Benchmarking Working Group Call #11
Tuesday 18th June 4pm BST (GMT+1), 3pm GMT, 11am EST

Attending: Kate Willett (KW), Victor Venema (VV), Enric Aguilar (EA), Lucie Vincent (LV), Matt Menne (MM), Ian Jolliffe (IJ), Claude Williams (CW), Peter Thorne (PT), Stefan Bronnimann (SB)

Not attending: Rachel Warren (RW)

ACTIONS FROM PREVIOUS MEETING:
KW: to see about guest speakers at next call - Manola (Spain), Andrea (Meteomet project update)? - DONE
LV: check Kate's write up of her presentation on the minutes - DONE
Everyone: please fill this list with contacts or contact them if you know them - COPIED INTO NOTES AT END OF THESE MINUTES FOR MORE ADDITIONS
MM/LV/EA and ALL INTERESTED: read Victor's error world document.
MM/CW: please provide details for USA - DONE

ACTIONS FROM THIS MEETING:
EA to check my scribbled notes above.
KW to add link to figure showing distribution for detected inhomogeneities plus estimated 'missing middle' - hadobs currently offline :(
VV to recirculate updated table - DONE
ALL think about seasonal cycle
KW/MM add in specific part for seasonal cycle discussion
KW/MM: find time for longer chats with Victor on Tuesday - the day of corruption!
KW: can put regional summary up on website/blog
MM/KW workshop schedule up - what we're aiming for – on website/blog
KW/VV overview of vague plan for worlds - 30000ft overview - Victor and Kate over a pint at IMSC? Could be Kate's presentation or a sub-set of it? – on website/blog
KW: investigate changing times of future calls - do a doodle

AGENDA:
1) Guest speaker - Enric Aguilar - Known inhomogeneities for Spanish data and South America
EA: Africa rather than SA. Guinea Conackry
KW: Ooops - still very useful though.
SDATS : 57 breaks + 20 known breaks due to screen effects over 22 stations - a little larger frequency than other networks 1850s begin for some
Conservative approach, only treated those breaks that were very obvious (conservative) - expect the frequency to increase in future homogenisation - using HOMER.
56% of breaks were in the metadata and due to relocation
35% of breaks were of an unknown nature
a couple of stations had more than 4 breaks, most had 2 breaks
In the analysis of another dataset, the Catalan Experiment (used also in HOME benchmark) - 30 stations from Catalonia, Southern France and Andorra - found

approx. 1 break every 16 years - Therefore estimate of 15-20 year HSP is an ok assumption.
Going back to SDATS, Tmax breaks were bell shaped, Tmin breaks had more negative than positive - think this is due to move from urban to rural.
Most obvious impacts are from screen effects. Cold bias on Tmin, warm bias on Tmax.
LV: Annual or monthly amounts - monthly would be a little bigger than annual?
VV: Define screen effects - EA: change in shelter type, 6 types of screen. NW and SE corners had >10yr experiments testing late 19thC Montsouris Screens. Temps in late 19th C - Tmax had a very large warm bias.
open Montsouris to closed Stevenson - 1893-1916 = Almost all changes are warm bias in Tmax (bigger), cold bias in Tmin (smaller) - annual cycle in Tmax, not so much in Tmin, annual bias in mean is positive.
We are studying a network of around 80 stations from the Agencia Estatal de Meteorologia (AEMET) and the Servei Meteorològic de Catalunya (SMC) As first results. Automatic weather stations - diffs are larger in Tmax and than in Tmin, larger annual cycle in Tmax. Very dependent on sensor type and especially where the sensor is located. 0.1 degs diff when thermometer and aws sensor are inside a stevenson screen. When aws sensor outside stevenson screen (small plated screens) differences are much larger.
MM: Non-aspirated? EA: yes
EA: Practically no metadata for Spanish station, except for those mantained by Servei Meteorològic de Catalunya
MM: Are the automated weather station screens inferior?
EA: Not inferior - not sure which one is superior/inferior - but differences are larger for change to automated station screens.
MM: In USA - find negative bias in automated weather stations (aspirated) - the opposite to that found in Spain.
EA: Conclude that it is related to exposure rather than the sensor itself. Uncertainty in automatic sensor is ~0.2deg.

Guinea Conackry, Africa:
12 stations - quite a different picture in size (much larger - 4, 6 degrees in the monthly) than in Spain.
Detecting the breaks but not adjusting them well because of missing data and station sparsity. Frequency is probably the same as in other parts of the world but poorer quality data, more missing data and station sparsity means that adjustments are not so good. This means that size may be miss-calculated - so we think that size is actually smaller than that found.
~1 break every 27 years.
Problem of two stations being merged as one, badly - is this a QC problem or a homogenisation problem - should we include these?

South America:
Paper ws written and workshops were held - so request made to provide some statistics for South America - fingers crossed :-)
EA: Update: got reply from Maria, she's preparing the dataset, so I will be able to provide some statistics for SA, hopefully before the WS. A clue: again only obvious breaks have been corrected.

ACTION: EA to check my scribbled notes above.

2) Team Corruption (Victor Venema)

VV: see the table circulated around overview_error_worlds.xls

LV: Is the bias amount for the annual data?
VV: yes - on the annual but this is the overall bias per century
VV: breaks drawn from gaussian distribution with st dev of ~0.7
VV: Value of 0.2 comes from Parker paper and GHCN
KW: This is the global average - so will vary by station and network
MM: Agree that 0.2 is consistent with what we know at this point in time, 0.4 for equator is a guess
VV: But actually for one world we want 'best guess for the west everywhere' so actually equator will be 0.2 here.
SB: Each value should be network dependent

LV: need units on the table

VV: local trend values would be varied around the global average - some negative and some positive.
EA: Spain - out of 22 stations detected local trends in 5 - ~20%. All of these stations are over 100yrs long
MM: USA - probably higher than 10% for stations with a local trend.
VV: And it will be higher than what we have detected.
LV: Spain local trends all because of urban effect?
EA: All urban effect - average 40-50 years, in some cases 80 years
CW: Can be trees growing, all sorts of things
CW: Saturation of urban warming though - Central Park
EA: yes - Madrid saturated in 1960s
PT: See http://onlinelibrary.wiley.com/doi/10.1002/wea.432/pdf for a paper by Phil Jones and David Lister for London effect (this is Open Access)
VV: easy 10%, hard 30%, others 20 % - agreed global value which will vary by network e.g. more for China
CW: distribution of warming rate? VV: yes -2 to +5 with a mean of 2 deg.

EA: What is the overall length of the series?
MM: For the most part our benchmarks will be the same length as the ISTI stations but the option is there to produce longer/spatially complete.

LV: How many stations will be created by Team Creation?
KW: 30000 - just like the ISTI databank - easy! :)
LV: Peterson & Vose GHCN overview - no. stations reporting from 1850 - they only have ~6000 stations reporting in 1970s.
PT: www.surfacetemperatures.org/databank - this website links to the ISTI databank comparisons with the GHCN databank. Top right hand panel - red line is what we have and black is +/- small changes what is in Peterson and Vose.

EA: Assume no errors in the station location coordinates?
VV: Yes

PT: Some processing has been done on ISTI databank to cope with some of these gross errors and some merging. See http://surfacetemperatures.blogspot.com/2013/01/blacklisting.html

VV: length of trends varied around a log-normal distribution. Any experience from other countries?
MM: Reasonable from a USA perspective
LV: Sounds good
KW: Keep length average same for all and allow variation by network - can do more specific analysis on this station by station.
VV: length ~ 20years for easy

VV: Ranges of the values are in the document - so not on the table.

VV: hard world - local trend should be smaller - harder to detect - so switch easy and hard local trend magnitude.
MM: 2 degrees for global average local trend seems large
LV: I agree, I would put 1 degree
MM: agree this would be better - there are larger local trends but not sure how prolific these are.
VV: Propose 1 varying between -2 and 4 degrees.
MM: Agree
VV: Therefore hard should be 0.5 or 0? PT/MM - go for 0.5. Easy - 2.5 to 3?
VV: What about the hard distribution of random
MM: have a long tail on the small side? Gaussian distribution is best of our knowledge
PT: merge two Gaussian populations - one large pop with small sigma and one small pop with large sigma.
VV: What about Peter's double distribution idea?
PT: We want some large breaks AND small breaks - even in the hard world. 80% from small sigma distribution and 20% from a large sigma distribution?
PT: Essentially building a non-normal distribution
KW: could justify using figure from my humidity paper which fitted Gaussian to detected adjustments but then had to apply a fudge to fit the wide tails and the narrow middle
ACTION: KW to add link to figure showing distribution for detected inhomogeneities plus estimated 'missing middle' - hadobs currently offline :(

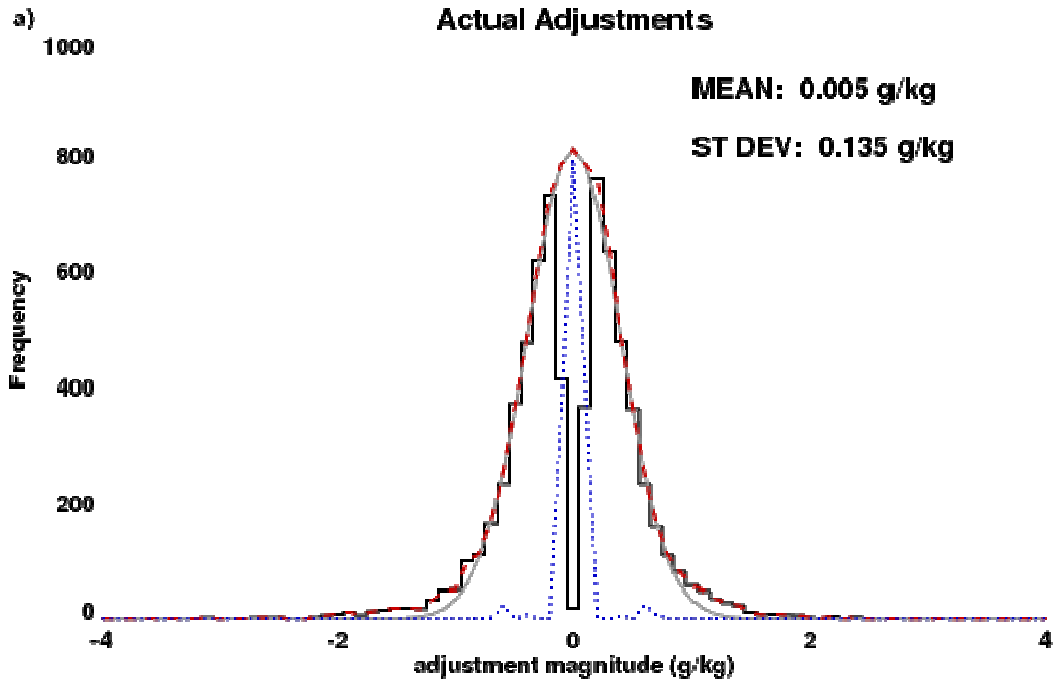

Fig. 3 Summary of adjustments applied to HadISDH during the pairwise homogenisation process. Figure a) shows the actual adjustments in black (stepped). The best-fit Gaussian is shown in grey. The merged Gaussian plus larger actual distribution points 'best-fit' is shown in dashed red. The difference between the merged 'best-fit' and the actual adjustments is shown in dotted blue with the mean and standard deviation of the difference.

PT: From a measurement science perspective there must be at least some possibility of frequent small breaks - people likely don't set out to introduce large inhomogeneities. So we may effectively have a tight middle of small breaks and real distribution may be non-normal.

EA: Will some stations have no breaks?
VV: Some will have none yes - through random distributions
EA: 1 in 30 stations we don't detect breaks - may just be that we don't detect them. Realistic to have a few stations without breaks.
MM: Haven't found a station without a break in the USA yet.
CW: Very short stations don't have breaks.
KW: open world #3 will have no breaks
VV: could have this in the exotic world - a few networks or countries without breaks
EA: stations longer than 50yrs tend to have breaks, shorter - have found some stations without breaks but maybe that they're just not detected.
VV: Keep it simple so no explicit stations without breaks, let it happen by chance from picking out distribution, can apply to the exotic or open worlds

VV: easy world - 0.9 (west) now 1.2 (outside the west) - for simplicity.

VV: Seasonal cycle - Lucie had much larger seasonal cycles in the inhomogeneities than the 0.35 estimated here.
VV: all worlds apart from one open world will have seasonal cycle in breaks.

EA: Will equatorial breaks have seasonal cycle? - more likely to follow dry/wet season rather than warm/cool.
VV: Should be smaller if actual seasonal cycle is smaller - make it a function of the seasonal cycle itself?
CW: Phase of cycle can change too.
MM: Don't see consistent seasonal cycle expressions. Maybe some noise needs to be added on top?
VV: Could apply only to a fraction of stations
ACTION: ALL think about seasonal cycle.

ACTION: VV to recirculate updated table.

Today we have to fix the default errors for mean temperature.
A - Size distribution of random (independent and unbiased) breaks.
B - Size of biased (ones that influence the region or global mean trends) breaks.
VV: Combine this category with E?)
C - Noise or walk
D - Frequency of biased breaks.
EA: I think a reasonable estimate for this is 2 or 3 every 100 years. For example, many networks change screens in the early 1900s and changed to AWS in the late 1900s or early 2000s.
E - Size, correlation and period of spatially clustered breaks.
F - Number of breaks.
G - Seasonal cycle of breaks.
H - Gradual inhomogeneities, local trends.

3) Workshop update
ACTION: KW/MM add in specific part for seasonal cycle discussion
MM: Can we have Victor online for longer than the standard call?
VV: timing difficult - in principle free
ACTION KW/MM: find time for longer chats with Victor on Tuesday - the day of corruption!

4) AOB
PT: Can we start putting some info up on the website and / or blog area to show what we are doing to RoW? Matt's document would make a good blogpost for example.
ACTION: KW: can put regional summary up
ACTION: MM/KW workshop schedule up - what we're aiming for
ACTION: KW/VV overview of vague plan for worlds - 30000ft overview - Victor and Kate over a pint at IMSC? Could be Kate's presentation or a sub-set of it?

PT: We should have these calls an hour or two earlier so that its in European working hours. 9.30am on East coast is 14.30 in UK and 15.30 in Europe.
Seems generally ok for everyone - doodle of times?
ACTION: Kate to change times of future calls - do a doodle

5) Next call:
Thursday 20th June 4pm BST/3pm GMT/11am EST - Team Validation

---

NOTES:
KW: Regional Inhomogeneity summaries - so far we have info from:
Spain
Switzerland
Poland
Slovenia
Russia
Australia
Canada
UK (pending)
Italy
Netherlands
Possible contacts?
South American countries - Mathilde Rusticucci?
EA: after and ETCCDI-like workshop we prepared a paper including homogenization of most countries in SA. Maria Skansi (Argentina) was the first author.
EA: now we're preparing another round of homogenization with Bolivia, Ecuador, Peru and Venezuela. Peru has paralell AWS/CON measurements and would be happy to contribute. Clara Rojas is the name of the person to be contacted.
VV: Great contacts.
USA - Matt/Claude? Cooler maxes, warmer mins
MM: Result of the change from Liquid-in-glass to MMTS instrumentation...
Africa
France
Germany
Austria
Czech Republic - Petr Stepanak?
Scandinavia
Nordhom project involves the Scandinavian countries. They're hosting a homogenization internal workshop in November. Erik Engstrom is the project manager.
Other European countries?
New Zealand
China
Japan
South East Asian countries
India
Middle East