The ISTI: Land surface air temperature datasets for the 21st Century

International Surface Temperature Initiative

WRCP Extremes, Feb 2015

Kate Willett, with thanks to many Initiative participants
The real world observing system is not perfect ...
Its more like these ...

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

More examples on www.surfacestations.org
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
Inhomogeneities: annual mean minimum temperature at Reno, Nevada, USA

(Matt Menne and Claude Williams, NOAA National Climatic Data Center)
Underlying these are four fundamental issues ...

• A lack of **traceability** to known standards and original hard copy data sources for most historical records

• A lack of adequate **documentation** of the (ubiquitous) changes (station location, shelter, observing time etc.) sufficient to characterize their changing measurement characteristics

• A lack of '**one-stop-shop**' for all land meteorological data (like ICOADS for land) – both raw and CDR/value-added-products

• A lack of set **benchmarks** with which to comprehensively test Climate Data Record development
No doubt that it is warming – the rate and temporal/spatial details are the issue

Is our climate changing?  What about the cows?

Figure 2. The average number of days (over 1973–2012) with THI > 70 for each of the 68 stations. Stations with no days with THI > 70 are shown as grey squares.
ISTI: Creating a framework to enable advances

1. Basic environmental data provision

2. Benchmark assessment of uncertainty relating to methodological choices

3. User advice
Step 1: Data rescue and provision

Jay Lawrimore, Jared Rennie and Peter Thorne (2013) Responding to the Need for Better Global Temperature Data, EOS, 94 (6), 61–62
DOI: 10.1002/2013EO060002
ISTI Stage 3 vs GHCNv3 data

GHCPM v3
Number of Stations: 7280

Stage Three (Recommended Merge)
Number of Stations: 31999

Number of Years
More than a little better?
Step 2: Benchmarking and Assessment

• With real world data we do not have the luxury of knowing the truth – we CANNOT measure performance of a specific method or closeness to real world truth of any one data-product.

• We CAN focus on performance of underlying algorithms (AKA software testing)

• Consistent synthetic test cases, simulating real world noise, variability and spatial correlations potentially enable us to do this
Bencharking 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
Daily Benchmarks for the USA using a GAM

Stations with temperature records from 1970 to 2011 in the contiguous USA:
Blue = Focus stations with no more than a quarter of the record missing
Step 3: Serving products and aiding users
What happens if you build a state of the art playground and nobody turns up?

• The Initiative will have provided a framework which should be conducive to scientists coming and having a ‘play’
• The Initiative cannot compel scientists to ‘come and play’
• Nor does it have dedicated funding support to offer ...
Q & A

www.surfacetemperatures.org

Bull. Amer. Met. Soc. doi: 10.1175/2011BAMS3124.1

General.enquiries@surfacetemperatures.org

Data.submission@surfacetemperatures.org
Parallel Observations Science Team (POST)

(http://www.surfacetemperatures.org/databank/parallel_measurements)
Long instrumental climate records are usually affected by non-climatic changes, such as relocations and changes in instrumentation, instrument height or procedures. They can hamper the assessment of trends and variability in climate records by distorting the climate signal, especially when it comes to trends in extreme weather.

The most direct way to study the influence of these non-climatic changes and to understand the reasons for these biases is the analysis of parallel measurements representing the old and new situation (for example instruments or locations).

**A GLOBAL PARALLEL CLIMATE DATASET**

Current studies of non-climatic changes using parallel data are limited to local and regional case studies. However, the effect of specific transitions depends on the local climate and the most interesting questions are about the systematic large-scale biases produced by transitions that occurred in many regions. Important potentially biasing transitions are the adoption of Stevenson screens or the introduction of automatic weather stations. Thus a large global parallel dataset is highly desirable as it enables the study of systematic biases in the global record.

A standard Stevenson screen and a historical French screen measuring in parallel at an observatory in Basel, Switzerland.
Multiplicty of data products

- Quantifying structural uncertainty is key
- Raw data is far from traceable to international measurement standards.
- Data artifacts are numerous and have myriad causes
- Metadata describing station histories is patchy at best and often non-existent
- Data is discrete in both space and time
- No “how to” … rather very many cases of “it may work …”
- Multiple subjective decisions required even in automated procedures (thresholds, periods)
- Different approaches may have different strengths and weaknesses
- No single dataset can answer all user needs
Stage 1 - Native format digitized
Stage 2 – common format

<table>
<thead>
<tr>
<th>Provenance / version control flags</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAO_BANG 22.6700 106.2500</td>
</tr>
<tr>
<td>CAO_BANG 22.6700 106.2500</td>
</tr>
<tr>
<td>CAO_BANG 22.6700 106.2500</td>
</tr>
<tr>
<td>CAO_BANG 22.6700 106.2500</td>
</tr>
<tr>
<td>CAO_BANG 22.6700 106.2500</td>
</tr>
<tr>
<td>CAO_BANG 22.6700 106.2500</td>
</tr>
<tr>
<td>CAO_BANG 22.6700 106.2500</td>
</tr>
<tr>
<td>CAO_BANG 22.6700 106.2500</td>
</tr>
</tbody>
</table>

International Surface Temperature Initiative
Stage 3 (under beta)

• Same format as stage 2
• Optimised station merging of non-unique records
• One unique version for each station – recommended version for most users
• Forms basis for creation of benchmarking analog stations (see later)
• Provenance tracking ensures an unbroken chain of stages
Station series example

01003713 (WHITECOURT) 1942–2012
LAT= 54.1500 LON= -115.7800 ELEV= 785.00

TMAX (°C)


01 35

TMIN (°C)


01 35

TAVG (°C)


01 35 39 41

International Surface Temperature Initiative
Benchmarking cycle
Benchmarking example

• For USHCN (lower 48 states)
• 100 member perturbed ensemble of the NCDC pairwise algorithm was run on 8 analogs (Williams et al., 2012, JGR-A, 117, D05116, doi:10.1029/2011JD016761)
• Consideration solely of timeseries and trends
• Analyses that follow are for the hardest analog with frequent predominantly small breaks added
a) Trends from Raw Input Data With Errors
Implication when applied to real-world observed record that warming magnitude and spatial patterns for the USA are reasonably well captured with some regional discrepancies.
d) Homogenized Data (Berkeley Earth Surface Temperature Method)

Implication when applied to real-world observed record that warming is robust but magnitude is overestimated and spatial patterns not well captured for the USA.
By analogy ...
The field is wide open ...

- We have thus far sampled only a small area of solution space which has many d.o.f
- We need to far more fully explore the plausible solution space
- Many possibilities exist

Bottom line:
Please please please come and play