Clim. Past Discuss., 7, C1401–C1404, 2011
www.clim-past-discuss.net/7/C1401/2011/

Climate of the Past Discussions

Interactive comment on "Benchmarking monthly homogenization algorithms" *by* V. K. C. Venema et al.

Anonymous Referee #2

Received and published: 12 September 2011

General Comments: This is a very interesting report on the work of the COST HOME project that discusses the use of novel methods for benchmarking homogenisation algorithms. This research is of high value to the scientific community and should be published promptly. However, while the paper is very thorough there are a few areas that I would like to see clarified for the reader. I apologise for the large number of comments but they are mostly very minor in nature.

Specific Comments: Abstract line 26: "Training" training in what? Use of the benchmarks? Building algorithms?

P 2664 line19+: More information is needed here. Why were those networks so poorly homogenised? What errors were added? We can learn from this if you provide sufficient information and justification for not including them. If they are realistic inhomo-

geneities then they could be useful for development of improved algorithms to some extent. Also, what is meant by "Selecting stronger did not changes the results anymore."

P2665 line 1+: Can you say more about the real data. Why were these stations chosen? Are they well studied and well documented? Are there references you can provide relating to them? Can the reader have access to them? Are you 100% all inhomogeneities within these series are 'known'? Are they monthly or have you averaged them? Have they undergone quality control by you or by the National Meteorological Service?

Sections 2.2 and 2.3: Are there any limitations of the IAAFT method? Perhaps not. It would be really nice to see an example of both a surrogate and a synthetic time series. Are there features of real data that were not well simulated i.e., ENSO, large volcanic eruptions such as Pinatubo etc.

P2666 line 25: Can you reword the "frequency is drawn from a uniform distribution between 2 and 8%" as I'm not clear what is meant by this.

P2667 line 10: Can you provide more information on the seasonal cycle of the breaks added? I wasn't clear on how this was done. Does the seasonal cycle vary from break to break in that sometimes a winter break will be larger than in summer and sometimes it will be smaller? Could breaks be of a different sign in winter compared to summer? This also applies to P2685 lines 6+.

P2668 line 0: What about random missing data? Many time series have missing months dotted throughout the series.

Section 4: You define "relative homogenisation" here. Please can you define "absolute homogenisation" here too as it is referred to in later sections.

P2671 line 19: Does "reconstitution of missing data" mean infilling of missing data?

P2672 line 10: How do you obtain the values for "true negatives"? Do you treat each

CPD

7, C1401–C1404, 2011

month/year without a break as a potential one or do you assume that breaks could occur with a certain frequency e.g., every 6 months. Also, how do you deal with location of a break 1 month too late or early? Is there any margin of error in the assignment?

P2672 line 22: “The reference r in Eq. (1)” - should that be “The reference rstd in Eq. (3)”?

P2673 line 20: Could add “ when the break locations and magnitudes were then known” to the end of this sentence to make it more explicit.

Section 6.1.4: Do you think the decrease in CRMSE from older to more recent data in the time series is anything to do with the way in which corrections are applied? Are older chunks of inhomogeneous data adjusted to match all of the more recent data or just the most recent homogeneous chunk? If there is a background trend in the data this may also have an effect. This perhaps links in with the last sentence “This fits to Climatol stating the correction of the breaks at the beginning of the series”. I’m not quite sure what is meant by this.

P2679 line 18: What is a “predicted break”? Are these the same as detected breaks?

P2680 lines 11-13: I’m not quite sure what is meant by this sentence.

P2686 line 3-4: Are cross-correlations the percentage of variance in one time series that is explained by another time series? I wasn’t quite sure what metric should be used here.

P2686 line 12-14: This sentence could do with a reference or a little more anecdotal evidence of where this has been observed.

Section 7.1: Do you have any ideas of how to improve/enable true objective intercomparison of algorithms in future research? This is a really valuable assessment but some guidance on how to progress given the lessons you have learned from this research would be really helpful.

CPD

7, C1401–C1404, 2011

P2687 line 8: “more strongly cross-correlated”?

P2687 line 19: How do you define the “all-over best”?

P2693 line 15: I disagree with this statement and it undervalues the very considerable and useful work that you have done. I think you have shown great value in the benchmarks both for validation of algorithms and also for development although as you say, these should be separate components as you cannot validate with a benchmark that has been used to tune an algorithm. A cyclical program would be useful and you do discuss creation of more benchmarks every few years. Something like: design benchmarks, invite blind testing, release results and then allow development using old benchmarks, design new benchmarks, invite blind testing etc. etc.

P2693 line 21+: This is a really useful summary of recommendations. Can you structure this more explicitly as recommendations to the community, perhaps using bullet points. I think this would make this stand out much more to the reader.

P2694 line 29+: Why should networks without added inhomogeneities be studied separately?

Gradual inhomogeneities are discussed in the input errors to simulate urbanisation or a growing environmental feature. However, the ability of the algorithms to find such inhomogeneities is not discussed.

Table 1: Please can you define “DP”, “HBS” and “MLR” in the table header?

Interactive comment on Clim. Past Discuss., 7, 2655, 2011.