Benchmarking Working Group Online Minutes #17

Wednesday 27th November 6pm GMT (10am Pacific Standard time (western Canada?), 1pm Eastern US, 7pm Europe, 5am Australian Eastern time - sorry Lisa!)

Attending: Kate Willett (KW), Victor Venema (VV), Peter Thorne (PT), Enric Aguilar (EA), Ian Jolliffe (IJ), Rachel Warren (RW), Thordis Thorarinsdottir (TT), Lucie Vincent (LV), Robert Dunn (RD), Matt Menne, Colin Gallacher

Apologies: Lisa Alexander (LA), Robert Lund (RL)

Actions from last meeting:
ACTION: CW: To talk to Robert Lund/Colin Gallagher about ways of analysing seasonal cycles between pairs of stations.
NOT DONE

ACTION KW: edit Glossary document with changepoint meaning
DONE

ACTION KW: Chat with Colin Gallagher to see how much he can help here.
NOT DONE

ACTION: Kate to try and draw up a flow chart and iterate with Victor. Some open source drawing program.
DONE see:
https://docs.google.com/drawings/d/1Hn7lDQJivDcKfZpkTyuIBzvYFE9MSyc8IPBNMwdPJCA/edit - an example of a station data object and its attributes
https://docs.google.com/drawings/d/1fb-kjL2J1oG1KmR2c_4hWhlSvV3Ats7xDGXP45aAU7w/edit - a flow chart showing how inhomogeneities might be added?
IJ: I prepared some notes on various distributions with I've just pasted in at the end of the agenda. In them I say I can't see a direct use for Poisson variables. Having seen the flow chart I now accept that Poisson variables can be used but I think it's unnecessary. Rather than choose the number of breakpoints with Poisson, simply generate positions of successive changes using geometric (for discrete time) or exponential (continuous) and the number of changes in the series comes out automatically (as binomial - discrete or Poisson - continuous)
CG: I agree.

ACTION: KW to doodle poll and set up
DONE

Actions from this meeting:
ACTION KW: Get the pad open earlier.

ACTION KW: Utilise the pad for discussion opportunities outside of calls.
ACTION KW; Revamp and circulate - discuss at next call. Finalise by Steering Committee call in January.

ACTION: Assimilate concepts paper comments and circulate next week

ACTION Matt to provide adjustments file from PHA to Peter. Peter to assess structure of changes found.

KW: ACTION - talk to Robert about ensuring cross-correlations are not too poor initially because they are built on unclean data and that they are not eroded unrealistically when errors are added.

ACTION LV to talk with CW? About retrieving seasonal cycle information from the data?
MM: This is probably a lot of work - CW has started it, but we'll probably have to guess this first time round.
MM: Grid os seasonality of time of observation bias from paper, Karl et al. 1986. We can at least provide the geographic/seasonal impact of the time of observation bias (say afternoon to morning relative to local midnight)

Karl, T.R., C.N. Williams Jr., P.J. Young, W.M. Wendland. A Model to Estimate the Time of Observation Bias Associated with Monthly Mean Maximum, Minimum and Mean Temperatures for the United States. Journal of Applied Meteorology, 25, Issue 2, pp.145-160, doi: 10.1175/1520-0450(1986)025<0145:AMTETT>2.0.CO;2, 1986.

EA: For the tropics and non-western countries I have some datasets I can look through but there is no metadata.

LV: What about the parallel stations?
VV: We only have Slovenian data and the COST action has just been rejected.
ACTION LV to send parallel data to Victor.-
https://ourproject.org/moin/projects/parallel

ACTION: Kate to work with IJ/VV? CG on reworking the flow chart.

ACTION: ALL TO EMAIL KATE WORDS THAT SHOULD BE IN THE ISTI GLOSSARY – OR PUT THEM IN YOURSELF:
https://docs.google.com/document/d/1xItD6yeQTxqwVnbfx-ZwUsh7hKJu1HqJVEf-OKstS4Y/edit

#----------

AGENDA

0.1 Please note that the steering committee will like a revised ToR and membership document before a posited January all hands telecon

ACTION KW; Revamp and circulate - discuss at next call. Finalise by Steering Committee call in January.

0.2 Kate, what is your plan for the concepts paper here on in? When do you need co-author comments?
VV: I would still love to work a little on the manuscript, but do not have any time in the next two weeks. Is that acceptable?
PT: I'm guessing even once we are 'signed off' there are still at a minimum UKMO and NCDC internal reviews to grit teeth thru ....
VV: Can you still make changes after such an internal review?
PT: Yes.
VV: Good.
KW: I haven't had time to assimilate all edits from V2 yet. I will hopefully get to this on Thursday. It would be great to get this submitted to internal reviews before Christmas
ACTION: Assimilate comments and circulate next week

1) What is a biased changepoint and a random changepoint and how do we apply them?
Its become apparent that I am confused when it comes to Victor's definition of biased and random changepoints:
http://variable-variability.blogspot.com/2013/11/break-inhomogeneities-random-walk-noise.html - Victor's helpful blog post.

Previously I had thought this meant that some changepoints added were inherently biased in that they always acted in the same direction. Actually I can't think of any types of inhomogeneities that would occur more than once and always in the same direction.
VV: Many will not occur more than once at one station, but they will occur in more than one station, often in all stations of one network and thus bias the network mean climate signal.
It is however plausible that several different types of inhomogeneity may occur that act in the same direction - or at least have their size picked from a distribution with a non-zero mean - one off biased changes could be:
PT: Early changes - room with open window to N (S in SH) facing wall in extra-tropics? Then N/S wall to screen? Okay, these affect very few stations, but still ... I can dig out several papers on this.

- Wild screen to Stevenson screen = cooler generally?

[PT: Mainly tropics only really] VV: Do you have literature on breaks in the tropics? I would be very interested in them. (I thought Wild screens were mainly used in high latitudes.). Maybe mis-understanding here - thatched screens in tropics - was assuming that was what was referred to here.
VV: At least the general move to Stevenson screen generally leads to cooling, but not always, sometimes it also leads to a warming (Parker, 1994). So I would describe that with a normal distribution with a mean of 0.1°C, but also a width of 0.1°C. (Estimated values). As the mean is non-zero, I would call such a process biased. The ensemble

average of all stations making such a transition would also be non-zero and biased. Any individual realisation of the process can have any sign.

- Move from cities to airports = cooler generally?

VV: As an aside, this interesting blog post on the clear climate code blog suggests that this transition was responsible for a trend bias of around 0.1°C. http://clearclimatecode.org/airport-warming/ Does anyone know any scientific articles on this? Would be important, we should not just research warming inhomogeneities (urbanization) that would bias the temperature curve.
PT: I recall early drafts of Lawrimore et al. 2011 alluded to this as a 30s/40s issue that may explain GHCNv2/v3 changes but do not know whether this survived through the final draft and do not have a copy immediately to hand.

- Stevenson screen to small AWS type screens = warmer generally?

PT: Tends to be associated with a change in ventilation efficiency which reduces Tx and increases Tn. Not clear which, if either, dominates. For MMTS in the states Tx change seems to be slightly larger so net effect on Tm is a slight cooling bias.
VV: Is the Doeske study the only pair of MMTS and Stevenson screens in all of the USA? That would be much too little for such an important transition, if only politically important as I would expect homogenization to remove it quite well given that the modern US network is so dense.

- Time of observation change = depends on change

PT: A change to morning I believe tends to increase both and vice versa but its a very long time since I read this literature and I may well be flat out wrong in this respect.
VV: In the USA the observers can select themselves the observation time. In the past they often selected the afternoon, now they often make the measurement together with precipitation and do it more often in the morning. This lead to more double counting in the Tn and thus to an artificial cooling trend. Due to the Watts et al. manuscript, that "forgot" to correct for the time of observation bias, there was a sudden interest in information on this inhomogeneity and I have a blog post about it. http://variable-variability.blogspot.com/2012/08/a-short-introduction-to-time-of.html In other countries the TOB can be in any direction and could be seen as a zero-mean (random) break.

A random change would then be anything else that could have either a warming or a cooling effect (size picked from a zero-mean distribution):
Station move
Instrument change
Observation practice change

We had decided to separate these two things so that we can ensure that a certain proportion of stations contain a certain size of bias introduced from inhomogeneities acting additively, in the same direction, and then add random breaks on top which may occur at the same time as a biased break.

I think Victor's definition of a biased changepoint is now a description of the time series progression such that multiple inhomogeneities are applied throughout the series, each using the previous homogeneous sub-period (HSP) as the new baseline to apply the change from so it becomes a sort of random walk away from the initial

baseline of the clean station. These inhomogeneities can be both positive and negative but with an overall movement towards one or the other.
VV: In principle these are two different matters. Breaks can either be biased or unbiased. And a break can have a permanent influence or only temporary until the next break. The proposal in my post is to make the biased breaks permanent and the unbiased break temporary. I would expect that that is close to reality based on the comparisons we made in the HOME benchmark. We could also study this in the larger global dataset of the GHCNv3. That would be additional work, though.

A random changepoint would be one that applies a change and then takes it away again, returning the series to the baseline of the clean station.

I don't think it really matters what we call it or how we model this. Our layers (1=clustered/known regional, 2a=biased abrupts 2b=random abrupts, 3= gradual with/without random abrupts) can still provide this. They can still also model various types of inhomogeneity discretely too - even though we don't perfectly understand all types (e.g, station move, shelter change) we can design a few examples and sample from them. The bias component will apply changes more likely to be in the same direction (non-zero mean distribution) and the random component can apply changes of any direction (zero-mean distribution). Importantly though, are we assuming that every change made will change everything after it or only that with the HSP? Are random breaks to be platform breaks that return to baseline __-----__
VV: As a further simplification I would propose to see the clustered breaks as the biased permanent ones. The clusters represent technological advances that were implemented in all stations of a certain network. These are also the transitions that cause biases if they have a non-zero average influence on the measurement, because they are implemented in all stations.

IJ: Random walk (non-zero mean) verses noise (zero-mean) makes sense as discussed in Victors blog
PT: We know that some changes are correlated over time. Additive? Basically temporary verses permanent changes.
VV: Mostly noise with a little bit of random walk is what comes out of the observations from HOME.
MM: Can we do this on the US network to find out the characteristics?
PT: Can do some analysis on an adjustment file provided by Matt before Christmas
ACTION Matt to provide adjustments file from PHA to Peter. Peter to assess structure of changes found.
LV: Will this random walk vs noise differ on monthly compared to annual assessment?

e.g., using the most recent period as our reference period and applying changes to all of the preceding record would be a permanent/random walk change - so all inhomogeneities are applied backwards
IJ: The reference period will be at the end of the series so perhaps inhomogeneities should be generated backwards in time. However, the mechanisms I can think of for generating inhomogeneities are reversible in

time, so I don't think it matters which way (forwards or backwards) it's done.
VV: This is a different reference period. I did not use that term in my post on purpose, was already afraid it would lead to confusion.

VV: I am not sure whether the beautiful drawings below help much. You can only see the difference between a biased break and an unbiased break in an ensemble. Same for clustered and non-clustered and temporary and permanent.
KW: Just assume this is one result from an ensemble - the possible distributions to pick the size of the biased breaks have a positive (non-zero) mean so are more likely to be positive.
KW: I like this definition of temporary and permanent and hadn't really thought about temporary changes before. Temporary changes could be:
A temporary station move during maintenance and then move back.
An instrument change that was noticed to be a bit weird and so changed back.
A bird nesting in a poorly attended screen for a while?
VV: Or more frequent relocations, change of observer, change in the surrounding, it does not have to be a very short period.

KW: Ah ok – you do not mean temporary in the sense that it will always come back to the same place it started from. So – permanent applies to entire record preceding HSP (a new baseline) – temporary applies only to HSP. So, essentially a temporary does return to the previous baseline but most likely a non-zero adjustment is applied at the same time so a new temporary baseline is created. In that case I'm happy with permanent and non-zero mean vs temporary and zero-mean categories.

STATION X

_____PRESENT

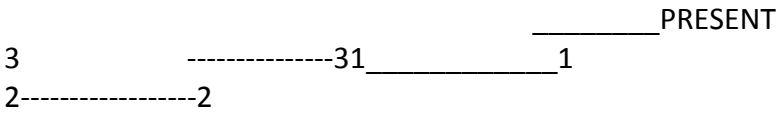
STATION X WITH ONE PERMANENT BIASED BREAK IN THE POSITIVE DIRECTION (makes earlier period more negative relative to present) e.g. stevenson screen to AWS

_____PRESENT
1_____1

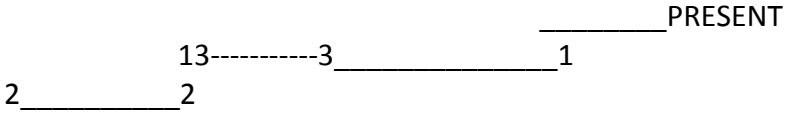
STATION X WITH TWO PERMANENT BIASED BREAKS IN THE POSITIVE DIRECTION e.g. Time of observation bias

_____PRESENT
1_____1
2_____2

STATION X WITH TWO PERMANENT BIASED BREAKS IN THE POSITIVE DIRECTION AND A PERMANENT RANDOM BREAK IN THE NEGATIVE DIRECTION (non-platform - applies to whole period prior)



STATION X WITH TWO PERMANENT BIASED BREAKS IN THE POSITIVE DIRECTION AND A TEMPORARY RANDOM BREAK IN THE NEGATIVE DIRECTION (platform with length until next changepoint)



I hope this demonstrates the permanent changes that apply to all data preceding the changepoint (and in this case are drawn from a non-zero mean distribution / biased) and the temporary changes that apply only to the HSP preceding the changepoint (and in this case are drawn from a zero-mean distribution / random).

KW: So - does everyone now understand the concept of biased changepoints and random changepoints and permanent verses temporary? I've tried to illustrate some of these things above using and example where the inhomogeneities are applied to the clean world station moving backwards through time. The layers can be modified:
Layer 1 clustered (known and estimated for the unknown countries)
Layer 2a permanent/random walk (non-zero mean)
Layer 2b temporary zero-mean
Layer 3 gradual changes

PT: really need a good estimate of the power spectra of the breaks
PT: Are we making problems by reducing the cross correlations further back in time by adding inhomogeneities?
VV: Should be applying inhomogeneities to the end so that the most recent period is not artificially highly correlated.
PT: Should we degrade the cross-correlations in the most recent period as this is realistic?
KW: Well, they're built on imperfect data so should be 'realistic' in the present day sense.
KW: Not sure how we're going to ensure realistic cross correlations further back in time - hope we've done a good job, check and if its terrible then redo with small changes?
VV: Cross correlations are key
PT: Compare clean worlds with homogenised data in the real world.
VV: Or without homogenized data, compare corrupted worlds with raw real worlds.
EA: This is good idea.
PT: Compare clean worlds with homogenised products and compare corrupted worlds with raw data - both need to pass tests.

KW: ACTION - talk to Robert about this.

2) Progress Summary from workshop on error world statistics and distribution probability framework - cont'd

iii) How do we build these things?
Flow chart to describe the layers, probabilities, types of inhomogeneity and flow through this process:
Layer one - known (regional summaries)/unknown regionally clustered changepoints (nationwide, airport moves)

function get_clustered_changes
- can apply a bias
- 1 or more one off changes
- specified types (shelter changes, time of obs change, manual to AWS etc)
- applying to more than one station in a country at similar times

VV: I would suggest to implement the clustered breaks as biased (non-zero-mean) and permanent breaks.
KW: Sounds sensible but I still argue for a separate layer 2a that can implement a series of permanent abrupt changes that are most likely to be biased in one direction (from a non-zero mean distribution) and that do not have a countrywide signature.
VV: Do we have a reason to expect that that happens in reality? What would be examples of that?
KW: I would imagine a large number of stations incur such things - if we know that bias exists in many stations but we do not have countrywide explanations for a series of non-zero mean shifts then these things are occurring more randomly over time.
VV: Good point. A relocation would be permanent (at least until the next relocation).

Layer two - non-clustered abrupt changepoints

a) function get_biased/permanent_abrupts
- some stations/regions are poorer quality than others
- apply bias (1 or more changepoints from a non-zero mean distribution)
- use specified types of inhomogeneities (governs shape and size) e.g. shelter change, time of obs change, move from city to airport
- not countrywide
- applies to all data preceding changepoint
- could add probability of a random adjustment occurring at same time (run function get_random_abrupt with high probability of certain types occurring?)

b) function get_random/temporary_abrupts
- some stations/regions are poorer quality than others

VV: I would implement this by change in the global average frequency of breaks 5 per century, with a random number per network/country and a random number per station. Both random numbers could have a sigma of 1 break per century. The same could be done for the typical size of the breaks (the standard deviation of the normal distribution).

- apply random (1 or more changepoints from a zero-mean distribution)

- use specified types of inhomogeneities (governs shape and size) e.g., station move, instrument change etc.
- not countrywide
- applies only to HSP preceding the changepoint

VV: I would suggest to implement the non-clustered breaks as zero-mean and temporary breaks.
KW: These could potentially be zero-mean and permanent though.
VV: Yes, maybe we should also have some unbiased permanent breaks

Layer three - gradual changepoints with or without abrupt random changepoints at start/end

function get_gradual_changes
- 30 % of stations have 1 or more gradual changes
- can be in any direction - greater probability of warming?
- can be assigned by type: urbanisation, forestation, shelter aging, regional irrigation
- not countrywide
- apply only for length of HSP
- probability of a random abrupt change occurring at beginning (low) and or end (higher) - run function get_random_abrupts

KW: Are we generally happy with this flow process - see also flow diagram: https://docs.google.com/drawings/d/1fb-kjL2J1oG1KmR2c_4hWhlSvV3Ats7xDGXP45aAU7w/edit

SOME DECISIONS TO BE MADE - PROBABILITY/DISTRIBUTION:
a) QUALITY RATING: Probability of being a terrible to excellent station/country grouped from 1 to 5 - 1= no breaks, 2=very few breaks, 3=moderate breaks, 4=quite a few breaks, 5=terrible
PT: Arguably some of the more actively managed networks will have more frequent breaks. Paradoxical?
KW: Good point, annoying.
VV: We could also implement the quality as a continuous random variable that determines the break frequency and magnitude.

b) How many breaks to apply (abrupt biased (permanent or not), abrupt random (permanent or not), gradual)

c) Where to apply breaks (and their length or just their start point (working backwards))?
PT: In terms of applying these it makes sense to 'reverse' the homogenization process which (invariably?) homogenizes relative to most recent segment?
KW: Sounds sensible

d) Specify a type of break using probability - some types are more common than others e.g., station moves.

VV: That may be a good point for discussion, I do not think we have enough information even for the most typical types of breaks (relocations, screen changes) on their probability and their distribution. It might be easier to just have two types: biased and unbiased inhomogeneities. The ones we wrote into our table. And what would be the advantage of making it more complicated by using multiple types of inhomogeneities? Would it make the benchmark easier or more difficult?
KW: But we know that some types of changes are more likely to have certain types of seasonal cycle. e.g., an instrument change may have no seasonal cycle, a station move may have either a seasonal cycle where all points are warmer but some parts of the year are very warm or a seasonal cycle where some seasons are warmer and some cooler, a shelter change may lead to warmer summers and little change in winters. Even if distributions for size are similar its nice to pick the shape in a vaguely informed way.
KW: I like the idea of customising a few types that can be picked from using bits of stuff we do know and guesstimating the rest for now.
e.g.:
ABRUPT CHANGES:
station move (cooling - city to airport - simple seasonal cycle)
station move (cooling - city to airport - complex seasonal cycle)
station move (random - simple seasonal cycle)
station move (random - complex seasonal cycle)
Time of observation bias (cooling, seasonal cycle) - do these have a seasonal cycle?
Time of observation bias (warming, seasonal cycle)- do these have a seasonal cycle?
Shelter change (cooling - wild/north wall/cotton region to Stevenson screen - seasonal cycle)
Shelter change (cooling? - Stevenson screen to ventilated AWS - seasonal cycle)
Instrument change (random)

GRADUAL CHANGES:
Screen aging (warming - seasonal cycle)
Urbanisation (warming - seasonal cycle?)
Vegetation growing (cooling (?) - complex seasonal cycle)
Increased irrigation (cooling - seasonal cycle (no non-growing season effect?))
Can't think of many cooling things.

PT: Seasonality will have very high latitudinal dependence - need to factor this in too. If its governed by radiation then this may be ok otherwise we need to think about applying things by latitude.
LV: No idea at present percentage of time/changes that have /do not have a seasonal cycle
ACTION LV to talk with CW? About retrieving seasonal cycle information from the data?
MM: This is probably a lot of work - CW has started it. but we'll probably have to guess this first time round.
MM: Grid os seasonality of time of observation bias from paper, Karl et al. 1986. We can at least provide the geographic/seasonal impact of the time of observation bias (say afternoon to morning relative to local midnight)

Karl, T.R., C.N. Williams Jr., P.J. Young, W.M. Wendland. A Model to Estimate the Time of Observation Bias Associated with Monthly Mean Maximum, Minimum and Mean Temperatures for the United States. Journal of Applied Meteorology, 25, Issue 2, pp.145-160, doi: 10.1175/1520-0450(1986)025<0145:AMTETT>2.0.CO;2, 1986.

EA: For the tropics and non-western countries I have some datasets I can look through but there is no metadata.

LV: What about the parallel stations?
VV: We currently only have Slovenian data
https://ourproject.org/moin/projects/parallel
ACTION LV to send parallel data to Victor.

e) Specify the size of a break (mean) using a Gaussian distribution with mean and st dev specified for that type
PT: Are these things truly normal in all cases or is this a necessary evil assumption to make the whole problem tractable?
KW: Normal is nice! Mean and standard deviations may be sufficient here.
VV: The NOAA study on breaks know in meta data showed that averaged over all break types, the normal distribution is quite good. For individual break types we have no information. The lack of information could be a reason to assume a normal distribution, for its simplicity.

f) Specify seasonal cycle shape (sine curve) with a mean and st deviation based on type of inhomogeneity - could be of opposing directions or same direction across year or only applicable for part of the season.

g) Probability of abrupt random break occurring simultaneously with a biased break?

h) Probability of a 1 or more gradual changepoints occurring

i) If this is a station with graduals assign a number of 1+ gradual inhomogeneities

j) Decide on the length/location of each gradual inhomogeneity

k) Probability of a random abrupt change point being applied at beginning of gradual?

l) Probability of a random abrupt change point being applied at end of gradual?

m) Probability of metadata being present? This may link to 'quality rating' of station/country.
PT: In realistic worlds also not just availability at issue but whether in some machine readable format of course. Realistically until we get our house in order in the real world will be hard to make use of more than US metadata.
VV: Currently mainly linked to the use of English in country.

KW: Can we put probabilities/appropriate distributions to help decide these on the above?

CG: Agree with Ian - no need for Poisson - better to use exponential and geometric.
CG: If we can get an estimate for % of additive vs non-additive and can then flip a coin to decide between - geometric
IJ: Simplify the flow chart and remove Poisson.
VV: Is this a simplification? Each time point has the same probability of having a break
IJ: Could be a seasonal change in probability over time.
CG: Can subdivide a random number 3 ways e.g. 0-0.95 no change, 0.95-0.96 additive, 0.96-1 non-additive.

ACTION: Kate to work with IJ/VV?CG on reworking the flow chart.

NB:
Changepoints occurring in periods of missing data forced to occur at beginning of missing data period for ease of assessment
Changepoints allowed to stack on top of each other - multiple changes can happen at once - considered as one changepoint for assessment though
Changepoints allowed to occur close to each other and within first and last two years of data - realistic problem.
No random degradation of the data to mimic poor quality data - assume everyone has (will in the real world) conduct a reasonable standard of quality control
KW: These were agreed previously - lets hope we still agree.

3) Using latest version of changepoint size/frequency distributions (overview_error_worlds.xls - circulated) - discuss to finalise if possible
See table in Notes below
Size of changepoints (abrupt and gradual) in the Tropics. Fewer gradual in the Tropics?
Kate summarise from chats with Blair

KW: Agreed numbers of table?
PT: Lets start developing on these numbers and then review some proto series that result (you could develop the bias series w/o needing to implement onto the clean analogs) and reassess. I'd like to see some just series of the posited break structures to be applied visually which would be more easily interpretable than a table of dry numbers.

4) Concepts paper - ready to submit?
Kate aiming to assimilate comments to date (v2) on Friday. Recirculate next week with aim for very minor alterations after that and submission, at least to internal review (MET O, NCDC) before Christmas.

5) Metadata - how to generate and how good/bad?

If time - any thoughts?
KW: Metadata available for US stations? Just replicate this? No worlds specifically state metadata inclusion. Could include vs not include for some countries in B8 that have the same 'exotic' changepoint structure applied.
PT: Metadata may want to be more complete for a world we wish to be 'tractable'. Thinking politically here in terms of what messages we may wish to put out AFTER the benchmarking, saying that x% (I'm guessing 90%) of algorithms performed better when metadata was quasi-complete would send a strong message to the community and funders about how its also important to concentrate on metadata rescue, preservation, digitization and standardization? Regardless, perhaps in the open worlds produce a world with perfect metadata. This would allow users to test the verity of their adjustment component to their algorithm in an isolated sense in that they could prescribe the breaks absolutely. I guess if the world is open this is in theory calculable anyway though?
VV: The default should be the generate artificial metadata for those networks where the database also has real metadata and the quality of the information should match the real dataset as well. To create some political pressure to generate better metadata, we could have one world where metadata is available for all networks/countries. I would not make such metadata perfect, but just as imperfect as real metadata.

6) AOB
Terms of Reference need to be updated

IJ: If time permits could we have a brief discussion of what should go in the glossary. It's clear we need things like biased/unbiased breaks and reference period, but how widely should the net be cast?
KW: Good point
ACTION - add in terms you would like to see in the ISTI glossary and I will add them in: EVERYONE!!!!
reference period
geometric distribution
permanent break
temporary break

7) Next Meeting: week of December (oops) 9th-13th A reversal of time -has Dr Who joined the group?
VV: Main topics: seasonal cycle of inhomogeneities and artificial metadata. Possible topic: ideas for additional open worlds.
EA: Off to Paraguay
IJ: Away on the Friday

#---
Notes:
Blind World Reminders:
Studying the influence of biases
Statistical inhomogeneities

#B1. Best guess world for the West everywhere. A mix of random and biased abrupt breaks with
some gradual inhomogeneities, some spatially correlated breaks, seasonally varying breaks, realistic
missing data.

#B2. Best guess world (#B1), but no spatially correlated breaks.

#B3. Best guess world (#B1),but more and smaller unbiased breaks and gradual IH. The properties
of the biased breaks stay the same.

#B4. Best guess world (#B1),but fewer and larger unbiased breaks and gradual IH. The properties
of the biased breaks stay the same.

Physical inhomogeneities
#B5. Best guess world, with a bias of 0.2°C per century at high- and mid-latitudes and 1°C near
equator. Implemented by making the bias a function of insolation and log(humidity) (or net IR
surface flux at night), if they are capable of producing biases).

#B6. Best guess world (#B5),but instead of ~2 breaks per century with a bias, it has ~4 breaks per
century with a bias on average. Total trend bias the same, thus the 4 biased breaks only have half
the bias size.

Random and biased breaks.
#B7. Best guess world (#B5), but exploring different background climate?

#B8.Best guess world (#B5), with national more exotic inhomogeneities. Next to the typical
exposure and relocation based inhomogeneities, there are many less frequent causes that have their
own specific signature. They typically happen in just one network and by implementing them only
in a small number of countries, we can try many different inhomogeneity problems.

Studying the influence of the seasonal cycle
These two blind worlds should be analysed together with #B5 and #O3 (a world without an annual
cycle in the inhomogeneities).
#B9.Best guess world (#B5) where the biases are implemented by using the equations of Auchmann

and Brönnimann (2012) taking insolation, humidity, wind and snow cover into account.

#B10. A more difficult seasonal cycle that only affects a small number of months, up to one season.
As in all cases with a seasonal cycle, this would include occasions where breaks were in opposing
directions for different parts of the seasonal cycle.

overview_error_worlds Table:

		Bias ****		Random Breaks			Local Trends		
		West	Equator	Size*	Length	Seasonal Cycle	Length	% stations	Warming rate***
		°C/100yr		°C (sigma)	years	°C (sigma)	years	%	°C/100yrs
Statistical Inhomogeneities									
B1	Best guess for the west everywhere	0.2	0.2	0.7	15	0.35	25	30	1(-2 to 4)
B2	B1+ no spatially correlated CPs	0.2	0.2	0.7	15	0.35	25	30	1(-2 to 4)
B3	B1+more/smaller random CPs	0.2	0.2	80%=0.5, 20%=0.7	10	0.25	25	50	0.5
B4	B1+fewer/larger random CPs	0.2	0.2	1	20	0.5	25	10	2.5
Physical Inhomogeneities									
B5	Best guess everywhere	0.2	1	0.7	15	0.35	25	30	2
B6	B5+more/smaller bias CPs	0.2#	1	0.7	15	0.35	25	30	2
B7	B5+different background climate	0.2	1	0.7(biased:0.35)##	15	0.35	25	30	2
B8	B5+ national/exotic CPs###	0.2	1	0.7	15	0.35	25	30	2
B9	B5+seasonalAuchmannBronnimann	**	**	**	**	**	25	30	2
B10	B5+complex seasonal	0.2	1	0.7	15	0.6	25	30	2
O1	No inhomogeneites	0	0	0	0	0	0	0	0
O2	Only large CPs	0	0	-1 to 1	15	0	0	0	0
O3	No seasonal cycle	0.2	0.4	0.7	15	0	25	30	2

Allow CPs in first and last two years

KW: May need a best guess open world or giving away too much?

KW: Seasonal cycle in O2?

KW: Why change bias for Equator in O3?

* The sigma of the unbiased breaks and the sigma of the random part of the biased breaks

** What equations produce

VV: Auchmannn and Brönnimann would only produce the biased breaks, we would still need aditional random ones, would suggest to do so like the default word.
*** Average warming rate in °C per century during affected period (average over benchmark will be smaller; every station will have its own value)
**** Bias in °C per century

Average bias per century the same, but double number of biased breaks with half the size of bias.
Size of random breaks 0.7, random part of biased breaks 0.35°C.
Exotic world has different inhomogeneities in many countries, the other countries will have the default values, listed here.

IJ: Some notes on Poisson, geometric, exponential and other distributions
The confusion here is partly because I was thinking of time as continuous. Of course, given that data are monthly, there is only a finite number of months in a station's time series, so things are discrete, not continuous. However, if the number of months is large it may be convenient to consider time as continuous, and move any chosen change time to the nearest month, for example.
Continuous time – the Poisson process. If changepoints for a station are equally likely to occur anywhere in time and are independent of each other, then changepoints follow a Poisson process. The number of changepoints in any fixed interval of time has a Poisson distribution and the time between one changepoint and the next has an exponential distribution.
Discrete time – if each month in the series has the same probability of being a changepoint, then the number of changepoints in a fixed number of months has a binomial distribution and the number of months between one changepoint and the next has a geometric distribution.
Thus, if time is considered continuous we can generate positions of changepoints using successive exponential variables. For discrete time we can use successive geometric variables, though it's equivalent and simpler just to independently choose each month to be a changepoint or not, with a given probability. The same distributions (exponential for continuous, geometric for discrete) could be used to generate the length of a gradual change, although other distributions with shorter tails (e.g. Gaussian) might be preferred for length of gradual change. I can't see a direct use for Poisson variables.