Nonlin. Processes Geophys. Discuss.,
https://doi.org/10.5194/npg-2020-28-RC1, 2020

Nonlinear Processes in Geophysics Discussions
Open Access
EGU

Interactive comment on "Ordering of trajectories reveals hierarchical finite-time coherent sets in Lagrangian particle data: detecting Agulhas rings in the South Atlantic Ocean" *by* David Wichmann et al.

Anonymous Referee #1

Received and published: 25 August 2020

General comments

Overall, I find this paper to contribute potentially useful technical refinements of clustering methods for Lagrangian trajectories. The paper offers a modest but meaningful contribution, and is wisely concise in its presentation. The reachability plot does seem to be a useful conceptual tool, and the ability to admit incoherent or 'noise' regions is a nice refinement.

I find there to be room for improvement on a few presentational issues. (i). There

seems to be an assumption of familiarity with other clustering methods. The paper would be more accessible, and therefore useful, if the authors took just slightly more time in defining new terms and in providing the intuitive content of mathematical concepts. (ii) I find it a little strange that some figures are presented in the appendix, but discussed only in the main text. Some of these make good illustrations of the performance of the method with respect to others, e.g. D1&D2. I feel this tends to negatively impact the narrative. If figures are discussed in the main text, I would present them there also. (iii) The paper has a highly technical focus throughout. More framing of the import of this problem at the start and end would have been appreciated. (iv) For a short paper, the abstract is perhaps disproportionately long.

There is one point raised in the paper that I felt required more elaboration. A selling point the authors bring up for this method is that it can in principle be applied to real-world trajectory data, see line 86 and also line 306. This is true but incomplete. Real-world Lagrangian instruments are sufficiently sparse that it is rare to find more than one in the same eddy at the same time. Thus, the application presented herein—finding eddies is idealized configurations—is not really relevant for how one would apply this method to real-world trajectories. The data density used here is orders of magnitude greater than for real-world instruments. Since the authors bring this up as an advantage of the method, a more fair and nuanced discussion of its potential and limitations with respect to real-world data is called for.

I would say, rather, that the method seems more suitable in application to model data or virtual trajectories from altimetry, where it benefits from a simplicity with respect to some other proposed methods.

Specific comments

Line 99. Do you not want to cite Bickley?

W.G. Bickley D.Sc. (1937) LXXIII. The plane jet , The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 23:156, 727-731, DOI:

10.1080/14786443708561847

My understanding is that the term “Bickley jet” itself is used to refer to a steady solution with a sechˆ2 u-velocity, see e.g. Swaters (1999). The authors' Eq. (2) is an added perturbation. As read, it sounds like the whole thing is the Bickley jet.

Section 3.2.2. I didn't really understand this section, or what B is encoding in Eq. (4). A more intuitive description would be helpful. When you say, “pairwise distances are approximately preserved”, this is with respect to what? Also, why are two dimensions chosen?

Line 193. The intuitive meaning of the ‘generating distances' that are not being used here should be mentioned.

Line 196. The definition of the epsilon neighborhood appears incomplete. Is it not the M-dimensional sphere of radius epsilon? Otherwise, what is the epsilon?

Line 200. It would be very helpful to write out in words the meaning of Eq. (6). My understanding is that c(p) is minimum distance epsilon such that the number of points in an epsilon neighborhood is greater than a specified number.

Line 213. I did not immediately understand how it arises that there are valleys in the reachability if you have sorted iteratively on the reachability. You might explain that this happens as you encounter groups of points that are all near to each other, thus replacing earlier high values of reachability with lower values.

Line 216. The phrasing here made me wonder if this was a second, different epsilon. It would be clearer to say that you choose a value for the parameter epsilon. Also, it appears this is conditional on a choice of s_min which should then be emphasized.

Line 228. What are the permissible values of k in condition (a)?

Figure 2, what are the units of the y-axis in the left column of plots?

Figs 2 and 3, some of the colored dots lie above the epsilon threshold.

Figure 4. I really don't understand the two dimensions of these plots, nor the star-shaped patterns, could you explain these more?

Data locations at Zenodo should be cited, not only the papers referring to them.

Minor typographic comments

Throughout the paper, the authors consistently omit the subject ahead of an infinitive, e.g. “which allows to detect”. I believe this is grammatically incorrect (in US usage anyway). “allows one to detect” or “allowing the detection of” sound better.

l 42 and 90. “sparse” should probably be used instead of “scarce”. The former means thinly distributed while the latter means hard to come by.

l 128. NumPy and Zenodo are the standard capitalizations

l 141. “method” should be “methods”

l 156. straightforward

l 191. “and as will become clear”

l 217. “is equal to” should be “set of points is equivalent to”.

l 243. “a priory” should be “a priori”

l 279. “large- and small-scale”

l 354. GitHub

l 359. There is a title of an appendix with no appendix.

l 360 & 361. “particle-based”

l 383. “ot”

l 389. There should be a period at the end of this sentence

Figure C1, “three” eigenvalues should be “two”, correct?

Interactive comment on Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2020-28, 2020.