Review of Manuscript egusphere-2025-908

This manuscript uses clustering algorithms to define settlement archetypes for the Italian territory based on variables known to be relevant metrics for risk/resilience from previous studies. The outcome is a set of 10 first-level archetypes, some of which are sub-divided and lead to a total of 18 categories, as well as a classification of Italian municipalities into each of these categories. As the authors well state, the definition of such archetypes has the potential to be relevant for supporting decision-making processes associated with vulnerability and risk reduction.

However, the manuscript appears as incomplete work in this regard. Clustering algorithms are numerical methods that group individuals (in this case, municipalities) as per their similarity with respect to the variables considered. The algorithms know nothing about the meaning of these variables. The manuscript is lacking: (1) a more in-depth interpretation of what it means or what consequences may stem from the fact that municipalities with some differences in certain attributes end up grouped together, as well as (2) a demonstration of some sort that these 18 archetypes are indeed associated with relevant risk metrics. In other words, do these groups actually make sense when confronted against risk analyses (single or multi-hazard)? Is the "vulnerability profile" (as the authors phrase it) meaningful? The fact that the variables selected as input for the clustering are known to be relevant from previous studies does not guarantee that the grouping is relevant and consequential in the risk domain, or suited for designing risk reduction actions.

I believe this manuscript should address this major shortcoming as well as the following main comments before it is considered for publication.

Main Comments

1. Title: The archetypes defined are not just urban. Please erase "urban" from the title and throughout the body of the manuscript.

2. As stated above, the major shortcoming of this manuscript is the lack of in-depth interpretation of the results and confrontation against risk data or models that demonstrate its relevance and meaning for the purpose that the authors state this work has. Here are some thoughts that aim at helping with developing the work further:

    a) The clustering was carried out using a series of parameters known to be relevant to vulnerability/resilience. The outcome is groups of data points that are "similar" to each other with respect to those same parameters. Without a comparison against actual risk metrics, it is not proven that these archetypes mean anything to risk. Moreover, there is no conclusion with respect to how exactly decision makers should use them. Which archetypes should be prioritised for in-depth risk assessments, for example?

    b) What would happen if you just enumerated all possible combinations of your categorical parameters (which could also be used by decision makers to prioritise areas of intervention)? What would the map look like in that case? Does the difference between this "full enumeration" map and the one you obtained say anything about the relevance of the clustering?

    c) Some combinations of parameters simply do not exist in Italy. This is one of the reasons that may prevent the exact same archetypes from being applicable to other regions (which is a potential limitation raised in lines 732-736 of the manuscript). A comparison across countries is only relevant in terms of risk metrics, otherwise the comparison is just about knowing how many combinations of parameters exist in each country.

    d) A clustering algorithm is agnostic to the meaning of the input data. The analysis could have included variables that are not relevant to risk at all, and they would have still come up in the clustering. This is not to say that the risk metrics should be included in the clustering. There is value in carrying out the clustering using the variables you used (because these are variables easily obtained from open datasets, because they can be used as proxies, etc), but the output needs to be interpreted in terms of risk metrics in order to be meaningful for vulnerability/risk reduction.

3. The selection of variables to include is difficult to follow. During my first read I could not understand why some variables from Table 1 were missing in Table 2. It only became clear as I kept on reading (around line 180, i.e. ~50 lines after Table 2 is mentioned in the text). Moreover, line 189 mentions 19 attributes, but Table 2 contains 15. It is only clear in Table 3 that some of the socio-economic parameters of Table 2 are further subdivided. Moreover, in Table 2 it is not clear how some of the parameters are measured (e.g., "quality of the buildings"). Then there is the issue that the number of numerical variables is further reduced due to their correlation (section 4.1), which is correct. But then the final variables used are only listed within the text (lines 386-388), which gets lost with respect to what the reader can easily identify from the tables. In lines 366-367 it is stated that the population variable is used as classes, which was already implicit in section 3.4 (consider moving the spirit of what is said in lines 361-367 to section 3.4). All these issues result in the manuscript feeling messy. Please consider unifying some of the three tables (Tables 2 and 3 are good candidates), using more columns or bullet points to sub-classify items, perhaps marking the 8 numerical variables used in the end, and indicating which variables are categorical and which numerical. Please make it clear from the beginning that Table 1 is just a discussion of potential parameters to include, but not the parameters used.

4. There is some imbalance in the level of detail provided for the first level of clustering (section 4.4), which is associated with the categorical variables, and that of the second level (section 4.5), which refers to the numerical variables. Please consider expanding section 4.5, perhaps including informative figures/plots.

5. Section 5 (results) is merely descriptive. Deeper insights on the meaning of the resulting clusters and their spatial distribution are completely missing. Lines 642-683 mostly repeat what is already said in Tables 4 and 5. Lines 684-705 mostly describe the spatial distribution of the archetypes with respect to Italian regions but (i) the regions are not marked in any map (please do not assume all readers will know where these regions are) and, most importantly, (ii) no insights are provided with respect to what this spatial distribution may mean or imply.

6. Section 6 (discussion) only refers to how the methodology used complies with the guidelines of Piemontese et al. (2022), in some cases inaccurately or obscurely. In some detail:

    a) L720-722: Employing WCD and ICD does not ensure the reliability or robustness of the clustering process. Again, too much faith is being put on the numerical algorithms. How is the accuracy (with which the dataset is represented, as per the wording in L721-722) measured?

    b) L730-731: While the authors imply that confronting the archetypes against risk data/models is outside of the scope of the paper, the classification work per se does not hold relevant meaning without this confrontation.

    c) L732-736: As stated above, comparing the clustering output of different countries without going into the risk domain is just a comparison of how these sets of variables are combined and spatially distributed in different countries. There would be no further implication.

    d) L737-738: The sentence "This study *demonstrates* the potential of urban archetypes in developing risk storylines, enhancing risk communication, and supporting stakeholder engagement" is simply not true. Where exactly in the manuscript are these three points demonstrated?

7. In the last paragraph of the conclusions (L759-761): "*its findings offer a valuable tool for policy makers to design targeted interventions based on specific vulnerability profiles*": The paper does not demonstrate that the archetypes are indeed associated with specific vulnerability profiles, it just assumes they are. Again, the paper needs to somehow show that the archetypes are indeed associated with vulnerability profiles that make sense for policy making.

Minor Comments/Edits

1. L13: "complexity in an otherwise too broad problem".

2. L24: "Over the last few decades".

3. L43: Exposure is not "intended". Do you mean understood? Defined?

4. L43-45: Please replace the semicolons (;) by commas.
5. L58: It says "we assume", but perhaps you mean "we hypothesize"?
6. L61: Erase "We note that". It is not necessary.
7. L83-85: Are capitals really needed for "design", "analysis" and "application"?
8. L95: No comma after "defined".
9. Table 1: It says "GPD per capita" instead of "GDP per capita".
10. L161: "complicates evacuation efforts and strains emergency response resources".
11. L167: Consider breaking the paragraph here and starting a new one to talk about the socio-economic factors.
12. L169: "highlight that the elderly".
13. L192: "proposed a grid-based approach".
14. L201: "adopting ~~to~~ the abovementioned" (remove "to").
15. Fig. 1: Please consider changing the colour of peri-urban areas. The map on the right shows six labels, but only three classes of urban centeredness degree are considered in the analysis. "Hub" and "municipal hub" can be seen as different shades of red of the final class "urban hubs", while the intermediate, peripheral and ultra-peripheral inland areas are indifferent shades of blue, so that is clear. However, peri-urban areas are in orange, which makes them very similar to the "urban hubs" class. A different colour for peri-urban areas would make it easier to visualise the final three classes used. Parentheses in the text (lines 243-245) could clarify this. E.g. "namely urban hubs (represented…, shades of red in Fig. 1), peri-urban areas (new colour) and inland areas (shades of blue)".
16. L258: "lead to the model becoming".
17. L259: "correlation that may exist".
18. Table 2 vs line 377 and Fig. 3: Is it population aged under 15 or 14 (different numbers in different places).
19. Table 2 vs line 311: Should it be "total population aged over 30" under "high educational index"?
20. L344: Clustering methods *per se* do not "ensure" that observations within the same group are highly similar. It is their goal, but it is not guaranteed. They are just numerical methods. They will return some sort of clustering no matter what you provide as input. Please re-phrase.
21. L354-355: "More detailed information".
22. L359: "is crucial for enhancing the quality of the clustering".
23. L376-383: The strong negative correlation between the proportion of employed and un
24. Fig. 3, labels (list 1-14):
    - (1) "Proportion of population under 14" (or another alternative, the current grammar is not right).
    - (2) "Proportion of population over 65" (or another alternative, the current grammar is not right).
    - (12) Replace "components" with "members".
    - (14) "Residents" (plural), not "resident" (singular).

    In general, please aim for consistency between the labels used in this figure and in table 3.
25. Fig. 3: It would be easier to interpret the matrix if the colour scale was based on three colours instead of two (right now, from yellow to blue, passing through green). This would make values around zero more visible.

26. Fig. 3: Is it correct that the correlation between features 7 and 8 is so low, numerically? These features are the proportion of employed and unemployed among the working-age population. Are they not one minus the other (1 – employed = unemployed) and, if not (because of some other category), are they not at least highly negatively correlated?

27. L406: Should the subscript of the x variables be "l" instead of "k", to be consistent with Eq. (1) and the text?

28. L515: "The attributes' importance for clustering is evaluated adopting the simplified procedure".

29. L516: "to assess their contribution".

30. L519: "WCD as the performance metric".

31. L526: Do you mean cluster 2, instead of 6?

32. L564: "The lower ~~is~~ the WCDl" (erase "is").

33. Table 4 is colour-coding the SoVI. Please mention the colours in the text (lines 614-616).

34. Table 4:

    - "Building poor state"? You used "poor" before. Please be consistent.
    - "Foreigners" or "Foreign residents".
    - Clusters 2a and 4b: "Settlements with high social vulnerability".

35. Table 4: As they are phrased, sub-clusters 2a (settlements with high social vulnerability) and 2b (settlements with aged population with high social vulnerability) overlap, with 2b seemingly a sub-category of 2a. If 2a explicitly excludes aged populations, then the label should reflect this

36. L696-697: "in the southern regions, with almost all cases located".