ISTI: Building a Land Surface Temperature Program for the 21st Century

International Surface Temperature Initiative

4th ACRE Workshop, September, 2011

Dr Kate Willett,
Met Office Hadley Centre, UK
Talk Outline

• Background and purpose
• The Land Surface Databank: data rescue, acquisition, provenance and version control
• Benchmarking and assessment of data-products
• Serving data and data-products
Background and Purpose
Examples of Temperature Reconstructions

IPCC Fourth Assessment Report (2007), Figure 3.1
Annual anomalies of global land-surface air temperature (°C), 1850 to 2005, relative to the 1961 to 1990 mean for CRUTEM3 updated from Brohan et al. (2006). The smooth curves show decadal variations … The black curve from CRUTEM3 is compared with those from NCDC (Smith and Reynolds, 2005; blue), GISS (Hansen et al., 2001; red) and Lugina et al. (2005; green). See http://www.ipcc.ch/publications_and_data/ar4/wg1/en/ch3s3-2-2.html.
The Big Question

- Can we create a process for producing a suite of independent verified estimates of land surface temperatures to answer scientific questions and societal demands of the 21st Century?
  - Open and transparent
  - Better understanding of fundamental instrument performance and measurement properties
  - Consistent performance evaluation
  - User tools
  - Not just monthly at the largest scales. Daily, sub-daily, regional and local
In the beginning...

- 2010 UK Met Office Submission to WMO Commission for Climatology
  - Call for creating new suite of products to meet 21st Century demands / expectations
- September 2010 kick-off workshop, UK Met Office, Exeter
  - 80 international experts including climate scientists, metrologists, statisticians, software engineers
  - White papers posted online and public comments solicited
  - Agreed project outline and governance structure
The Land Surface Databank
Data Acquisition and Provision

Proposed International Land Meteorological Databank

- Stage 0 Data: Digital image and hard copy
- Stage 1 Data: Keyed in native format
- Stage 2 Data: Converted to common format
- Stage 3 Data: Consolidated master database
- Metadata Database: Machine readable format

International Surface Temperature Initiative
Databank Progress

• Working Group set up
  – Data rescue task team
  – Provenance and version control task team
• Development version posted

http://www.gosic.org/GLOBAL_SURFACE_DATABANK/GBD.html

• First version release and accompanying documentation / paper submitted spring 2012
Stage 0 Data in Hardcopy / Image Format

- Millions of rescued images in the NOAA Foreign Data Library
- 2000+ boxes in NCDC library
- Holdings in other libraries and repositories, particularly former colonial powers
- Holdings rotting away or not valued in many countries
- Data not held by NMSs

http://docs.lib.noaa.gov/rescue/data_rescue_home.html
Data currently digitized:

- Numerous public holdings
  - NCDC, UCAR, RIHMI
  - Regional and national holdings
- Some records are restricted for commercial or geopolitical reasons
## Stage 2 – Common Format

### Provenance / version control flags

<table>
<thead>
<tr>
<th>Location</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Temperature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAO_BANG</td>
<td>22.6700</td>
<td>106.2500</td>
<td>-9999</td>
<td>101/105/101/102/103</td>
</tr>
<tr>
<td>CAO_BANG</td>
<td>22.6700</td>
<td>106.2500</td>
<td>-9999</td>
<td>301/105/101/102/103</td>
</tr>
<tr>
<td>CAO_BANG</td>
<td>22.6700</td>
<td>106.2500</td>
<td>-9999</td>
<td>18980222</td>
</tr>
<tr>
<td>CAO_BANG</td>
<td>22.6700</td>
<td>106.2500</td>
<td>-9999</td>
<td>301/105/101/102/103</td>
</tr>
<tr>
<td>CAO_BANG</td>
<td>22.6700</td>
<td>106.2500</td>
<td>-9999</td>
<td>18980223</td>
</tr>
<tr>
<td>CAO_BANG</td>
<td>22.6700</td>
<td>106.2500</td>
<td>-9999</td>
<td>301/105/101/102/103</td>
</tr>
<tr>
<td>CAO_BANG</td>
<td>22.6700</td>
<td>106.2500</td>
<td>-9999</td>
<td>18980224</td>
</tr>
<tr>
<td>CAO_BANG</td>
<td>22.6700</td>
<td>106.2500</td>
<td>-9999</td>
<td>301/105/101/102/103</td>
</tr>
<tr>
<td>CAO_BANG</td>
<td>22.6700</td>
<td>106.2500</td>
<td>-9999</td>
<td>18980225</td>
</tr>
<tr>
<td>CAO_BANG</td>
<td>22.6700</td>
<td>106.2500</td>
<td>-9999</td>
<td>301/105/101/102/103</td>
</tr>
</tbody>
</table>
Stage 3 – Consolidated Land Surface Database

- Same format as stage 2
- One unique version for each station – recommended version for most users
- Basis for creation of benchmark analogs
- Protocols used in merging sources are as yet to be finalized
- Provenance tracking will ensure an unbroken chain to earlier stages

Release of version 1 planned April 2012
Partnerships are Essential

• Bring together existing efforts, augment and ensure pull through. e.g. ACRE project (http://www.met-acre.org/), IEDRO (www.iedro.org) and other national / international programs.
• Pursue innovative approaches (crowdsourcing building upon success of oldweather.org, www.data-rescue-at-home.org etc.)
• Build on ICOADS model for sea surface temperatures (http://icoads.noaa.gov/) – easy submission and access to data
• Recognize key partners and contributions
• And you can also help ...

www.surfacetemperatures.org/databank
Data.submission@surfacetemperatures.org

http://www.surfacetemperatures.org/databank/DataSubmission-Stage1-Guidance.pdf?attredirects=0

http://www.youtube.com/watch?feature=player_embedded&v=CEaTQjzc0zo
Benchmarking and Assessment
Benchmarking and Assessment

- With real world data we do not have the luxury of knowing the truth – we CANNOT measure closeness to real world truth of any one data-product.
- We CAN focus on performance of underlying algorithms
- Consistent synthetic test cases, simulating real world noise, variability and spatial correlations potentially enable us to do this

**Inhomogeneities:**
*annual mean min temp at Reno, Nevada, USA*
Benchmarking progress

- Working group and task teams instigated
- Create c.10 analog-error-worlds
  - Climate model basis (maintains plausible spatial correlation structure) tweaked with real station climate characteristics
  - Add random and systematic errors to approximate the real world error structures which may exist
  - Algorithm assessment based upon ability to recover original data
  - Error structures should enable answering a range of questions / assumptions regarding the true error to avoid over-tuning

Analogs to be made available November 2012 based upon version 1 release of databank
Benchmarking Cycle

Example use of benchmark data for USHCN
Data-Product
Portal
Serving Products and Aiding Users

- Data formats – netCDF, ASCII?
- Degree of user interaction – data-subsets?
- Tools - Ability to create graphical and tabular output on the fly
- Limited progress to date
  - Largely a reflection that this data provision is some way down the road?
  - Ideas and suggestions welcome …
Progress Summary

• Steering Committee (www.surfacetemperatures.org/steering-committee)
  – Terms of Reference
  – Endorsed by World Meteorological Organization (WMO) and The International Environmetrics Society (TIES) - International Bureau of Weights and Measures (BIPM) pending

• Working groups on databank and benchmarking
  (www.surfacetemperatures.org/databank)
  (www.surfacetemperatures.org/benchmarking-and-assessment-working-group)
  - Databank prototype made public and data sources coming in

• Implementation Plan published

• Progress documented on Initiative website at:
  www.surfacetemperatures.org
You can help make this a success

• Help to find raw data sources
• Come up with novel ways of analyzing the data
• Create data-products using the databank
• Partake in the benchmarking exercise
• Help to make a data portal a reality
• Provide constructive feedback
Questions and Answers

www.surfacetemperatures.org

General.enquiries@surfacetemperatures.org

Data.submission@surfacetemperatures.org