

Data rescue task team call 3/13/12
8EST (12Z / GMT, 13 CET, ...)

Dial in details circulated separately

Present on call: Peter Thorne, Rod Hutchinson, Jay Lawrimore, Hermann Maechel, Jared Rennie, Stefan Bronimann, Rick Crouthamel

Apologies in advance: Manola Brunet, Juerg Luterbacher, Rob Allan

Agenda

1. Review of actions from last call - Peter

ACTION: PT to contact Lisa Alexander and NCDC normals team to put in touch with IEDRO about making climdex and normals processing facilities available to IEDRO. An email was sent. Unclear whether anything accrued. RC noted that ball was now in his court. PT urged that IEDRO use pre-existing algorithms where possible to aid intercomparisons and avoid any redundancy.

2. Update on overall initiative - Peter

The initiative continues to move forwards. A progress report was recently posted for the initiative as a whole. This is available through surfacetemperatures.org. The initiative is now recognized in some shape or form by each of WMO, BIPM and TIES. There will be a session and keynote on the initiative and its aims at the major metrology once a decade ITS9 meeting next week.

3. Update on progress with databank - Jay

Databank holdings now exist of 37 sources on the ftp area with several more queued. Work is ongoing to create a merged set. <ftp://ftp.ncdc.noaa.gov/pub/data/globaldatabank/monthly/>. The processing code is modular and open source and will be made publicly available. Latest estimate on first release is the summer. We aim to release two papers describing the databank, one more generalist in BAMS, the other more mechanical, probably in JAOT, Databank WG members want us to err on the side of assigning as unique in the merge. We are merging based on metadata, overlapping data and non-overlapping data. Metadata consisting of comparisons of station name, location and elevation. We work through each of the sources based on a previously established priority. Currently GHCN-Daily is set with the highest priority. The priority is based on a process - strawman is available at http://editthis.info/intl_surface_temp_initiative/Source_Hierarchy. Based on metadata comparisons a small set of candidate stations for potential merging with the master source are then compared - either overlapping or non-overlapping data tests. Currently evaluating overlapping tests based on 1) normalized RMSD and 2)

Index of Agreement. Following this we'll develop tests of overlapping data - possibly based on differences of means and variance.

Location of Databank Wiki: http://editthis.info/intl_surface_temp_initiative/. This provides more in depth detail.

4. Update on activities - All (round robin)

NCDC staff continue to try to develop a crowdsourcing portal with zooniverse but have thus far been unsuccessful with getting the requisite funding.

IEDRO have made some efforts on funding crowdsourcing. May have some funding. Josh Wexler. PT has been in contact. Will call Arfon Smith on Monday. Pibal data is currently the focus.

ACTION: PT / JL to send the surface proposal to RC.

IEDRO are proposing a gaming aspect to digitization. Funding through VCP? May still be open.

Juerg Luterbacher (sent in advance): please find below our current digit effort. More station data will follow within 2012. Maybe this information is of use for you and the others

- Mauritius data (1852-1861 and 1909-1915, DONE)
- Christmas Islands (1901-1923 DONE, 1923ff data in progress)
- Jerusalem, Israel (1900-1907; 1910-1913, DONE)
- Lagos, Nigeria (1890-1900, DONE)
 - Papua New Guinea (1892-1895, DONE)
- Abbassia, Egypt (1900-1904, DONE)
- Alexandria, Egypt (1901-1904, DONE)
- Assiut, Egypt (1900-1904, DONE)
- Assuan, Egypt (1901-June 1904, DONE)
 - Helwan, Egypt (1904, DONE)
- Mongalla, Egypt (1904, DONE)
- Port Said, Egypt (July 1900-June 1904, DONE)
- Berber, Sudan (1902-1904, DONE)
- Omdurman, Sudan (1900-1902, DONE)
 - Wadi Halfa, Sudan (1902-1904, DONE)

- Wad Medan, Sudan (1902-1904, DONE)
- Limassol, Cyprus (1933, 1938, 1940, DONE)
- Seychelles (1912-1919, DONE)
- Campo Major (1861-1861, DONE)
- Freetown, Sierra Leone (1910-1913, only pressure, DONE)
- Farafanga, Madagascar (1888-1892, DONE)
 - Maritzburg, South Africa (1859-1866, DONE)
- Aden, Yemen (1847-1859, only pressure, DONE)

- Perim (1891-1898, work in progress, DONE are the years from 1891-1893)
- Zomba, Malawi (1911-June 1946, work in progress, DONE are the years from 1911-1925)
- Barbados (1853ff, work in progress, DONE are the years from 1853-1856)
- Auckland, New Zealand (1853-1869, work in progress, DONE are 3 years)
- Mogadishu, Somalia (1922-1922, work in progress)
- Duem, Sudan (1902-1904, work in progress)
- Giza, Egypt: (1902-1904, work in progress)
- Khartum, Sudan (July 1903-June 1904, work in progress)

E-Mail to send data: jared.rennie@noaa.gov

Rod Hutchinson visited East Timor W/E 2nd March 2012. Agreements in place to digitise East Timorese 1950-1974 Portuguese era and Dili (Comoro Airport) 1982-1999 Indonesia era daily and sub-daily observations. If time permits we will digitise East Timor monthly held by Lisbon/NOAA.

East Timor target manuscripts known to exist:

Records from 1916-1936 - Lisbon Archives – Monthly Data from Portuguese era for 27 East Timor locations. Only three (3) sites have been digitised; Dili, Hato-Lia and Manatuto.

Records from 1950-1974 - Dili Archives – Daily and sub-daily data from Portuguese era for 30 East Timor sites.

Records from 1982-1999 - Darwin BoM - Original SYNOP observations from Indonesian era for Dili (Comoro Airport) with gaps.

Rod Hutchinson will update Rob Allan spreadsheet (<http://badc.nerc.ac.uk/browse/badc/corral/images/metobs>) with his knowledge of digitised/undigitised Pacific data. Have already added East Timor Portuguese era records known to exist.

Monthly sugar mill records for Fiji located in Australia for 1st half of 20th century.

Australian Aid program (PACCSAP) has set aside about AUD\$350K for in-country digitising. Countries have been invited to make an application for funds.

CliDE climate database has been distributed to 14 Pacific Countries and East Timor.

We have a certificate that we can give to countries / providers co-signed by WCRP / CCI / GCOS

ACTION: JL has e-mailed the certificate to the group.

Discussion around rehosting NIWA holdings, which are substantial and would greatly improve the coverage over New Zealand (and potentially some Pacific island regions).

ACTION: RH to chase about what, if any data can be released from NIWA holdings.

John Christy has rescued Ugandan data.

Stefan, still updating ERA-CLIM, Chile has been funded and have started to work. Will be lots (400) of stations, incl. Antarctic peninsula. Inventory will be added to the BADC. East Timor data.

ACTION: Stefan to ensure data is uploaded to databank.

IEDRO will be starting a Data Rescue site at the national meteorological service in Bolivia in April and believe we will begin another site in Namibia later this year. Additionally we are beginning a GIS project which will provide visitors to our web site a map showing all the WMO recognized observation sites in the world. Also loaded will be a listing of all the sites for which NCDC has digitized data in one or more of their data bases. Eventually we hope to have the specific data available and the years of coverage for each station.

Our precipitation strip chart program is nearly operational. Once that is on board we will modify the core program to provide digitization of scanned thermograms which should add a considerable amount of temperature data.

5. AOB

Suggest next call around May / June.