

## ***Interactive comment on “Multiresolution analysis (MRA) classification of plurennial to multi-decadal climate drivers to streamflow in France using Wavelet Transform and Geostatistical Euclidean Distance Clustering” by Manuel Fossa et al.***

**Anonymous Referee #2**

Received and published: 18 January 2017

The subjects treated in the paper "Multiresolution analysis (MRA) classification of plurennial to multidecadal climate drivers to streamflow in France using Wavelet Transform and Geostatistical Euclidean Distance Clustering" are consistent with the subject areas of the journal and they are interesting for the scientific community who reads it. In particular, the article addresses the problem of clustering hydrographic stations according to 6 – 10 – 21 year trends and to study the pattern that drives them. The algorithm to solve this problem consists in: 1) Signal decomposition in different Time Scales of Variability (TSV), 6, 10, 21 years, by using wavelet analysis 2) run correlation analysis between each hydrographic station and each point of a variable field. This allows to

C1

build a correlation map called Climate Correlation Field (CCF) 3) run GEDC, derived from Image Euclidean Distance (IMED) proposed by Wang et al.(2005), in order to clusterize hydrographic stations with a similar behavior 4) a CCF composite analysis is performed for each cluster, TSV and atmospheric variable. Although the application is interesting for the scientific community and the method is original I do not recommend it for publication in the present form. The article is incomplete in some places, rambling in other parts, and wrote in a hasty and conversational style in other parts. I have some doubts about the decomposition by using wavelet, too. Going deep into analysis:

- Gaps in the paper: The Bibliography is poor (25 references) for a topic that would requests at least twice. In particular references could be extended in the introduction and in the discussion paragraph. E.g., the discussion of results is structured into three subparagraph and it is long enough, but the authors focus only on their own results and do not compare them with the extensive literature about European Climatology. In particular the first subparagraph only cites 3 sources, the second 0, the third 3 ... I feel this part could be much improved.
- Rambling paragraphs: Paragraph about methodology is not easy to follow. Some of the 4 algorithm points are well developed (the one on Wavelet decomposition is very discursive, but it is enough explained). However some others consist of just one sentence, i.e. correlation is only explained in line 107, in which I have to guess which measure of correlation has been used... I guess the Pearson one. To facilitate comprehension of the text and the algorithm, I suggest to divide the method in four sub-paragraphs and to explain them sufficiently, both formally and discursive.
- Style related problems (hasty and conversational): Throughout the paper references are really bad managed, please use a software to manage references (such as mendeley which has a nice plug-in in Microsoft word). In the paper there are a lot of assertions without citation or demonstration (ie, 9-12). There are some sentences like: "It is commonly reported in literature...", "...is still extensively studied": please use citations. In some points of the paper, the style becomes a bit too much colloquial, with a lot of parenthesis (28, 33, 34, 42, 43, 51, 59, 79, 80, 81, 83, 90, 91... and I didn't checked more), please use commas.
- About the methodology: I applied wavelet

C2

methods in the past, but I'm not an expert. However I have some comments about the decomposition methodology: 1) Regarding the part of the decomposition of the time series in different TSVs, I'm wondering if it must be used a significance testing in order to take into account the 95% confidence for a red noise process (Torrence & Compo 1998 "A practical guide to wavelet analysis") 2) In the case it is appropriate to calculate the significance testing of the time series, I'm not sure that it makes sense to calculate the correlation over the entire time series, but only in the part that passes the significance test. 3) In the case it is appropriate to calculate the significance testing of the time series, I'm not sure with timeseries ranging from 1968 to 2008 if it is possible to effectively evaluate the TSVs 21 years.

I suggest to the authors to verify the comments about the methodology, which are the most critical.

Please also note the supplement to this comment:

<http://www.hydrol-earth-syst-sci-discuss.net/hess-2016-395/hess-2016-395-RC2-supplement.pdf>

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-395, 2016.