

International Surface Temperature Initiative Steering Committee call 7/11/12

Present on call: Peter Thorne, Kate Willett (KW), Richard Chandler (REC), Antonio Possolo, Michael de Podesta, Jay Lawrimore, Xialoan Wang, Blair Trewin, Akiyo Yatagai, Chris Merchant, Albert Mhanda

Apologies in advance: Rob Allan (on travel)

1. Review of actions from last call

PT: ACTION next call to discuss the five issues raised in the progress report to which steering committee members were assigned to work on.

See item #3 below

ACTION: JL to send BT the data certificate of appreciation

JL sent it. However, no additional data was obtained from Pacific Island countries as a result of the May New Caledonia workshop.

ACTION: REC to follow up with steering committee and databank WG.

See item #4 Significance big data issue discussion.

2. Reviews of progress

2a. Databank - Jay

Good progress continues to be made although the original goal of releasing version 1 of the Databank has slipped past April. We are now to the point of refining the merging methodology and conducting ensemble runs with various methodologies based on variations of source priority and thresholds for metadata and data comparisons. To aid us in this effort three people on the Databank WG have volunteered to provide their own personal source priorities and threshold preferences for merging. These will be combined with others we have devised in-house, all evaluated, and a single best merged dataset released along with 4 or 5 other reasonable variants.

KW: Met Andreas Becker last week. Change in WMO identifier. He can provide a list of such id changes for Germany.

JL: We aren't using the identifiers. But list would still be useful.

KW: What is the latest estimate of release date?

JL: No firm date yet. September is the earliest. Need to do internal NCDC process once happy.

BT: Quite a few countries seem to be putting their own data online. Russia and Sweden just did this.

JL: Yes, Russia have given us around 400-500 stations.

PT: Please send us any tips as to such sources.

JL: Need to find some Brazilian data w/50 years that we have been told about.

[papers are to be discussed under item #4]

2b. Benchmarking - Kate

This has really slowed up (IPCC deadlines, SotC) but Robert and I are still working away at the methodology behind the clean analog-known-world data. Methods so far:

- use climate model gridbox monthly mean standardised anomalies
- take all stations within that gridbox - monthly mean standardised anomalies
- get the station anomns minus the gridbox mean anomns calculated from the stations to get the residual station noise

- smooth (12 point filter) to remove low frequency variability (most likely due to ENSO/other modes/non-climate changes) and create 'clean' residuals - assume that station-gridbox mean should result in a relatively constant offset in a perfect world (?)
- build an AR (1) model fit to the residuals (test that this is a good fit) - take into account spatial correlations too
- reproduce a synthetic AR (1) time series of residuals using the derived model parameters - test match up (autocorrelation) - repeat if not good enough
- add synthetic AR (1) residuals for each 'station' to the GCM gridbox mean standardised anomalies and put back in the actual station means/variance.
- should result in 'real' feeling data clean from non-climate features?

PROBLEMS: missing data, large inhomogeneities, single station gridboxes, removal of any station variability relationship with background low frequency variability such as ENSO/other modes - annual cycle variance should still be there.

More complicated than I had first envisaged.

Started drafting two papers: concept paper and methodology paper.

ACTION: KW to take a call of the benchmarking group in August

REC: What is the motivation for removing things like ENSO modes from the data?

KW: They will be in the model data. Don't want to double-count.

BT: ENSO strongly affects autocorrelation in certain regions

KW: Agreed. This is part of the big headache here in making things realistic.

2c. General.

Matt Menne attended GCOS AOPC meeting in Geneva on our behalf. The Initiative progress was acknowledged. GCOS did not foresee taking on a surface reference network oversight capacity at this time.

Peter Thorne attended a meeting on frontiers in detection and attribution in Banff and highlighted the initiative in his talk and also gave a slightly updated version of the generic overview poster at the meeting. The poster had considerable interest.

Kate Willett attended the WCRP 4th International Reanalyses Conference and presented a poster on the ISTI databank. There was a lot of interest in the databank:

- when will it include subdaily data which are more relevant to reanalyses?
- data provided by Reanalyses creators from their collections?
- data rescue work funded specifically by reanalyses grants e.g., ERA-CLIM.
- where do reanalyses sit within the framework? Can reanalyses background fields be somehow attached as metadata to specific stations or stored/linked somewhere useful? Use of QC'd/Homogenised data as 'gold standard' data or validation data?

Michael de Podesta visited Met Office (Mike Molyneux) in Exeter to discuss issues around temperature and humidity calibration issues. Spoke with Robert Dunn and Colin Morice about relevance of instrumental uncertainty on final data products.

3. Review against within year objectives

As noted on the last call we wished to assess how we were doing against the set of priority actions noted on our progress report. This is copy and paste below from that report.

Data portal working group creation – Kate Willett, Michael de Podesta, Jayashree Revadekar

Getting groups engaged in product creation – Richard Chandler, Antonio Possolo, Xiaolan Wang

Crowdsourcing – Peter Thorne, Jay Lawrimore, Rob Allan, Albert Mhanda

Over-arching Initiative funding – Blair Trewin, Greg Strouse, Akiyo Yatagai

Data portal working group creation

Nil known to date

Product creation

Both NIST and University of Washington have made some progress to date. Some interest from groups in France and Norway expressed. Not sure how likely these are to proceed. We need to be ambassadors for groups to do this. The more groups the better.

REC will initiate an off-line discussion to move this forwards.

Crowdsourcing

Michael de Podesta, Jay and Peter have been trying to get together a flyer on this and it has started to be used.

NCDC now has approval from Department of Commerce legal counsel to use volunteers to inventory and image paper records in its Foreign Data Library. The next step in this arduous process is to work with NOAA's human resources department to further explain the specifics of the project and get further approval.

Initial tranche of images will shortly be posted as stage 0 and advertised through the blog along with instructions as to how to digitize (crowdsourcing 0.0). Many more images and inventorying remain to be done.

Funding

Nil known to date

4. Literature status

Metrologia piece progress? - Greg / Michael

MdP: Still getting buried by paid work deliverables.

ITS9 invited 10-page proceedings manuscript got accepted as is

Significance big data issue piece? This didn't happen by the original deadline of end of May. REC contacted the editor to apologise and was given an extension. This will now not appear on the special issue on "big data", but the editor is interested in a piece on climate data nonetheless. He writes "Yes, of course the piece need not now be so focussed on big data - but I think it is still worth mentioning that aspect of it somewhere, partly to give continuity to the magazine - just cos the last issue was all about big data that doesn't mean we ignore big data in all subsequent issues - and partly cos it seems an integral part of the subject - I assume you cant do climate research without a reasonably large amount of data ... I'll nudge you again in about six weeks time, to ruin the summer break". REC summer break is already ruined however, and unlikely to be able to produce anything for the foreseeable future. It is worth keeping this ball in play, however, because Significance has a wide circulation (including high schools etc., in addition to the statistical community).

Planned papers on databank - Jay

Two papers are in the works. A brief overview paper of 1500 words or less will be submitted to EOS to advertise availability broadly. A second paper with much more technical detail on the development of the Databank and merging methodology is being prepared by Jared Rennie.

KW - 2 Benchmarking papers: Concepts, Methods

5. Inaugural Earthtemp meeting, Edinburgh, June 2012

Several active participants in the initiative attended this meeting. The meeting was hosted by Edinburgh University and organized by Chris Merchant's NERC facilitated project. About 60 participants from land, marine, cryosphere and in-situ and satellite communities attended. There were a lot of breakout group discussions and activities. There are a number of potential visiting scientist support positions to support up to a month long exchange in the coming year. Exchanges need not include UK researchers but this is preferred, and the project furthered by the exchange needs a UK angle (because the scheme is presently UK funded). What opportunities do we foresee? It was a useful way to start to think about how we might use the satellite land surface temperatures (skin temperatures)

There will be a paper of recommendations from the workshop, the headlines being:

- Build understanding of the relationships of different types of surface temperatures, where presently inadequate
- Demonstrate new underpinning applications of surface temperature data in meteorology and climate
- Make surface temperature data easier to obtain and exploit, for a wide constituency of users
- Consistently provide realistic uncertainty information with surface temperature data
- Undertake large-scale systematic intercomparisons of surface temperature data and their uncertainties
- Communicate differences and complementarities of different types of surface temperature data in readily understood terms
- Rescue, curate and make available valuable surface temperature data that are presently inaccessible
- Maintain and/or develop observing systems for surface temperature data
- Build capacities to accelerate progress in the accuracy and usability of surface temperature data

KW: Restrictions on scope of visiting scientists?

CM: Only really must be UK+int'l angle to it.

6. Possible workshop support submission

There is a possibility to apply for a research grant to cover costs of a meeting in Banff which may logically coincide with the end of the first benchmarking cycle. Is this something which we should pursue at this stage?

The Banff International Research Station for Mathematical Innovation and Discovery (BIRS) is now accepting proposals for its 2014 program. The Station provides an environment for creative interaction and the exchange of ideas, knowledge, and methods within the mathematical, statistical, and computing sciences, and with related disciplines and industrial sectors.

JL: Make a proposal. We do need to get together.

REC: Pretty mathematical. Success would be increased by getting some maths / stat groups on board. Washington?

PT: Probably need Stats / NIST / maths. Generic mathematical and statistical issues around benchmarking and homogenization.

General agreement to take this forwards. Will form an organizing committee and take forwards.

ACTION: PT to lead a proposal in August once IPCC is submitted.

7. AOB

Suggest next call just prior to October when the WG progress reports are due. Agreed to 9/19 @8EDT