Clim. Past Discuss., 10, C353–C356, 2014
www.clim-past-discuss.net/10/C353/2014/

Climate of the Past Discussions

Interactive comment on "Identifying homogenous sub-periods in HadISD" *by* R. J. H. Dunn et al.

E. Aguilar (Referee)

enric.aguilar@urv.cat

Received and published: 6 May 2014

General comment:

This is an excellent article, suitable for publication with just a few minor modifications. Presents relevant scientific questions which fully match the scope of CP. It not only introduces HadISD, but provides assessment on its homogeneity and usability with an state-of-the-art approach. The paper reflects the authors knowledge on the topic, through an adequate background and including the necessary references. The title, though, does not entirely cover the contents of the manuscript: I suggest to include on it a reference to the study of the homogenization process performed, as this is a very important part of the article. On the contrary, the abstract does cover this part and provides a good summary. The rest of the article is well structured, English usage is excellent and very easy to read, except for a couple paragraphs that may need some

rewriting for the sake of clarity (see detailed comments below). The supplied figures are informative and necessary, although I will suggest to improve a couple of them. The results are fully traceable. Detailed comments: - Page 1569, line 3: Where says "Two main approaches exist for determining the location of change points" I would add a reference to the superiority of relative methods over absolute methods: "It is widely accepted that relative homogenization (based in comparison between candidate and neighboring series) is preferred to absolute homogenization (based in the analysis of candidate stations data alone). - Page 1569, line 12: Please, notice that MASH is the homogenization package and MISH is an interpolation pacakge; SPLIDHOM is different than the rest as it does not detect inhomogeneities but adjusts previously detected inhomogeneities in daily data. I also miss here a little more information on the properties of HOMER, MASH and ACMANT. For example, a reference to the detection principles in HOMER (based on the ancient prodice (Caussinus and Mestre, 2004 and on a joint segmentation algorithm initially code in genetic science (Pickard et al, 2011). Also, it should be necessary distinguish the scope of HOMER - very suitable for medium size networks, were the climatologist input is possible, meanwhile other approaches, such as PHA will do a much better work with large networks, such as HadISD or larger, where in-depth station by station analysis is not practical/possible.

- 1570, line 23 and adjacent: distinguish between methods which adjust daily data from those which detect inhomogeneities in daily data. Most of the quoted references rely on other methods applied to lower resolution data (monthly or annual) for detetection.

- 1571, line 3: where says "Following the terminology used in the ISTI ... " I would say : "Following the widely accepted terminology, adopted in the ISTI ... "

- 1571, line 15: could you give some details on which studies are sensitive and which are not?

- 1572, line 12: COST-HOME networks were small networks, i.e. no larger than 20 stations, in ocasions as small as 5. This was most likely a drawback for PHA when

CPD
10, C353–C356, 2014

applied to COST-HOME I think it is worth to mention here.

- 1573, line 1: I wonder how PHA would perform replacing in this step 2 SNHT by other algorithms. Any available analysis on this?

-1573, line 12. “The PHA code ... converted to monthly values” ... Although you discuss in the following sections how the monthly means are computed, I suggest to give some details here or refer to adequate sections.

- 1574, line 11: this comment is linked to the previous one. I wonder if you have assessed potential problems and inhomogeneities introduced by how monthly values are computed. Most likely, in some stations the available hours, even the number of daily observations are changing from through the data record. Also, computing a month with 20 values is far, for example, from WMO’s 5/3 rule. Any evaluation on the impact of this?

- 1575, line 1: according to the claim here, roughly 1 out of 5 stations is homogeneous. This is a but surprising. In the forthcoming sections you discuss on the role of the length of record and the availability of well correlated neighbors. I think it would be necessary to advance some hints on why so many “homogeneous” stations.

- 1575, line 7: “There is no pattern to the stations which could not be processed”. I think here it is necessary to stress two concepts, additionally to the lack of completeness which you mention: low station density and large decorrelation due to complicated geographical patterns. Low density surely applies to Africa and, for example, in western South America, the presence of the Andes range introduces changes in altitude and very different climatological characteristics in relatively small distances, thus making the selection of well correlated neighbors more difficult.

- 1576, line 18: does this mean underdetection in low density areas?. I think it is important to stress it.

- 1579 line 20 to -580 line 13: this part is difficult to read. I suggest to reword it. - 1580,

CPD
10, C353–C356, 2014

line 29: any explanation on this? Incomplete metadata? Homogenization artifacts? Averaging artifacts?

- Figure 3: indicate which two methods.

- Figures 4 and 9 are difficult to interpret.

Interactive comment on Clim. Past Discuss., 10, 1567, 2014.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper