Homogenisation algorithm skill testing with synthetic global benchmarks for the ISTI

International Surface Temperature Initiative

Budapest, May 2014.

Kate Willett (Met Office Hadley Centre) and the Benchmarking Working Group

www.surfacetemperatures.org

Talk Outline

- The ISTI – facilitating robust climate analysis
- The basics of benchmarking for the ISTI
- Creating a 'clean' synthetic world
- Creating a set of dirty/error filled worlds
- Assessing homogenisation algorithm skill against the benchmarks
- Where are we now...

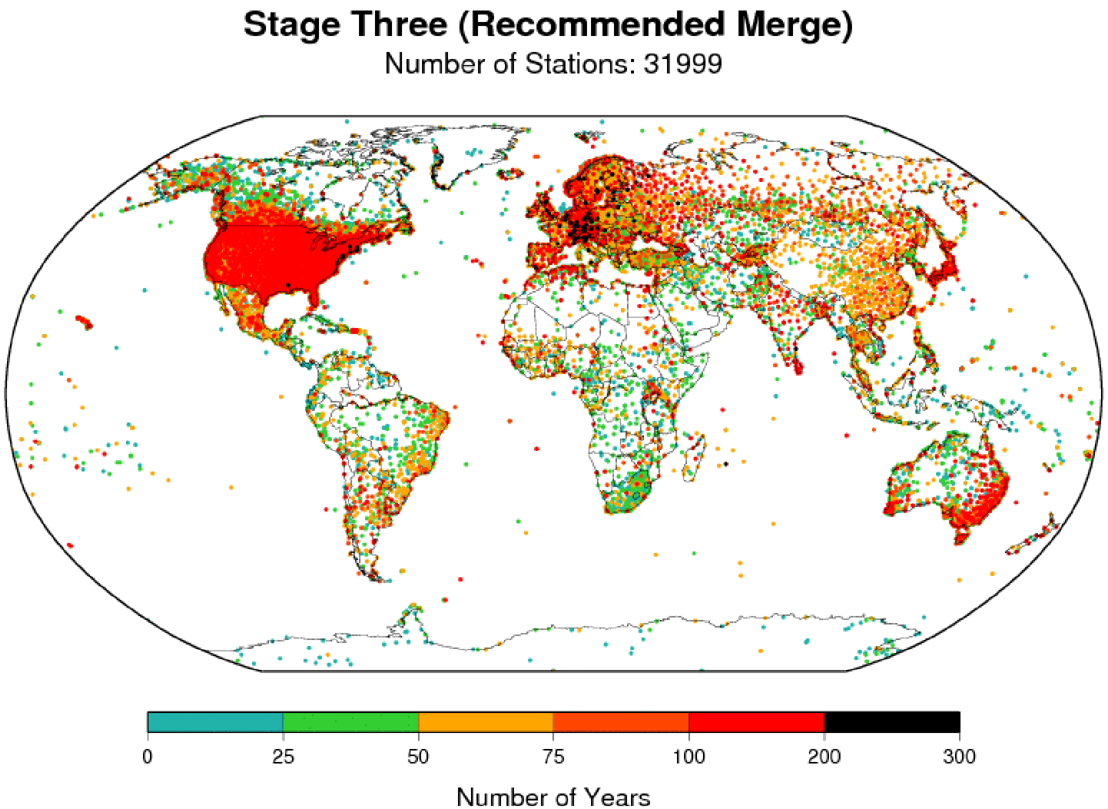
International Surface Temperature Initiative

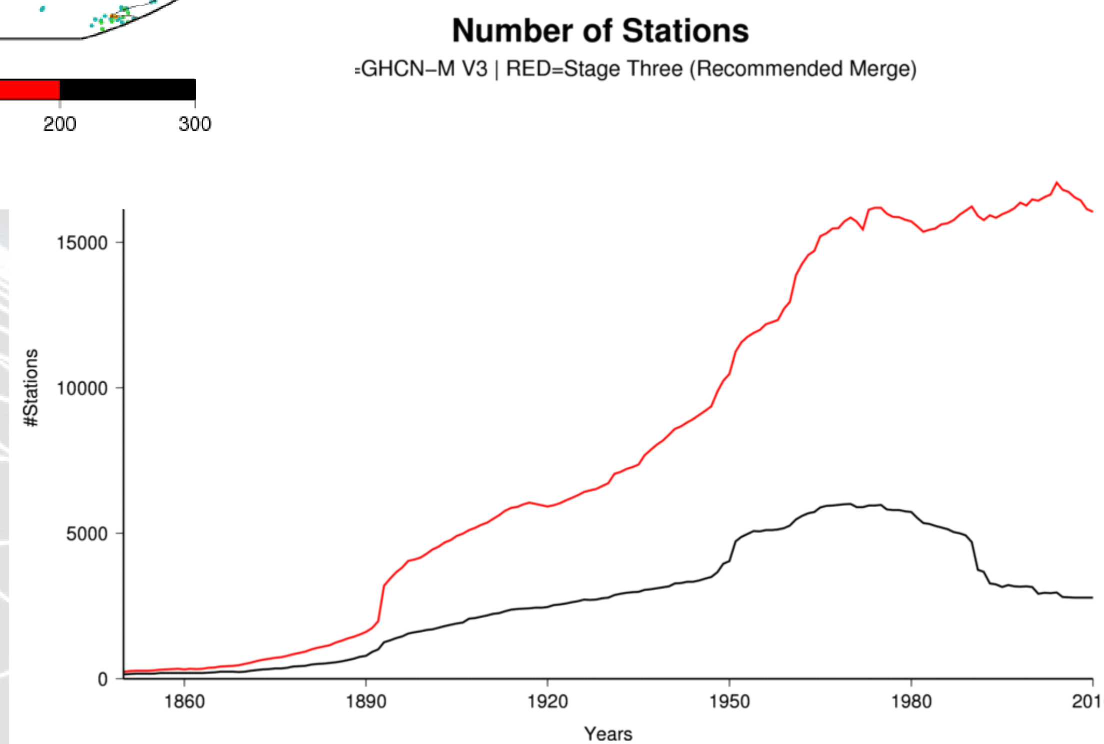
The ISTI Facilitation of Robust Climate Analysis

The ISTI

- A framework for creation of multiple robust independent estimates of land surface temperature to answer scientific questions and societal demands of the 21st Century
  - DATABANK: open, transparent processing, traceability to standards and source
  - BENCHMARKING: Consistent performance evaluation and methodological uncertainty estimates
  - USER TOOLS: fit for purpose, visualisation, intercomparison

The rest is down to the global science community...





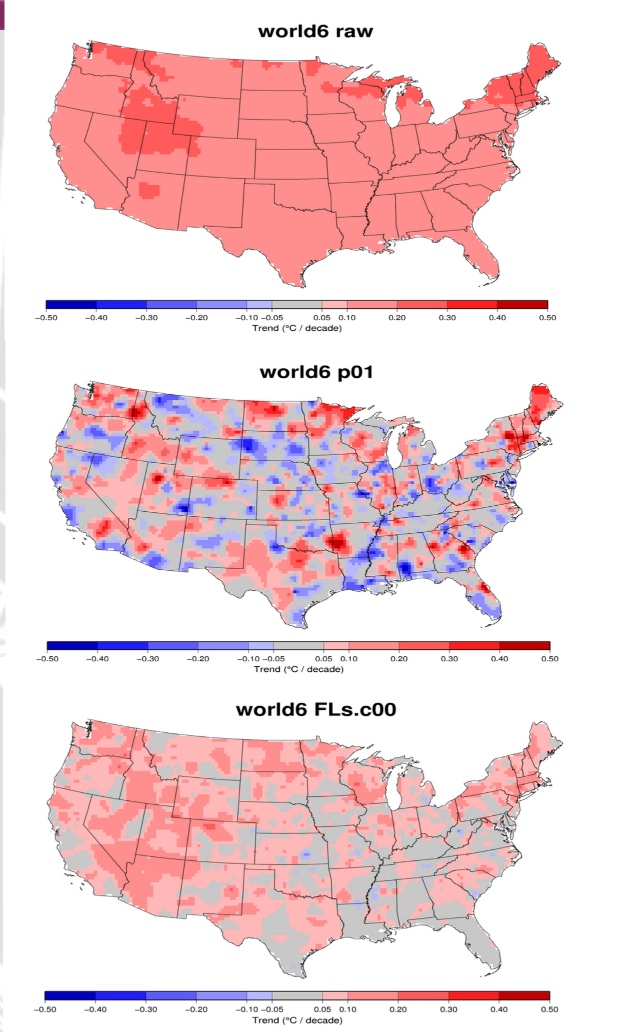
International Surface Temperature Initiative

The Basics of Benchmarking for the ISTI

Benchmarking Cycle

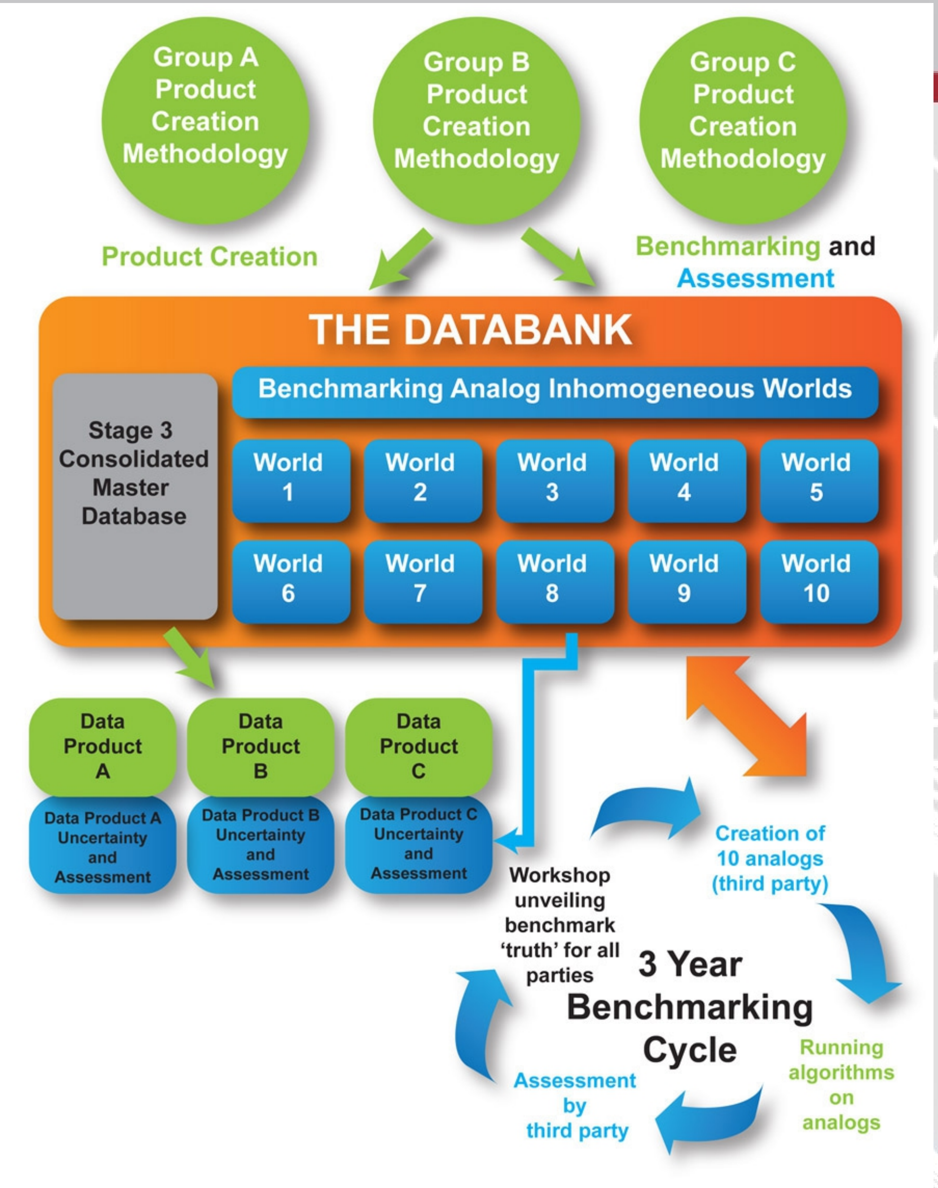
Create c.10 analog-error-worlds

- Simulate 'clean' spatio-temporal characteristics of actual stations underpinned by low frequency variability from a climate model to maintain plausible spatial correlation
- Add abrupt and gradual changepoints to approximate our best guess real world error structures
- Run homogenisation algorithms on the test data and assess ability to recover original 'clean' data
- Useful for further improvement of algorithms



Example use of benchmark data for USHCN

Benchmarking cycle



International Surface Temperature Initiative

Creating a 'Clean' Synthetic World

The real world observing system is not perfect …

US Climate Reference Network website

Its more like these ...

Huge range of instrument types, siting exposures etc. regionally, nationally and globally with many changes over time.

Many more examples on www.surfacestations.org

'TRUTH' UNKNOWN

$$XTRUTH_{t,l} = S_{t,l} + T_{t,l} + V_{t,l} + M_{t,l}$$

XTRUTH = a climate element at time *t* and location *l*

S = seasonal cycle

T = trends (long-term signal)

V = variability (ENSO,
NAO, Volcanoes, Solar Cycles...)

M = microclimate (topography, proximity to coast,
prevailing wind, local environment...)

$$XOB_{t,l} = XTRUTH_{t,l} + \varepsilon_{t,l} + \lambda_{t,l}$$

XOB = observation at time *t*, location *l* and height *h*

ε = random error at time/place/height
(recording error, instrument error etc.)

λ = systematic error at time/place/height *possibly correlated*
(station move, exposure change, instrument change,
observing practice change, urbanisation etc.)

Team Creation: How To...

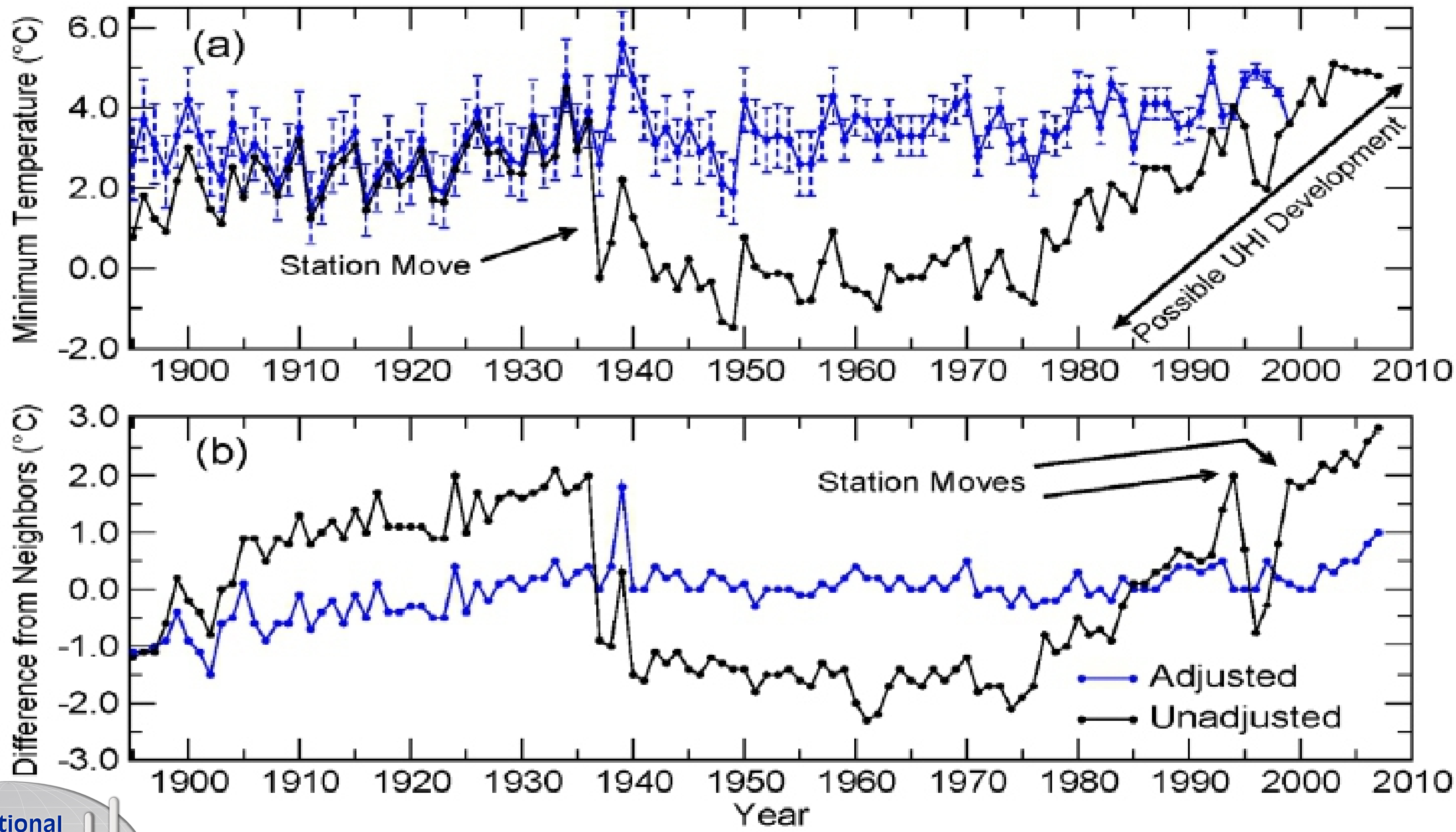
Seasonal Cycle (S) from the real station

Trend (T) and some Variability (V) from a GCM gridbox time series

Variability (V) and Microclimate (M) from both the standard deviation of the real station and a Vector Autoregressive (VAR) modelled time series

Creating a set of Dirty/Error-filled Worlds

Inhomogeneities: annual mean minimum temperature at Reno, Nevada, USA



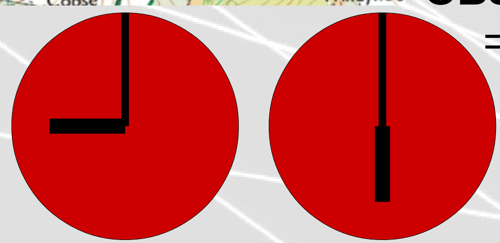
(Matt Menne and Claude Williams, NOAA National Climatic Data Center)

International Surface Temperature Initiative

Effects of Changes that are not of Climate Origin

STATION MOVE: EXPOSURE AND MICROCLIMATE = abrupt change in mean and diurnal extremes - may affect seasonal cycle extremes

SHELTER CHANGE: EXPOSURE = abrupt change in diurnal extremes - may affect seasonal cycle extremes



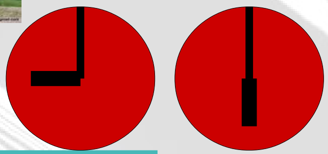
OBSERVING PRACTICE CHANGE: SAMPLING = abrupt change possible in mean and extremes

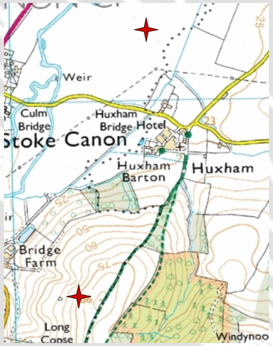
INSTRUMENT CHANGE: CALIBRATION = abrupt change in mean and possibly extremes

LANDUSE CHANGE: EXPOSURE AND MICROCLIMATE = gradual change in mean and diurnal extremes - may affect seasonal cycle extremes

Team Corruption

$$XERRORWORLD_{t,l} = XTRUTH_{t,l} + \lambda ERRORWORLD_{t,l}$$

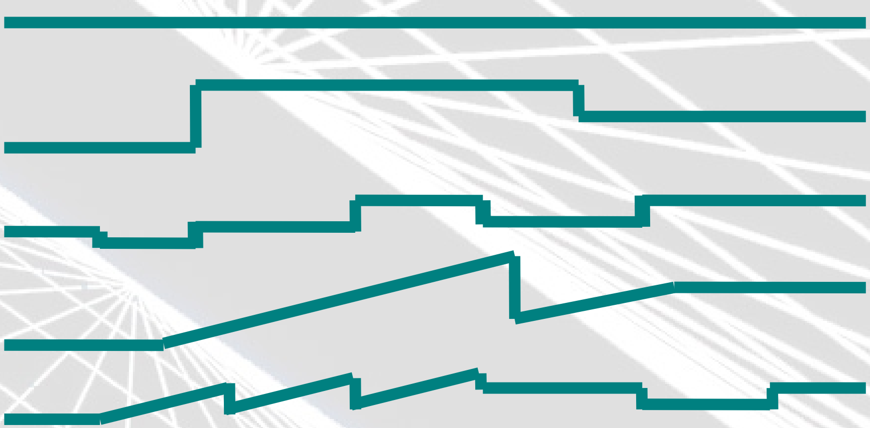




SURFACE TEMPERATURE DATABANK

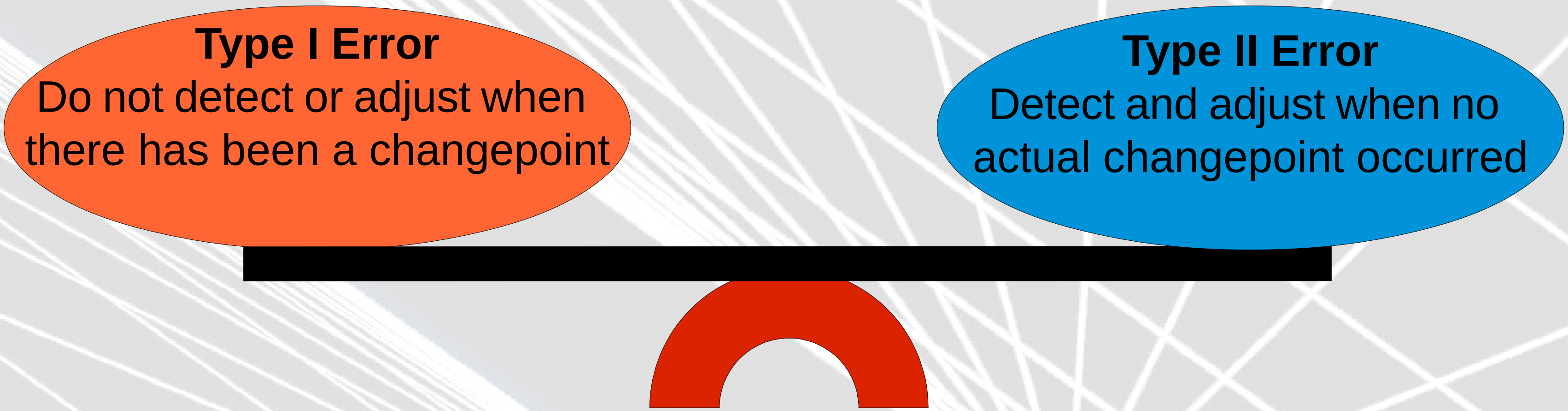
- World 1: no breaks
- World 2: few large simple breaks
- World 3: many small simple breaks
- World 4: few large complex breaks
- World 5: many small complex breaks
- etc.

Example error models applied to stations



Assessing Homogenisation Algorithm Skill Against the Benchmarks

Confidence in Adjustments Made?

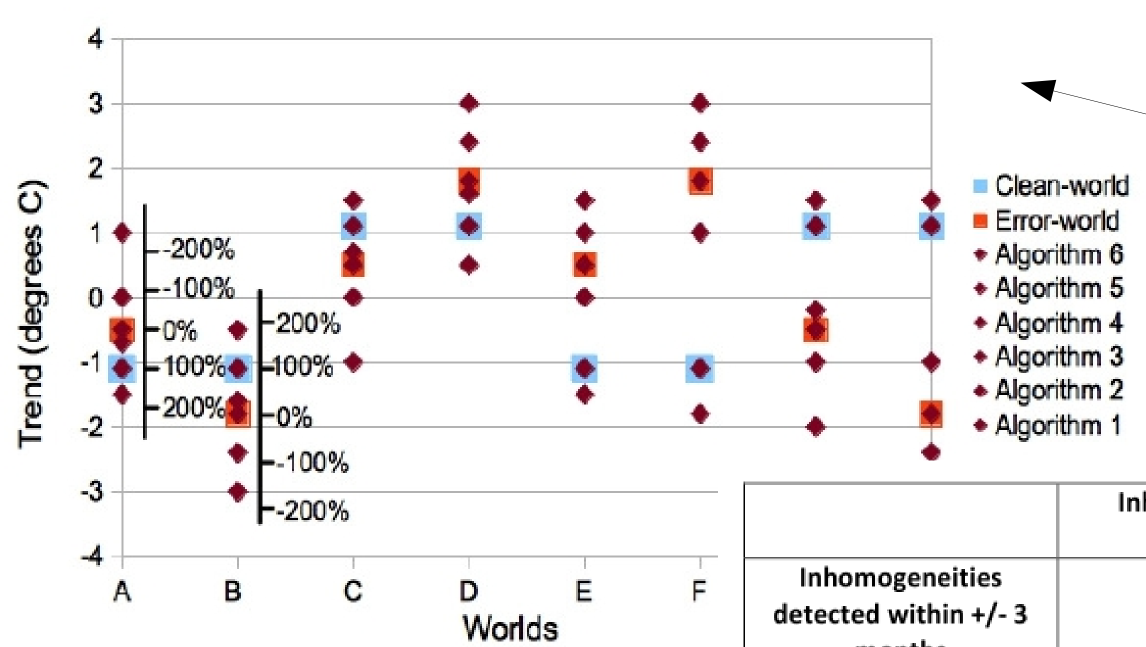


Missed adjustments vs false alarms: which is worse?

What about adjustments in the wrong direction?

Adjustments that are the wrong size/length or do not correctly adjust across the seasonal cycle?

Team Validation

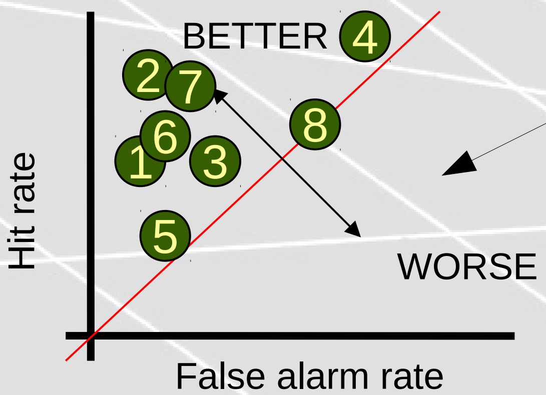


Trend recovery percentage

Contingency Table

	Inhomogeneity present	No inhomogeneity present	TOTALS
Inhomogeneities detected within +/- 3 months (Adjusted value must be correct sign (+/-) and within +/-1 °C)	HITS: 5 (4)	FALSE ALARMS: 3 (3)	8 (7)
Inhomogeneities not detected within +/- 3 months (Adjusted value incorrect sign and not within +/- 1 °C)	MISSES: 2 (3)	CORRECT MISSES: 42 (42) (potential detections)	44 (45)
TOTALS	7 (7)	45 (45)	52 (52)
Heidke Skill Score		61%	
Probability of Detection Hit Rate		71%	
False Alarm Rate		7%	

ROC Curve



Where are we now...

Team Creation – getting there

Team Corruption – error worlds now defined

Team Validation – levels defined

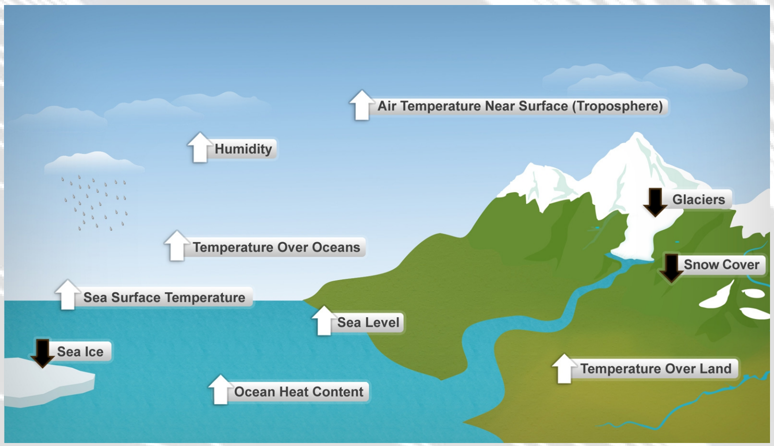
Aim for v1 release: mid-2014

Questions

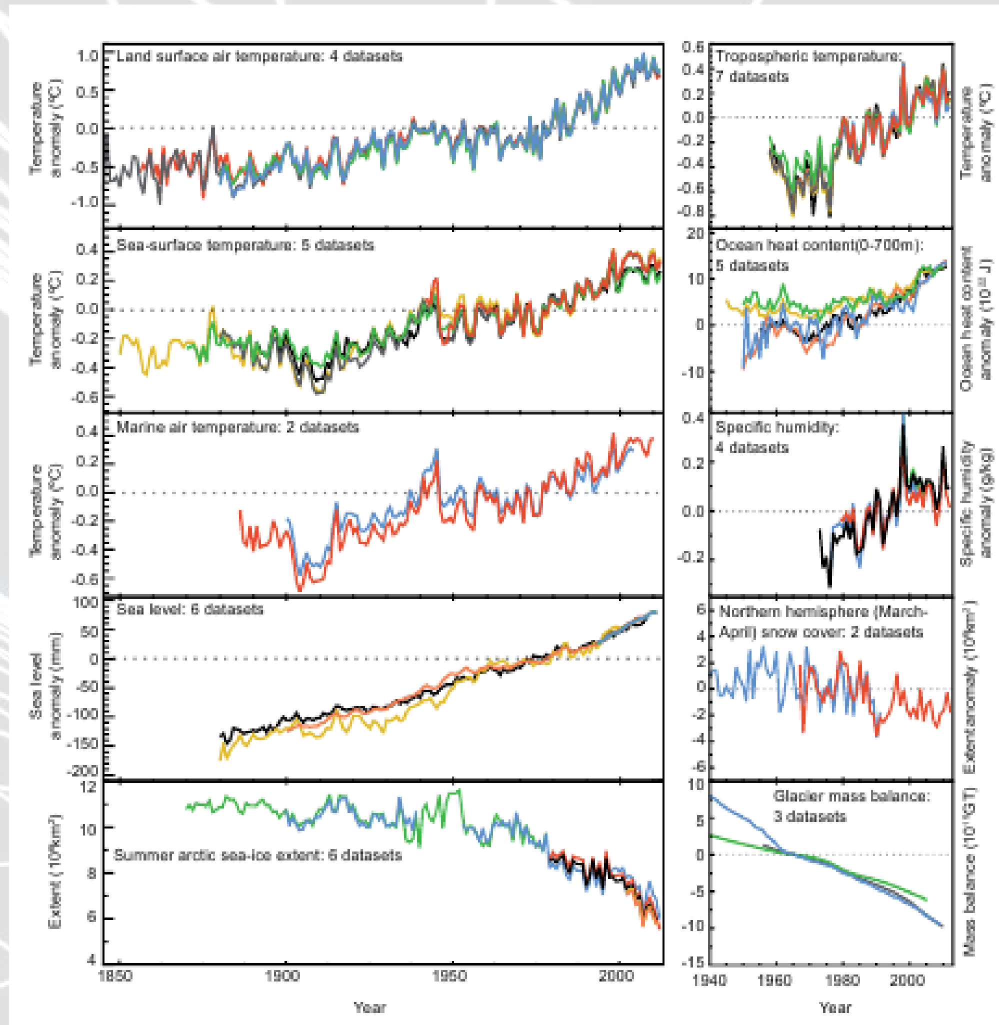
www.surfacetemperatures.org

kate.willett@metoffice.gov.uk
@Kate_M_Willett

No doubt that it is warming – the rate and temporal / spatial details are the issue



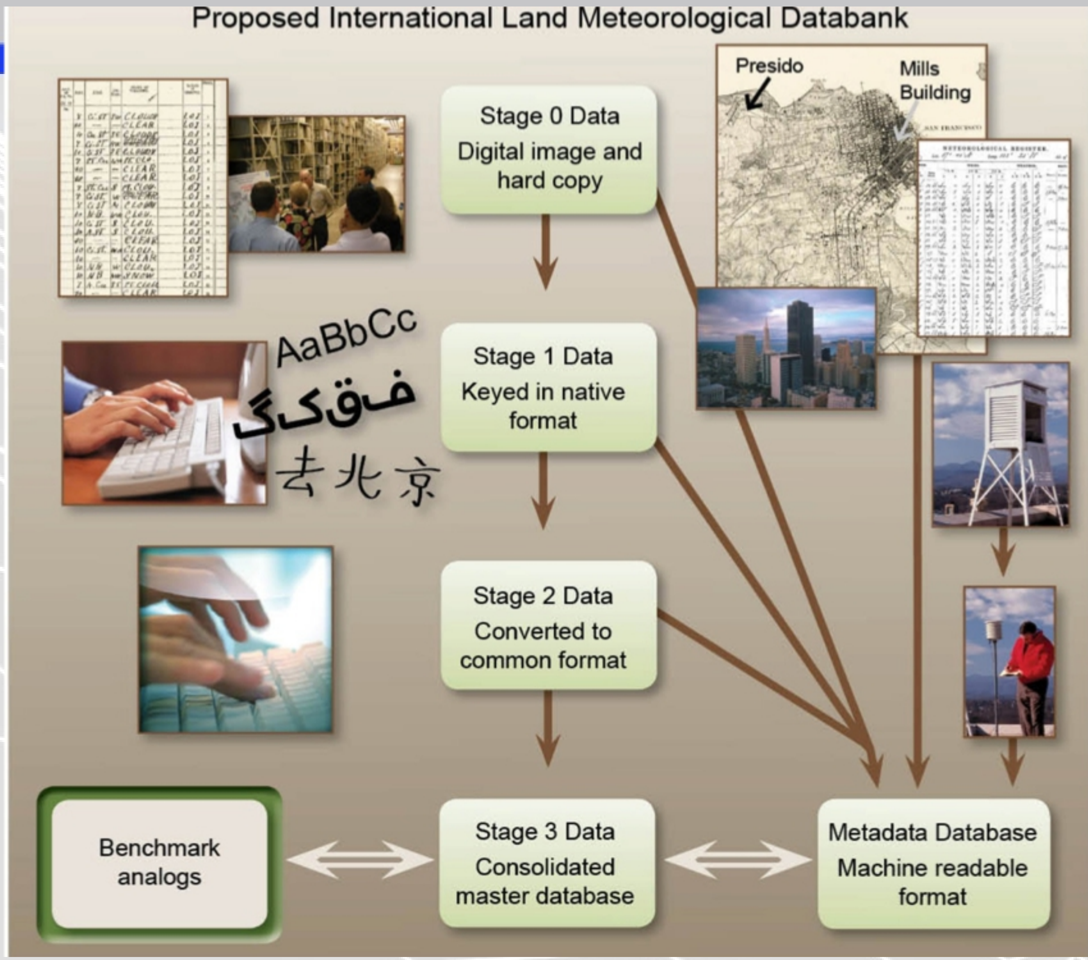
(NCDC Graphics Team (above), John Kennedy, Met Office Hadley Centre (right))



IPCC AR5 Ch2 FAQ2.1 Fig 2

International Surface Temperature Initiative

Step 1: Data rescue and provision

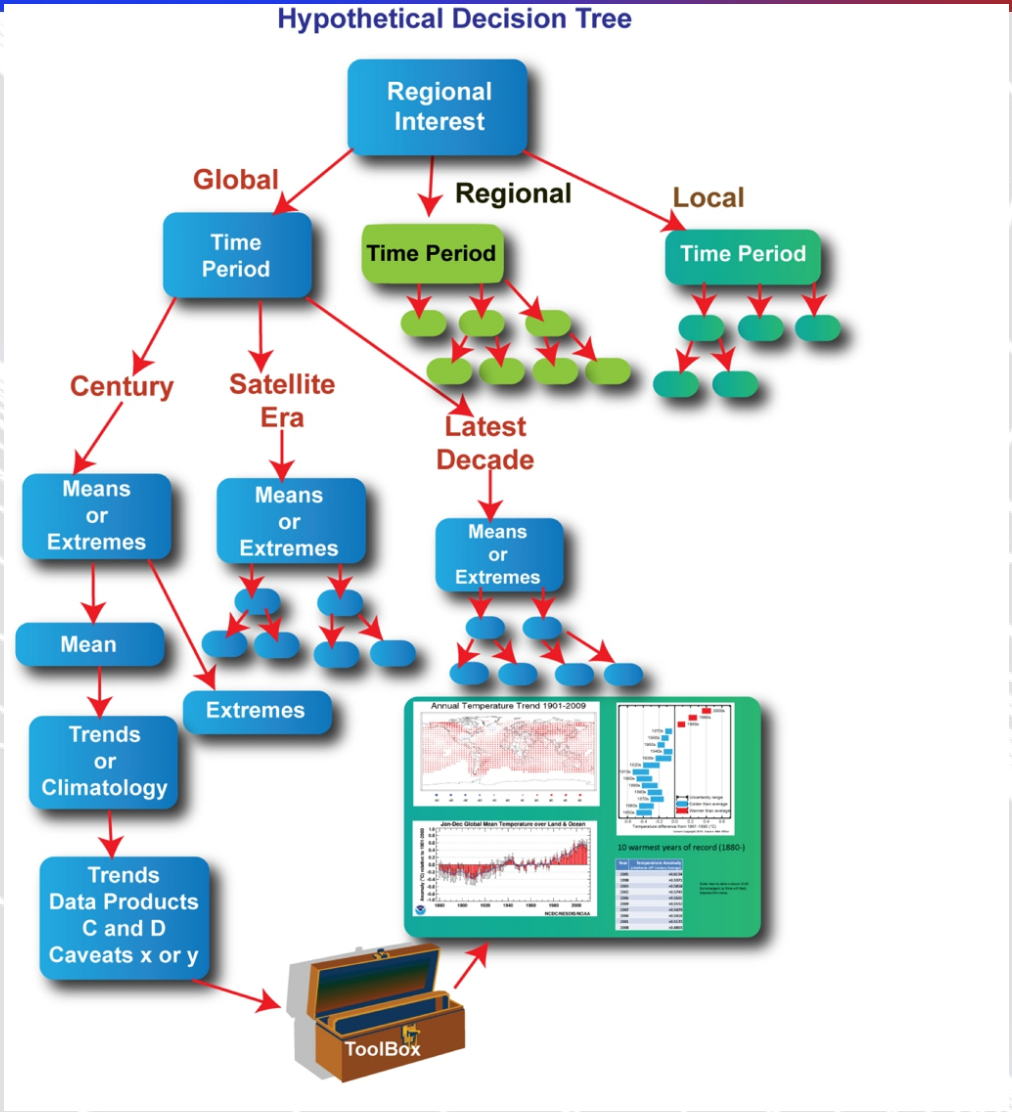


www.surfacetemperatures.org/databank

Lawrimore et al., 2013: Responding to the Need for Better Global Temperature Data, **EOS**, 94 (6), 61–62 DOI: 10.1002/2013EO060002

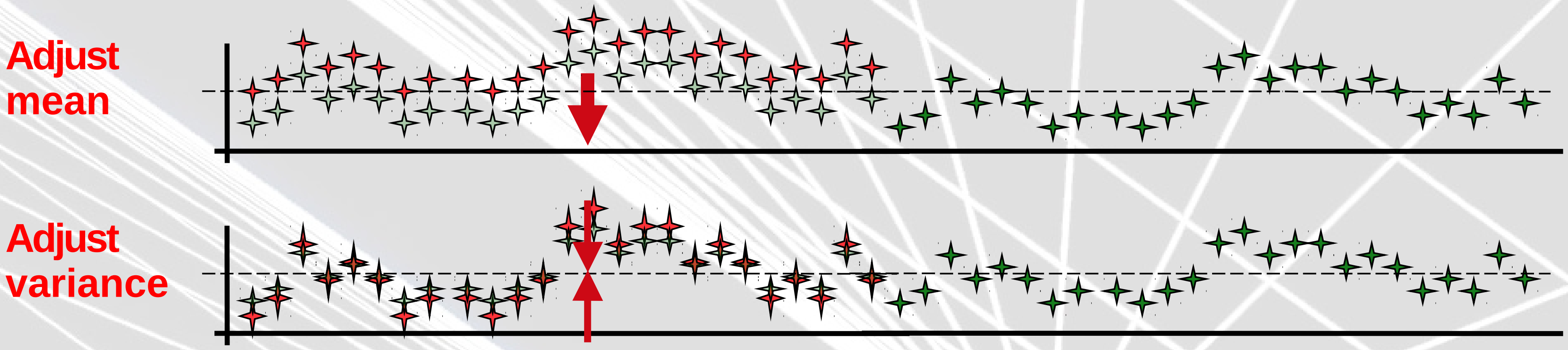
Rennie et al., accepted: The International Surface Temperature Initiative Global Land Surface Databank: Monthly Temperature Data Version 1 Release Description and Methods. **Geosciences Data Journal**

Step 3: Serving products and aiding users



International Surface Temperature Initiative

Adjustment



By month? By season? By full homogeneous period?

Use candidate – neighbour fields?

Use SNHT, PHA, MISHMASH, SPLIDHOM etc?

Fit a model?