

July 9, 2013 Databank WG Conference Call Summary Notes

1200 - 1320 UTC

Members In Attendance

Jay Lawrimore (NOAA NCDC, USA, Chair)

Albert Klein-Tank (KNMI, Netherlands)

Byron Gleason (NOAA NCDC)

Waldenio Gambi de Almeida (CPTEC/INPE, Brazil)

David Lister (CRU, UEA, UK)

Matt Menne (NOAA/NCDC, USA)

Colin Morice (UK Met Office)

Jared Rennie (NOAA NCDC, USA)

Madeleine Renom (Universidad de la Republica, Montevideo, Uruguay)

Steve Worley (National Center for Atmospheric Research, USA)

Peter Thorne (CICS-NC, USA, Chair, Surface Temperatures Initiative Steering Committee, ex officio)

Apologies in Advance

John Christy (Univ of Alabama-Huntsville)

Matilde Rusticucci (Univ of Buenos Aires, Argentina)

ACTIONS RESULTING FROM THIS (JULY 9, 2013) CONFERENCE CALL

ACTION: ISTI IP comments to Peter / Jay. By Friday.

ACTION: Working Group members to submit metadata and/or parallel measurements now so we can see what exists and inform the data formats.

ACTION: JR to follow up with WG members when review comments are returned.

ACTION: JR to undertake stage 3 version 1 release.

ACTION: Waldenio to assess Brazilian data and advise Jay on potential for gaining additional period of record data.

ACTION: DL to make inquiries regarding whether anyone can download data from CLARIS now.

ACTION: PT to follow up on location and focus of EarthTemp meeting with eye toward gathering additional source data from Africa.

Agenda

- I. Welcome and review action items from recent ISTI conference calls (Lawrimore)
- II. Status of Databank; Beta4 and changes implemented (Rennie, Gleason)
- III. Update on activities of the Int'l Surface Temp Initiative Steering Committee and Benchmarking progress (Thorne, Menne)
- IV. ISTI Implementation Plan update and planned activities associated with the Databank

(Thorne, Lawrimore)

V. Data collection activities; Gaps in coverage and areas of focus at the Regional level (All)

VI. Summarize Activities and Next Steps for coming months

VII. Other Business

I. Welcome and review action items from recent ISTI conference calls (Lawrimore)

From September 27, 2012 DWG call

JL and JR to follow up on outstanding data source issues, BADC and Russian source.

BADC: Conference call with David Lister and Ag Stephens – Not successful.

Russian Source: Ingested into GHCN-D (provided by visiting scientists) and some additional monthly data (web download).

JL to aide MR (Matilde Rusticucci) in pursuing new Argentine source data by providing a letter encouraging contributions to the initiative's efforts.

Sent open letter on October 17. Data still outstanding.

Beta release to occur during first week of October. Accompanied with Overview story, updated surfacetemperatures.org website, and blog postings.

Completed. Beta4 release now online.

JR to continue to add new sources up until 1 December.

Total of 53 sources now (in recommended merge). 58-60 in stage 2 (some sources redundant, e.g, GHCN-M v3, ECA already in GHCN-D). Some new sources still to be added.

Stage 3 data issues to be addressed as comments come in.

Beta 4 release now online.

PT to work with Benchmarking working group to initiate benchmarking of Stage 3 data.

Progress being made. Update later in call.

NCDC to conduct QC and Bias correction testing and development of Stage 3 data.

Ongoing.

From June 4, 2013 ISTI Committees and Working Group Conference call

PT to follow up with GCOS secretariat regarding request to garner metadata for GSN and RBCN network stations and include JL.

Response – they will take to a regional meeting and advocating for metadata. PT to follow up further.

Stefan Bronnimann to ensure ERA-CLIM phase 1 data are submitted to databank.

JR has not heard from Stefan on this. The website (<http://www.era-clim.eu/products/>) says the data will be available in 2013, but no indication when.

II. Status of Databank; Beta4 and changes implemented (Rennie, Gleason)

Since the first beta release in October, there have been 3 additional beta changes relating to formatting, blacklisting and other method tweaks. Beta 4 was released in June. There are 31,427 stations in the Stage3 merged dataset. Graphics showing coverage, number of stations, etc. are available on the surfacer temperatures.org website.

The databank code and logic is largely unchanged. The methods paper has been submitted and is still under review. Blacklisting had the greatest impact. Accounts for things that code logic can't. Things like flipped lats / lons. A copy of the Stage3 dataset in NetCDF format has been completed. It is CF compliant. It is made available in hopes that others will download and test usage.

Compared to other global datasets (GHCN-M v3) there is much better sampling throughout the period of record and no drop-off in stations in the recent past like existing holdings. Improvements are less marked in early record period (although still there).

Since beta 4 the focus has been on how to undertake an update system. Now working on how to append data to existing stations each month and then more periodically do a period of record update adding sources and stations. A few sources are in the queue to be added. Once the journal article is accepted the dataset will go live and start to be updated.

CM: Monthly update is the plan to update just a subset or more holistically?

JR: Plan is to focus on GHCN-D and real-time CLIMAT messages.

CM: Then more bulk updates?

JR: Yes, all sources in PoR updates

AKT: Netcdf files for each station?

JR: Yes. They are station-wise. They have all elements and metadata. Can send the link.

III. Update on activities of the Int'l Surface Temp Initiative Steering Committee and Benchmarking progress (Thorne, Menne)

NCDC hosted a workshop June 1-3 to further engage working group members. Prior to that weekly calls had been held in recent months. Momentum is now building.

There are three teams – creation, corruption and validation. Creation is creating the synthetic datasets. Corruption is adding the breaks. Validation is quantifying performance. 10 blind worlds are to be created and then 3 open worlds where errors are known up front. Blind worlds will be available for a year before analyzed. We included a survey into the known nature and frequency of inhomogeneities they see in their data they have looked at. General nature in blind worlds will be known but the specifics will not be known. This will ensure an even playing field.

CM: Will clean worlds be available for other purposes? Interpolation technique training.

MM: At least three cases (open worlds) we will provide. Consensus was to try to replicate entire PoR.

JL: Any sense for number of groups?

MM: Two for sure. A third may well and we hope more than that. The worlds should be out there at end of '13.

IV. ISTI Implementation Plan update and planned activities associated with the Databank (Thorne, Lawrimore)

WG members are encouraged to review the 2013-2015 Implementation Plan with particular focus on the Databank section (pages 4-9). We are happy to take comments on any part. Long term aims remain to do sub-monthly and multi-elemental. But that is truly long-term. Metadata needs to remain a focus. A short to medium-term goal is the improvement of metadata collections. Also added is a goal to create a Parallel measurements database. This will help in understanding the physical understanding of breaks. For example the Wallops test bed, UCAR, CRN, Camborne.

SW: Metadata is hugely important. Move forward to standard format including station histories is important. UCAR has four metadata cases. Does this refer to only temperature?

PT: Parallel measurements should be more than just temperature but limit to the surface.

AKT: Are parallel measurements designed to be separate?

JL: Should be an adjunct to the databank so might be similar overall format but will need some bespoke extras not in the databank standard formats.

AKT: Ideally want metadata for all stations.

JL: Side-by-side measurements will be dealt with somewhat differently.

Albert: WMO requires 2-years. Have built some holdings from such studies in Europe.

JL: Main thing is collating, discovering, and making available.

PT: Metadata and parallel measurements are where action items need most critical assessment and possible revision.

JL went through the actions table on pages 8-9. Would be nice to have some examples of both metadata and parallel measurements holdings to start to define the formats as called for in the current actions in the IP redraft.

JL presented the update protocol for the Databank

Near real-time: Update from a small number of sources each month.

Currently GHCN-D and preliminary CLIMAT.

ECA-M – not currently updated. Unsure how many are unique – already getting a lot of European stations via GHCN-D. Also getting CLIMAT for areas such as Africa.

ECA-D being added to GHCN-D monthly.

Annually: Will perform POR updates and add new sources (likely after start of each calendar year; Feb or March)

Semi-annually: A POR update but no new sources added (approximately 6 months after the annual update (Aug or Sep))

SW: How are we going to react to user feedback. Adding to blacklist?

JL: Depends upon severity of issue. Preference would be to do annually if minor. Might want to document what the procedure is so an end-user can understand when it is going to be changed and how.

SW: Would be helpful to make sure it is well documented and open and transparent.

V. Data collection activities; Gaps in coverage and areas of focus at the Regional level (All)

Jared discussed where we have good data holdings and where we have gaps remaining. Africa and South America remain a requirement for focus. Ongoing work.

Jay asked Waldenio whether there were any additional Brazilian sources. Waldenio said that he is still actively pursuing this and will advise by email.

Waldenio to assess Brazilian data and advise Jay.

CM: Next year's Earthtemp network has an African focus. Looking to invite African participants.

PT to follow up.

Dave Lister: CLARIS project data (<http://www.claris-eu.org/>) will be publicly available soon. Looking at the data in many Brazilian 'long' series there is a gap between 1979 and 1997.

JL: When will it be available.

DL: Potentially now. Quite difficult to download everything. Will advise.

DL to make inquiries regarding whether anyone can download data now.

AKT: Coming back to African data working with ACMAD – Albert Mhanda will be working with KNMI on a regional daily data holding in West Africa.

VI. Summarize Activities and Next Steps for coming months

ACTION: ISTI IP comments to Peter / Jay. By Friday.

ACTION: Working Group members to submit metadata and/or parallel measurements now so we can see what exists and inform the data formats.

ACTION: JR to follow up with WG members when review comments are returned.

ACTION: JR to undertake stage 3 version 1 release.

ACTION: Waldenio to assess Brazilian data and advise Jay on potential for gaining additional period of record data.

ACTION: DL to make inquiries regarding whether anyone can download data from CLARIS now.

ACTION: PT to follow up on location and focus of EarthTemp meeting with eye toward gathering additional source data from Africa.

VII. Other Business

Will aim for next call around first half of November. Tentatively November 5th.