Review of: "When a Cluster Is a Cluster"

Subhrajyoty Roy[1]

1 Indian Statistical Institute

Potential competing interests: No potential competing interests to declare.

The article provides an overview of the importance of spatial clustering methods with applications to epidemic spread. The author argues that understanding the geographic distribution and clustering of cases is critical for characterizing and controlling epidemics, especially in the early stages. The author also provides a brief review of methodologies for detection of spatial clusters and advocates for a Topological Weighted Centroid (TWC) algorithm.

Overall, the central argument of the paper is sound, but it lacks substantial evidence or demonstration to support the argument. A few major concerns are listed as follows:

1. When referring to the dot maps for the spatial clusters, it is not clear whether one should consider the location of the testing center/detection center or the location of onset of the disease.
2. On the same point, in the early stages of epidemic spread, when there are no serious lockdown impositions, there are migratory movements of the people. In this case, it is difficult to represent them as dot maps; one has to consider the spatial and temporal information together to account for movements.
3. The author should provide a brief description of the TWC algorithm.
4. "The algorithm produced the coordinates and heat map of the area" - the author should provide the heat map or summarize the results in figures or tables.
5. Just applying the method to a single instance of Covid-19 in Italy is not statistically significant to conclude anything. The authors are encouraged to apply their methodology to other available datasets from other countries.
6. "Through other functions of TWC, it was possible to build a prediction for the near future and the future of the future." - How is this prediction performed? Also, "future of the future" - is this correct?
7. Whenever some acronyms are used (FANNY, PAM, CLARA, etc.), the author should indicate their full names in brackets.
8. In the literature survey, the author advocates for using a fuzzy algorithm. Is TWC a fuzzy-based method? If not, how does it help, and why is the discussion on fuzzy methods relevant?

In summary, the article does not establish clearly the novelty of it compared to the previous works of the same author, mentioned in references [22]-[26] of the paper.