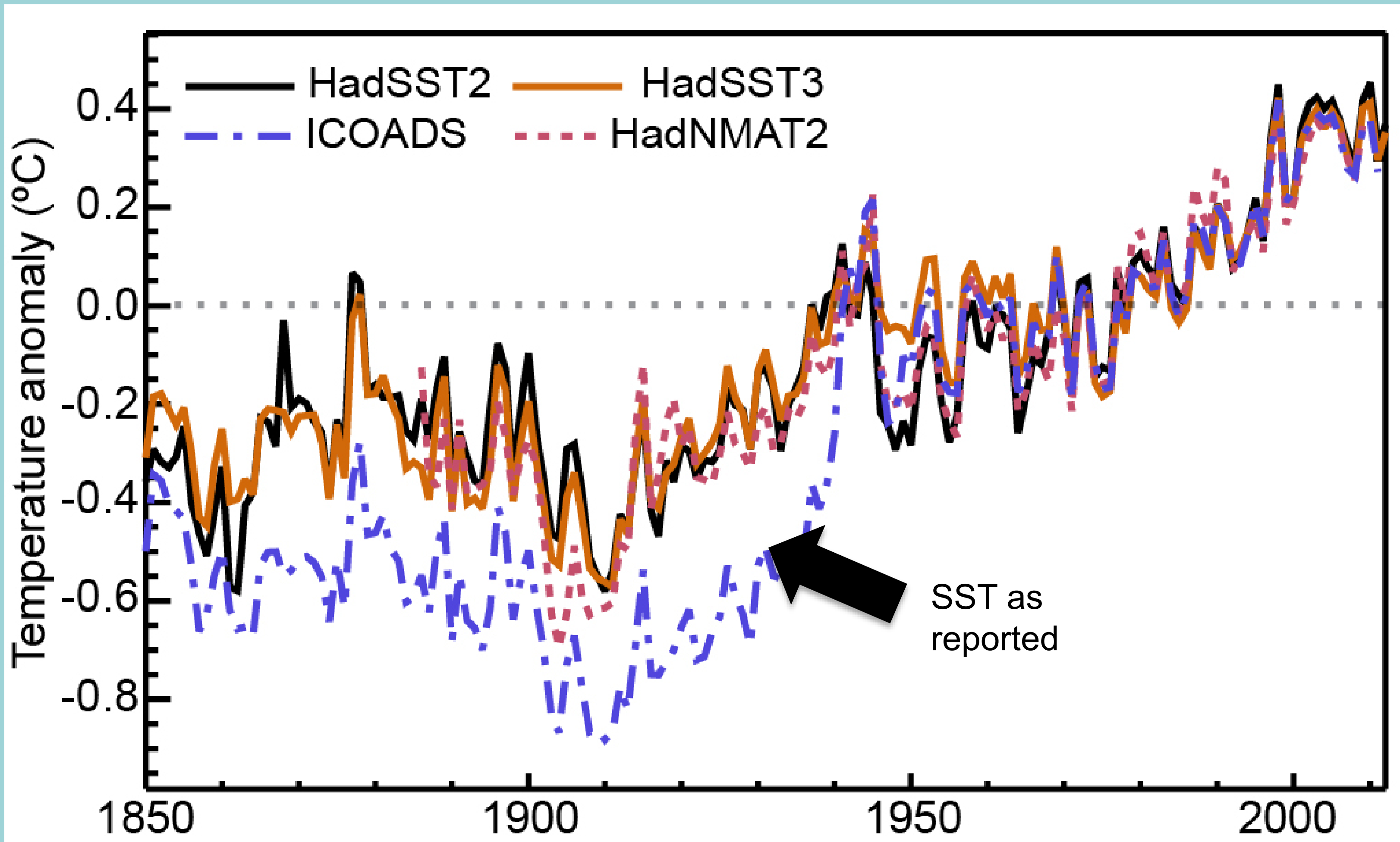
Key Points:
- We describe the construction of HadSST.4.0.0.0, a climate data set of sea surface temperature change from 1850 to 2018
- A range of bias adjustments was generated to create an ensemble of

An Ensemble Data Set of Sea Surface Temperature Change From 1850: The Met Office Hadley Centre HadSST.4.0.0.0 Data Set

J. J. Kennedy[1], N. A. Rayner[1], C. P. Atkinson[1], and R. E. Killick[1]

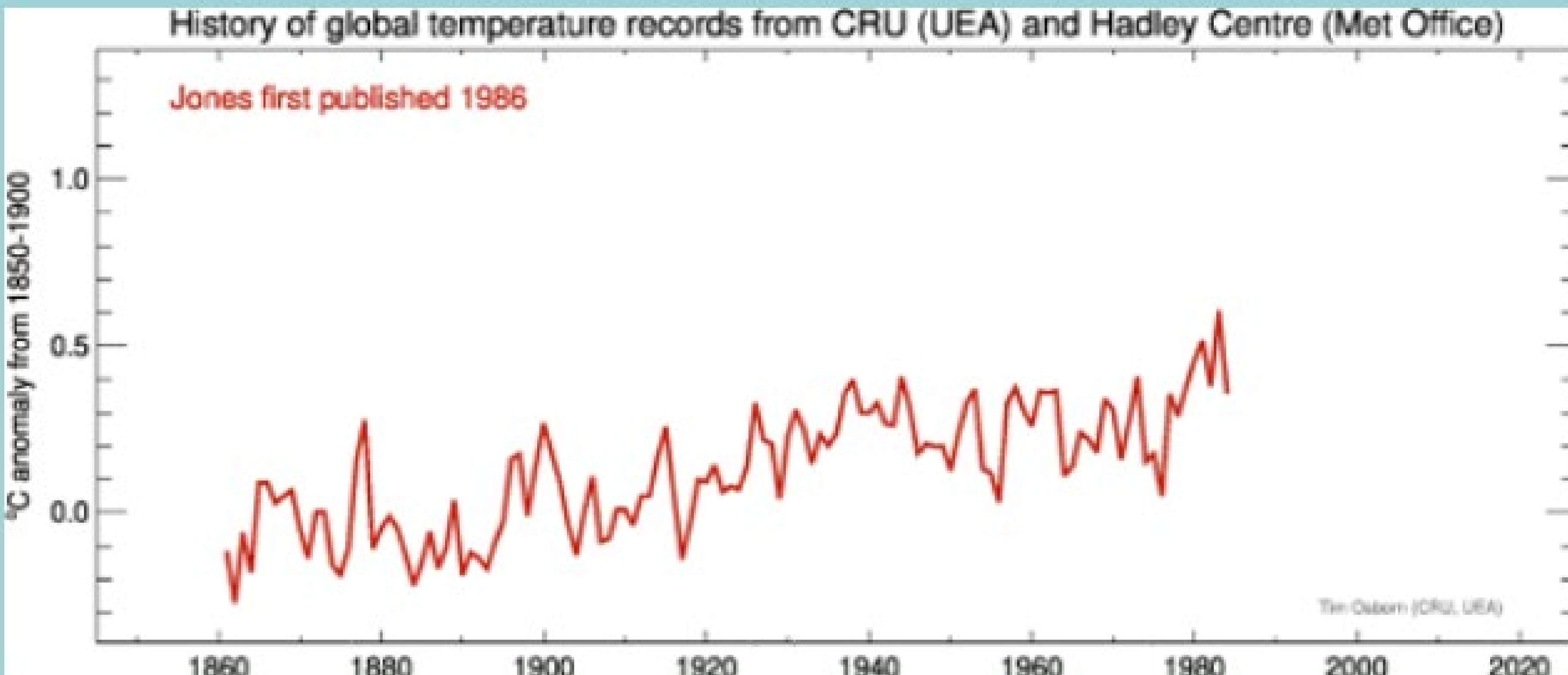
[1]Met Office Hadley Centre, Exeter, UK

Adjustments to SSTs

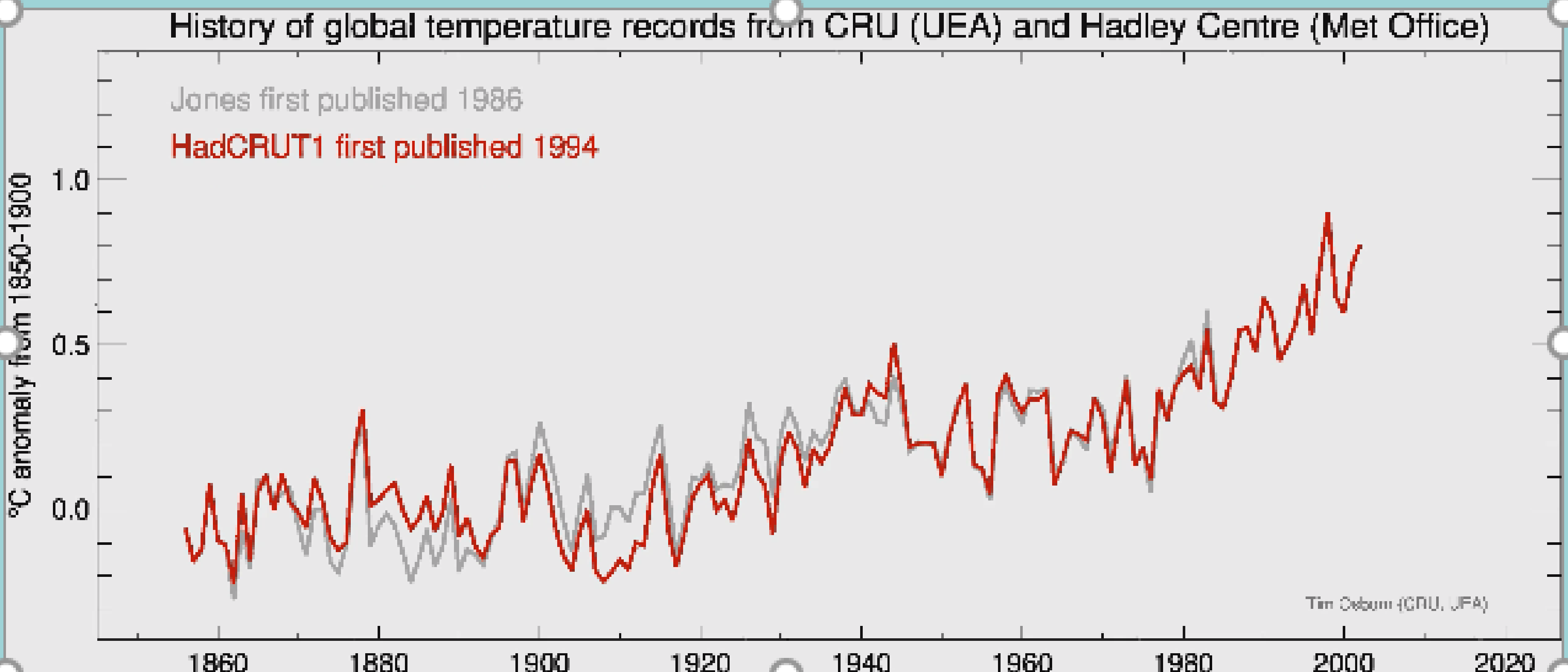


Slide from Liz Kent (NOC, Southampton)

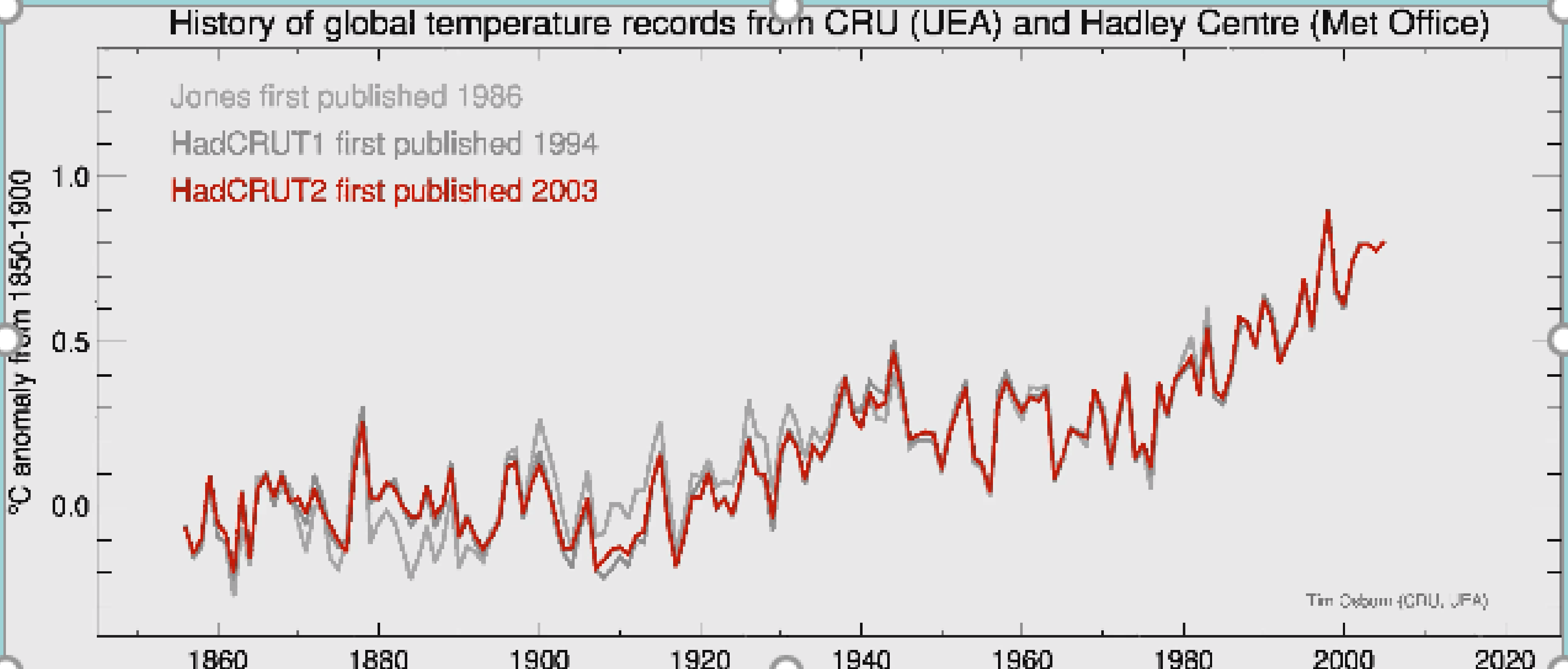
HadCRUT versions, all wrt 1961-1990



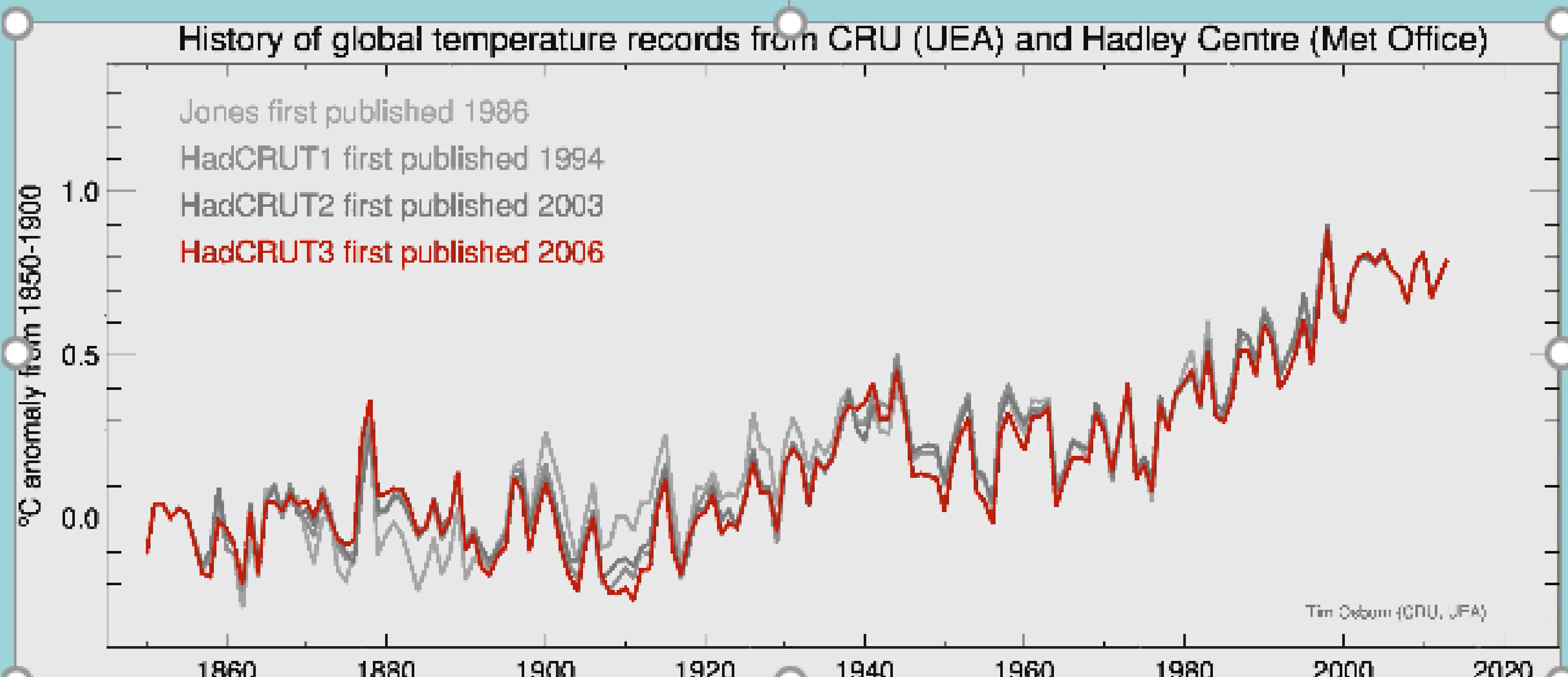
HadCRUT versions, all wrt 1961-1990



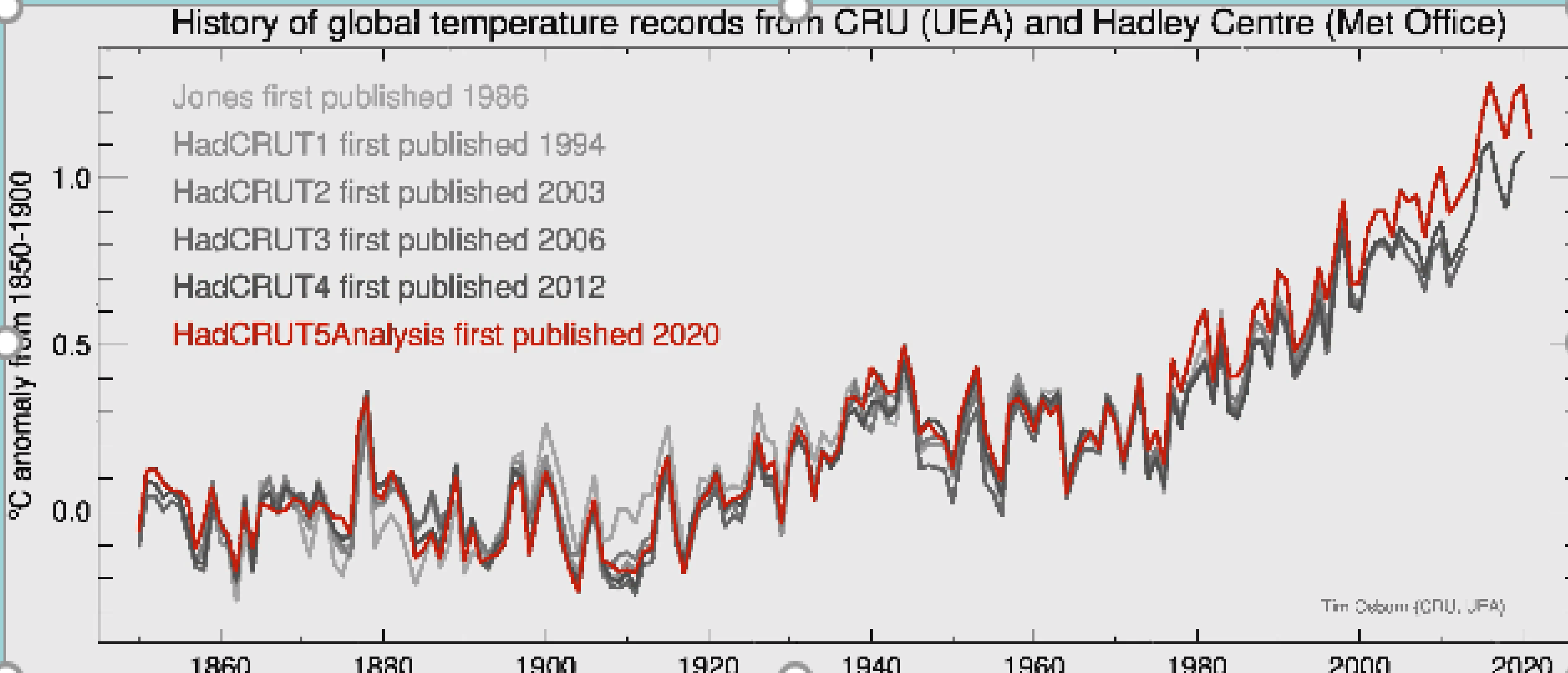
HadCRUT versions, all wrt 1961-1990



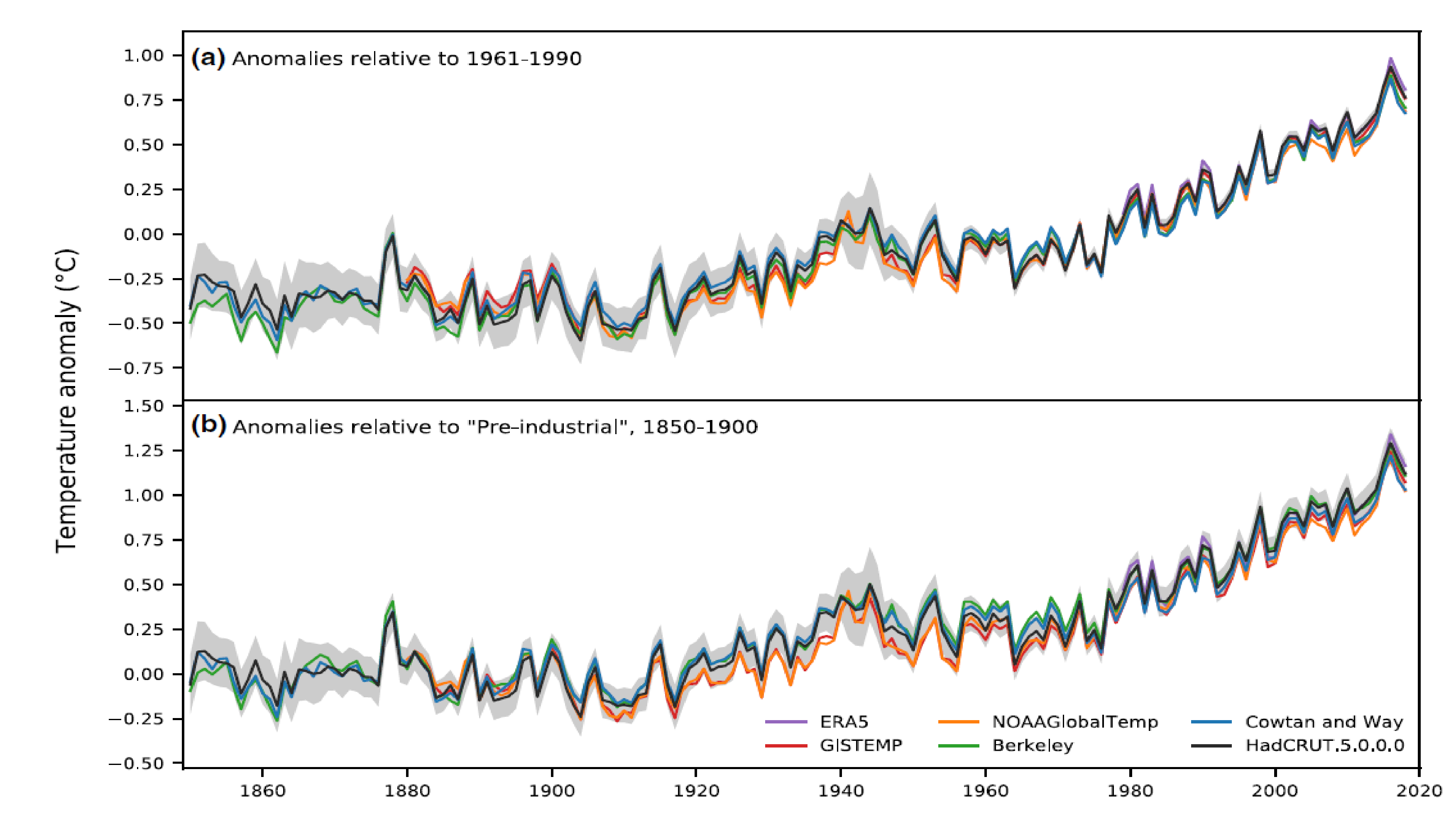
HadCRUT versions, all wrt 1961-1990



HadCRUT versions, all wrt 1961-1990



Comparison of HadCRUT5 with other series



Very small differences at these scales. American groups use spatial infilling. HadCRU4/5 do to a small extent, but the best is to find additional data. More ships being found, but here we'll discuss finding more land data

Recent Work (land stations)

- Adding extra station data
- Making use of short segments – stations that don't have the 1961-90 base period
- Extensions back to 1781. Mostly European, but MAT data from sailing ships incorporated
- More ship logbook data being found in a variety of locations (e.g. Finnish clipper ships – in Aland Island archive, Whaling ship archives). Important to understand that different shipping fleets recorded in different ways.

Extra Station Data

- C3Surf – (KNMI led project, so mostly European)

- Earlier European data on some NMS sites. Some are good, but generally only go back to the foundation of the Met Service (e.g. Switzerland in 1864, Sweden ~1861). Need to get NMSs everywhere (especially in developing countries), to realize that data exists before the country became independent or its NMS existed.

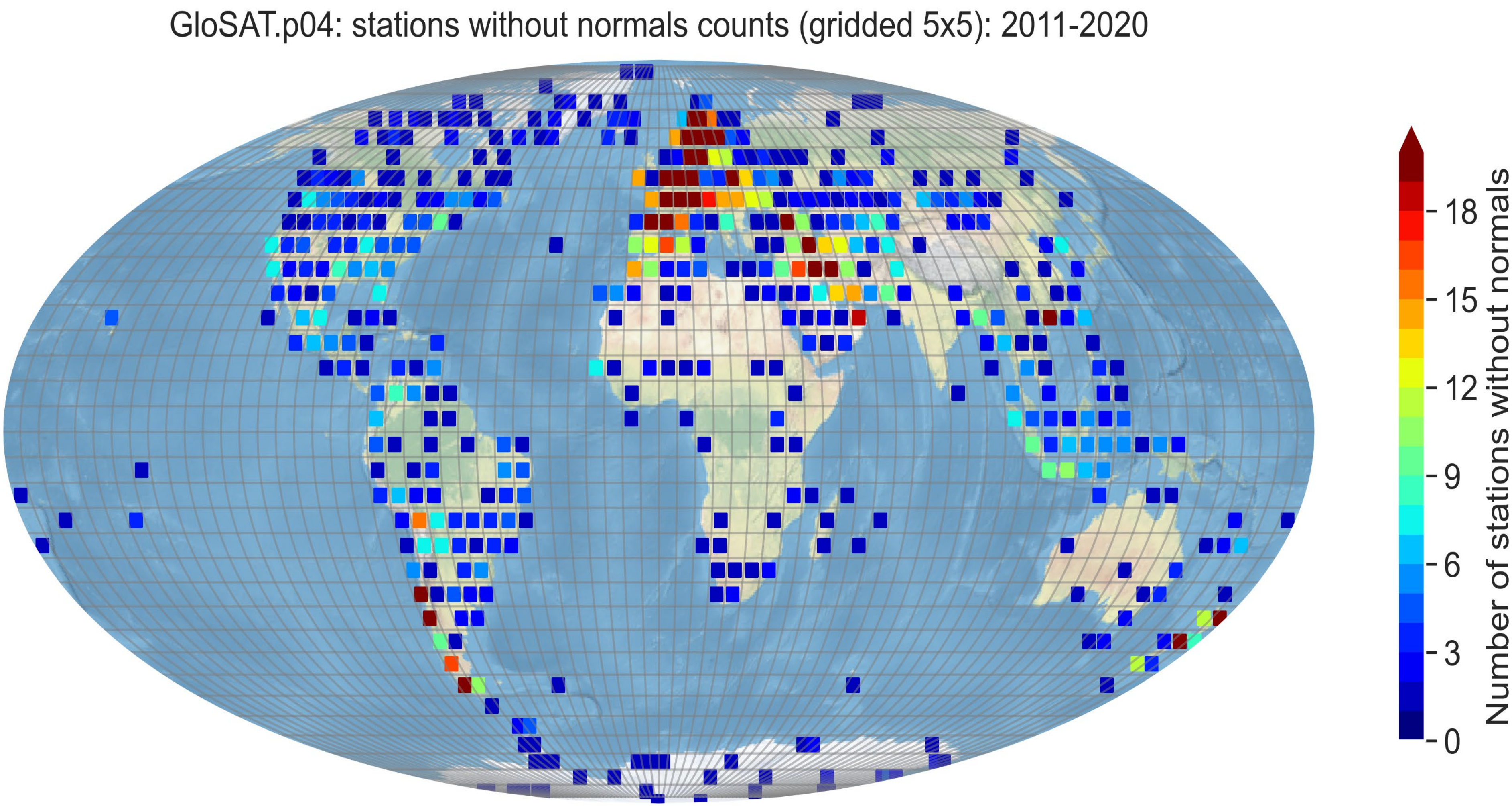
- Getting more station data – that's not on the CLIMAT/WIGOS network

- NMS websites (Russia- when the site is up, Israel, Brazil, Indonesia)

- Requesting data from authors of Scientific Papers (added data from Brazil an also Svalbard/Frans Joseph Land)

- Contacting people I know (Saudi Arabia, Argentina, China, New Zealand). China has a network of about 3000 stations, but they are designated primary, secondary and tertiary

Additional stations reporting in 2011-2020 but no 1961-90 normals



Also short stations in the 19th century – not linked to a modern record
Hardly any of these efforts improve Africa

Estimating Normals

- Using a technique called LEK (Local Expectation Kriging) –uses neighbours
- Estimates the absolute temperatures for each month in each year of 1961-1990. Uses elevation as well.
- Can show it works, by estimations where we know the result
- Problems occur near coastlines and islands, and where sea-ice occurs, but the errors are relatively small
- The Svalbard/Frans Joseph sites and Antarctic Peninsula sites used for testing
- Despite this, some islands are just too far from anywhere to do this

Problems

- Adding short segments (also some long stations without 1961-90 normal) makes little difference to regional and hemispheric averages
- Adding short series with just the 2010s decade tends to slightly lower warming (but only by a few hundredths of a degree)
- Similar adding short series before 1900 tends to very slightly lower temperatures
- Both these are compared to what we had before the data were added. Hardly any long-term change in warming. More grid-boxes have values and many grid boxes have more stations per grid box
- Early data (pre-1900) may have issues that the monthly means are calculated differently from that used by most stations during 1961-90 (i.e. fixed hours and (Tx+Tn)/2)
- Coverage issues can be looked at with complete ERA5 and JRA55

What are other groups doing?

- NOAA, GISS get their data from NOAA – they produce homogeneous series (and as received series). Homogenization via the PHA algorithm (Pairwise Homogeneity)
- Berkeley Earth get all these series from NOAA and also from CRU. They may add a few series, but they have few resources (at least since the Koch Foundation withdrew their financial resources when the funders didn't like the answer)
- CMA (in Beijing) now have a dataset. They have more stations in China, but mostly use NOAA and CRU sources (and all the NMS links I have given them). CMA has better contacts with a few countries in eastern Asia
- Others take the NOAA and CRU data and attempt mathematical infilling using various approaches

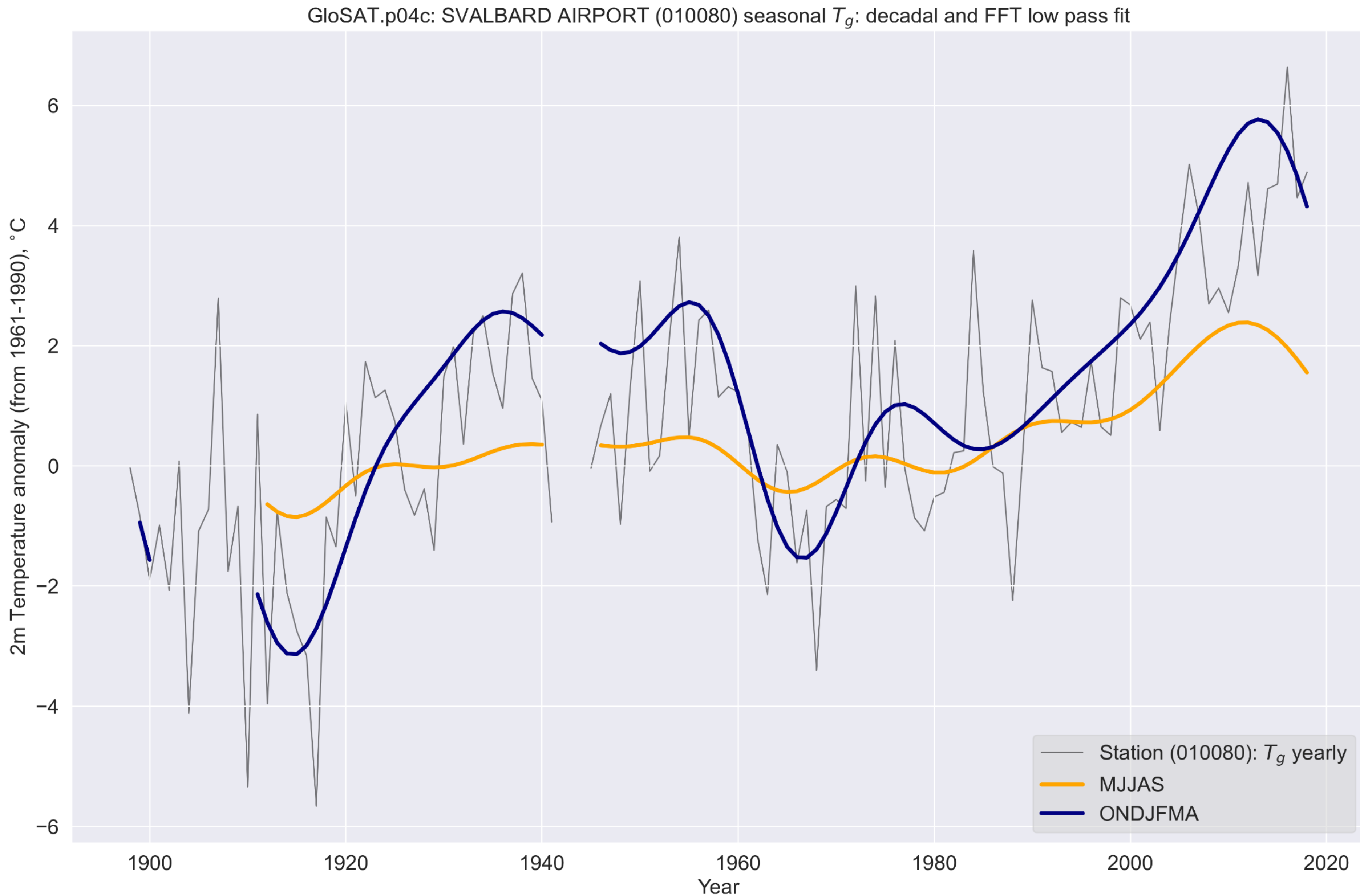
Asking scientists for their data

- Klingbjer, P., and Moberg, A. (2003). A composite monthly temperature record from Tornedalen in northern Sweden, 1802-2002. *Int. J. Climatol.* **23**, 1465–1494. doi: 10.1002/joc.946

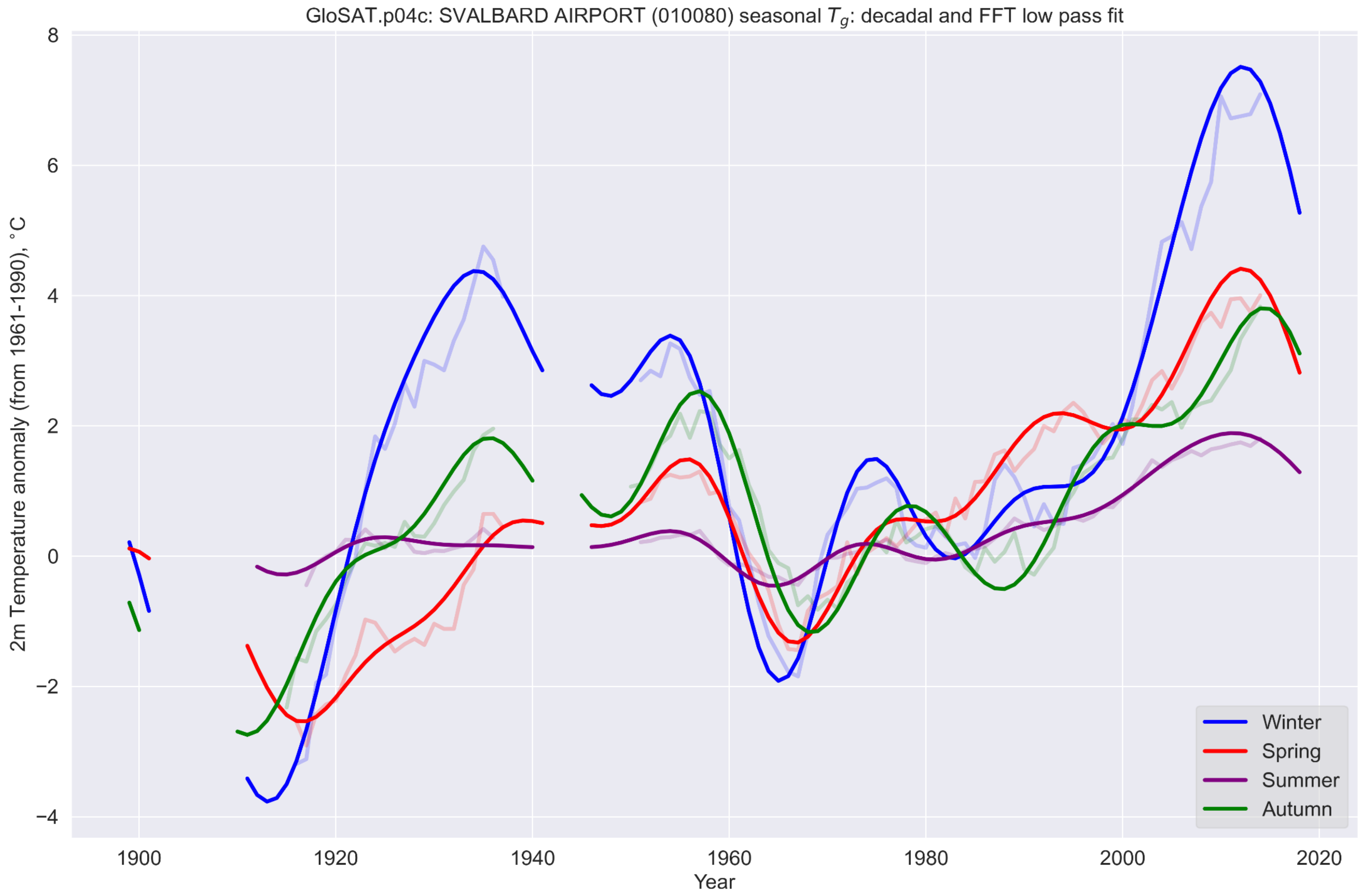
- I knew that this paper had monthly averages for Haparanda (in N, Sweden) in the paper from 1802, but no-one has used it.

- This was added to the CRU dataset earlier this year. It will find its way into datasets once the CRU data get downloaded by the other groups.

- I've requested data from scientists working in Brazil and Indonesia, and a few other countries. I have never been successful requesting data from Indian climate scientists. Have been told they are the property of IMD (Indian Met Department), even the record summaries held in Britain for years pre-1947.



Met Norway paper and report on the record – includes Rajmund


GloSAT.p04c: SVALBARD AIRPORT (010080) seasonal T_g: decadal and FFT low pass fit
2m Temperature anomaly (from 1961-1990), °C
8
6
4
2
0
−2
−4
1900
1920
1940
1960
1980
2000
2020
Year
Winter
Spring
Summer
Autumn

Causes of Large-Scale Temperature changes

Explaining the temperature record has focused on two aspects

- High-Frequency Variability – related to ENSO

- Longer timescale trends – related to Anthropogenic forcing

All IPCC Reports from 1995 onwards

Jones, P.D., 1989: The influence of ENSO on global temperatures. Climate Monitor 17, 80-89.

Hansen, J., I. Fung, A. Lacis, et al. 1988. Global climate changes as forecast by the Goddard Institute for Space Studies three-dimensional model. *Journal of Geophysical Research* **93**: D8, 9341–9364.

Wigley, T.M.L., 2020: How Good Are Past Predictions of Global Warming? *Skeptical Enquirer* **44**, 45-49.

Conclusions

- Biases generally much more important than individual station homogeneity issues
- Exposure issues pre-Stevenson screens an important issue in Europe before about 1880. Important for proxy climate calibration in Europe
- Urbanization issues relatively unimportant at large-scales, but maybe issues at local scales
- Biases with SSTs much more important than biases over land. Important issues with the introduction of drifters since the 1990s
- Early assessments (e.g. Callendar in 1938 and 1961) agree well with modern land-based estimates
- These were good due to the limited number of spatial degrees of freedom for monthly-mean temperature. We don't need tens of thousands of stations to measure large-scale averages. We do need lots of stations to get local details in gridded products