

Benchmarking Working Group Online Minutes #15
Monday 28th October 1pm GMT

#-----

Attending: Kate Willett (KW), Peter Thorne (PT), Ian Jolliffe(IJ), Rachel Warren (RW), Victor Venema (VV), Enric Aguilar (EA), Lucie Vincent (LV), Robert Lund (RL), Stefan Bronnimann (SB)

Apologies: Thordis Thorarinsdottir (TT), Renate Auchman (RA), Matt Menne (MM)

Actions from last meeting:

ACTION: Kate to push for research proposals on all aspects of the benchmark project in the UK statistics/climate community

PT: Note that a proposal to a SAMSI summer workshop in 2014 is approved pending logistics confirmation.

ACTION: Kate is to complete the existing concepts paper draft and circulate ASAP – August at the latest.

DONE

ACTION: Claude and Lucie to conclude findings on seasonal cycle in changepoints in the next month – how frequent are seasonally varying changepoints? What are their common characteristics.

LV/CW - may be able to provide something on seasonal cycles in changepoints later - see timescales listed below.

ACTION: Victor to contact the homogenisation list about characterising seasonal cycles in changepoints and their frequency.

DONE.

There are relationships with insolation, soil temperature, vegetation and precipitation that cause seasonal cycles. For station moves almost any climatological phenomenon can cause a seasonal cycle in the inhomogeneity: seasonal circulation patterns, orography, distance from coast, land use, etc. In case of station moves almost any change can lead to seasonal cycles. No new quantitative information was given. Found to occur regularly so should be included.

ACTION: Kate to contact Blair Trewin about size of changepoints (and character) in the tropics.

DONE:

Very strong seasonality found in station correlation decay distances – much smaller in the wet/monsoon region for Northern Australia. Vegetation greenness also plays a big role. A change from non-irrigated to irrigated land can result in no change in the wet season but changes > 1degree C in the dry season. Similarly, moves from town (watered lawns) to airports (non-watered lawns) can have large changes in the dry season. Moves from town to agricultural region can be +1deg in summer and -1deg in winter – so no change annually.

May expect largest changes in Tmax in the Tropics due to any changes to insolation – shelter changes, station moves. Largest changes in Tmin due to station moves with topographical changes.

Nationwide conversion from imperial to metric September 1972 but no clear changes found. Pre-1963 all stations staffed 24hrs (midnight to midnight recording), post 1963 staffing 9-5 only (9am to 9am recording) = modest change in Tmin ~ 0.1 deg.

ACTION: get excel file of known changes into nice format to share.

KW: posted current version on website – anyone can update as necessary. I will keep a static scrape every so often to ensure it does not get corrupted.
<https://docs.google.com/spreadsheets/ccc?key=0Al6ocsUAaINSdHpTREJzVkRZUTdfVjNPRlh0Q1V3WUE&usp=sharing#gid=0>

KW: Email link to group and ask VV to email around homogenisation list

ACTION: Kate to pass R code for Team Creation to Enric to look at and play with – double check it does what we think it does

KW to email R Clean World code to Enric ASAP as now is a good time.

ACTION: We need standard deviations on the table – overview_error_world.xls to define the distribution.

We had discussed making distributions that were not strictly normal in Asheville workshop.

KW: Find latest version of table of shift types that was edited in the workshop. Call needed on this to finalise beginning with a summary of where we got to in the workshop.

RL: need more than St Dev - can pass on to Statts to help out here - normal/skewed/fat tailed.

PT: Defining a distribution that includes the missing middle.

ACTION: Kate keep in touch with Douglas Maraun about how assessment and reporting of assessment is being done for VALUE as they are planning on an online automatic assessment.

KW: Continue to keep track of VALUE progress.

ACTION: Kate to start an ISTI glossary document in googledocs and invite all ISTI members to contribute.

DONE - but not circulated

KW: Posted a working version on website – anyone can update as necessary. I will keep a static scrape every so often to ensure it does not get corrupted:

<https://docs.google.com/document/d/1xltD6yeQTxqwVnbfX-ZwUsh7hKJu1HqJVEf-OKstS4Y/edit?usp=sharing>

KW: to email around list and possibly ISTI main list

ACTION: Ian to think about use of ANOVA or other methods for comparing the variance between two time series.

*IJ: Suggested on a call and not thought through. ANOVA decomposes variances but compares *means*. Thinking about 'other methods' to compare variances still to be done.*

KW: Lets have a focussed call on methods of comparing assessment statistics over regions where we can summarise where we got to in the workshop.

VV: Needs definition of reference period

RL: Easy to compare the variance of two zero-mean time series and for autocovariances.

IJ: Trying to compare specific features like the seasonal cycle will depend on how they are created in the first place.

VV: Autocorrelations are of interest - hope that what we change will be realistic.

ACTION: Discuss how best to deal with hits/misses/false alarm rates in future conference call.

IJ: Contingency table document put together a while back. Discuss and circulate for a future call.

ACTION: Victor to put some numbers on how big a window we should allow surrounding changepoint location – suggestion that the window should be larger for smaller changepoints as these are harder to locate accurately.

UPDATE: Victor is no longer sure whether such an equation makes sense as it depends on the SNR and thus on the reference, which may be computed in many ways. Maybe first discuss this in Team Validation.

LV: Keep it simple - a flat window size before and after changepoint. 10 years?

EA:VV: suggest too large for Europe

VV: suggests that window size perhaps should depend on location

LV: Variability so large in Canada so we need very large windows

PT: How about a penalty function that relaxes away from 1 year to be zero skill at n-years away? VV: could be a function of signal to noise ratio? More complex though.

IJ: Keep it simple first time round and get more complex next time round.

VV: Another topic for a call.

KW: Link discussion of acceptable window of detection in with focussed call on comparing variances of different climate features – Team Validation call.

ACTION: Claude to look at PHA results to try and characterise some of the gradual changepoints in particular the frequency of the slopes, their size, their duration and if anything can be said about their character – are they linear or stepped or other?

KW: To chase Claude –

CW: I do not think we can say when an abrupt change within a gradual change /local trend can happen nor the characteristics of a local trend during which it may happen. Depending upon which random, or multi-random events that may occur to cause the change, a local trend may start, stop or change intensity (slope) or all of the above.

CW: Gradual changes may as well be assumed to be Gaussian in distribution – there may be evidence that there are slightly more "cooling" than "warming" local trends!

VV: Do we know how this varies by climate region?

ACTION: Kate to get GCM data in workable format to Robert

KW: Get data to Robert - next week.

ACTION: Kate and Robert are to work on their respective methods utilising GCM data and the VAR(?) method. This will result in a publication describing the methodology. We aim to finalise this work in August.

KW/RL: to sort Clean Worlds out and report back.

ACTION: Follow on teleconferences will focus on completing the distribution specifications within the overview_error_world.xls document and the building the probability model to allocate station changepoints. This work is planned for completion by October 2013. Release of the benchmarks is scheduled for November 2013.

KW: Team Corruption changepoint distributions topic for the next call, summarise where we are with this.

ACTION: Team Validation will continue to develop the specific assessment tests and work these into a software package that can be easily implemented. This work is hoped to be completed by December 2013, but there is more time available as assessment will take place at least 1 year after benchmark release.

Actions from this meeting:

KW: SORT OUT A STABLE SET OF PHONE NUMBERS THAT WORK

DONE Science Admin have some but awkwardly available

KW: Draft Progress Report and circulate

DONE

KW: GCM data to Robert and complete CleanWorld code
DONE

LV/CW - may be able to provide something on seasonal cycles in changepoints later
- see timescales listed below.

KW: Email link
(<https://docs.google.com/spreadsheet/ccc?key=0A16ocsUAaINSdHpTREJzVkrZUTdfVjNPRih0Q1V3WUE&usp=sharing#gid=0>
) to group and ask VV to email around homogenisation list
DONE

KW to email R Clean World code to Enric ASAP as now is a good time.

KW: Find latest version of table of shift types that was edited in the workshop. Call needed on this to finalise beginning with a summary of where we got to in the workshop.

KW: Continue to keep track of VALUE progress.

KW: to email link (<https://docs.google.com/document/d/1xltD6yeQTxqwVnbfx-ZwUsh7hKJu1HqjVEf-OKstS4Y/edit?usp=sharing>
) around list and possibly ISTI main list
DONE

KW: Lets have a focussed call on methods of comparing assessment statistics over regions where we can summarise where we got to in the workshop.

IJ: Contingency table document put together a while back. Discuss and circulate for a future call.

KW: Link discussion of acceptable window of detection in with focussed call on comparing variances of different climate features – Team Validation call.

KW/RL: to sort Clean Worlds out and report back.

KW: Team Corruption changepoint distributions topic for the next call, summarise where we are with this.

KW: Recirculate concepts paper by end of this week - aim for GIMDS

KW; Send around Doodle poll for that week.
DONE

KW: Plan for next meeting:
- summarise progress from workshop
- find and circulate latest version of distributions - discuss shape and size.
- Concepts paper - ready to submit?
- Metadata - how to generate and how good/bad?

#-----

Agenda:

1) ISTI Update (PT)

The paper describing the databank has been formally accepted and is in the process of page proofs being returned. This acceptance WAS to trigger the formal version 1 release. The release was to go live Oct 1st but the US Gov't shutdown saw to that. Then it was due to go live today. However, NCDC senior management have decided to hold pending yet further internal reviews before a formal version release. The diplomatic term here is 'disappointing'. The current plan is to put out a 'release candidate' when the paper goes out which for all intents and purposes is version 1. I would strongly advocate your using this release candidate version as the basis for all your work otherwise there is just unnecessary further delay of little value. I have local permission to host a frozen version of the release on servers at my work and we could also, theoretically hold the benchmarks. Given the difference in ethos this may be a good route forwards vis-a-vis hosting and promoting the benchmarks. Alternatively the hadobs server may have room if its choke issues have ever been resolved Kate?

KW: May well be an option as we're pushing everything over to CEMS at BADC which is much bigger. Time frames for this are not in our favour though ~3years.

The IP refresh was published (and is now immediately dated due to issues beyond reasonable control)

We need from the Working Group a progress report. Does not need to be long. This forms the basis for a longer report on the initiative as a whole to the 'sponsors'

KW: Draft Progress Report and circulate

As noted above we are tentatively good to go on a workshop hosted by SAMSI next summer on advancing homogenization techniques and the use of the databank and the benchmarks. Without wishing to put too much pressure on this does now yield a reasonably hard deadline of May 2014 to at least have the open corrupted worlds available and ideally some rudimentary means of assessing the output ...! Further advancement beyond that minimum set will obviously increase the value :-)

I will give talks at a Norwegian conference early next month (11th/12th) and also at the 2014 AMS meeting on the Initiative. The initiative also featured in my talk to Royal Society in early October which may get streamed online in the fullness of time (hopefully not, clearly ...)

Over the weekend a letter went out from Michel Jarraud (DG WMO) to all PRs requesting support for ISTI activities. Includes building parallel measurements database.

2) Benchmarking Working Group Update (KW)

Progress stalled since workshop due to holidays and Kate being off work a lot

Concepts paper draft circulated and edits compiled - apologies for sending the wrong version first time round.

Invited ISTI/Benchmarking talk at Reading University Meteorological Department - Monday 4th November (KW)

3) Actions pending (KW)

Reminder of who needs to do what - still relevant?

See Actions list above

4) Concepts Paper (KW)

VV: Which papers are in the planning and who will write them?

KW: I would like to see a paper for each team if possible?

Is it generally ok in scope?

VV: Quite general –

RW: Too much stats may make it difficult to submit

LV: Important to establish where the paper is going.

PT: GIMDS most appropriate. Needs to be only what and how – not who. What are these benchmarks and what is their purpose?

LV: Need to convince scientific community the importance of what we're trying to do. Submit to: Environmental Research Letters? Geosciences Data Journal? Both open access but may not be appropriate Journal of Atmospheric and Oceanic Technology, BAMS? – not open access.

VV: Be more specific about uncertainty that our benchmarking work can help to quantify – 30-40% in global mean trend actually. At present we say that the global mean will be changed very little.

PT: BAMS is open access. But I'm not sure the paper is within the BAMS remit per se - a bit too niche for their generalist audience aims. EGU journals?

<http://www.geoscientific-instrumentation-methods-and-data-systems.net/home.html>

This is where the Earthtemp whitepaper was submitted (and accepted) which constituted a similar charge and on a similar topic so I would see a reasonable chance of success.

KW: Recirculate concepts paper by end of this week - aim for GIMDS

VV: Description of the benchmarking cycle could come more in the beginning. PT:

PT: Use cycle diagram.

5) Updated plan of action:

Original time line:

August 2013 - produce clean worlds

November 2013 - produce error worlds

December 2013 - Formalise a system of assessment with specifications of what data/statistics must be returned

By Christmas? [Dear Santa ...? ;-)]

Kate and Robert to sort out clean worlds

RL: Kate - get me some clean GCM and the stations in ASCII format

Call to complete the distribution specifications in the overview_error_world.xls document

Finalise statistical methods for allocating various changepoints to the Error Worlds

By March 2014 - must be available by summer 2014?

Release of Benchmarks V1?

PT: Need to have at least open worlds done by June. Closed ones not essential for workshop.

6) AOB:

i) Next Meeting:

11th-15th November (not 14th)?

PT: I am on conference 11th and 12th if my attendance were required avoid these two days.

RW: Away that week?

KW; Send around Doodle poll for that week.

KW: Plan for next meeting:

- summarise progress from workshop
- find and circulate latest version of distributions - discuss shape and size.
- Concepts paper - ready to submit?
- Metadata - how to generate and how good/bad?