Review of Manuscript egusphere-2025-908R1

I thank the authors for their detailed replies to my previous comments and for having taken them into account in the revised version of the manuscript. While the lack of confrontation of the proposed archetypes against risk data or models that demonstrate the relevance of the classification is still a shortcoming, the more in-depth discussion of this limitation as well as the new section on the vulnerability profiles have contributed to improve the quality of the paper. I recommend this work be published after addressing the following minor comments and edits.

Response: The authors thank the reviewer for their additional comments and suggestions. Please find below the responses to each comment, along with the corresponding edits made to the main text.

Minor Comments

1. L190-192, L205-206, L413-415, Tables 1 & 2, Fig. 3, etc.: While the authors have improved the presentation of the variables they considered with respect to the original manuscript, it is still not straightforward to navigate this matter throughout the manuscript. Section 2 should be more explicit regarding the fact that a larger number of indicators were considered for inclusion but only a sub-set was used, and enumerating the reasons for doing so (e.g., correlation analysis, availability). For example:

   a) "Dependency ratio" is mentioned in L190, Table 1, Fig. 3 and the lines before it, but then it is dropped. Only in the analysis around Fig. 3 it becomes clear that the variable was considered but then discarded due to it being strongly correlated to other variables that were selected in the end. But when reading L190 it gives the impression of being used.

      Response: To clarify which indicators were adopted in the study, the original sentence: 'The complete list of socio-economic indicators considered includes age, dependency ratio, level of education, family structures, commuting rate, quality of buildings, race/ethnicity, employment rate, percentage of women in the workforce (Table 2)' (lines 190–192), has been replaced with: "*A comprehensive list of socio-economic indicators considered is presented in section 2.5, though some indicators were not used for the final clustering process due to their strong correlation with other selected variables (Table 2)*."

   b) The indicator names in L190-192 is different from that in Table 2, and line 192 directly refers to this table. This makes it very confusing for the reader, who might also wonder why the dependency ratio is not in the table (if they go and see the table when directed to do so in L192, before reading the correlation analysis).

      Response: Lines 190–192 have been revised, as indicated in the response to the previous comment.

   c) L205 mentions 19 attributes, Table 2 includes 12 variables, Fig. 3 includes 14. This is confusing for the reader if it is not accompanied by an explanation early on in the text (i.e., Section 2).

      Response: The sentence in line 205 ("*The dataset includes 7960 objects, representing the 7960 Italian municipalities, and 19 attributes (both numerical and categorical) related to the vulnerability factors outlined in sections 3.1 through 3.5*") has been revised to "*All data are collected at the municipal level, aligning with the administrative boundaries adopted for the*

analysis. The dataset includes 7960 objects, representing the 7960 Italian municipalities, and the numerical and categorical attributes related to the vulnerability factors outlined in sections 3.1 through 3.5.". Additionally, to improve clarity regarding the indicators ultimately used for clustering, the following sentence has been added at line 333: "*It is important to note that only 12 of the 14 previously presented social vulnerability indicators are used in this study (Table 2), as a correlation analysis - described in Section 3.1 - was conducted*". Finally, the sentence in line 405 has been modified to (edits reported in orange): "Figure 3 shows the correlation matrix obtained for the 14 numerical variables presented in section 2.5."

Perhaps it would help to add text around certain lines to clarify early on to the reader that not all initially-considered indicators were actually used for the clustering analysis. For example, L190-192 could read: "The complete list of socio-economic indicators considered includes age, dependency ratio, … […] … and percentage of women in the workforce, though some indicators were not used for the final clustering process due to their strong correlation with other selected variables (Table 2)", or similar. I am not saying that the details of Section 3.1 should be stated earlier, but the overall strategy (considered a large number of indicators initially, made a selection of a sub-set to be used for the clustering process) should.

Response: Lines 190–192 have been revised in response to the reviewer's comment, as outlined in the response to Comment 1a

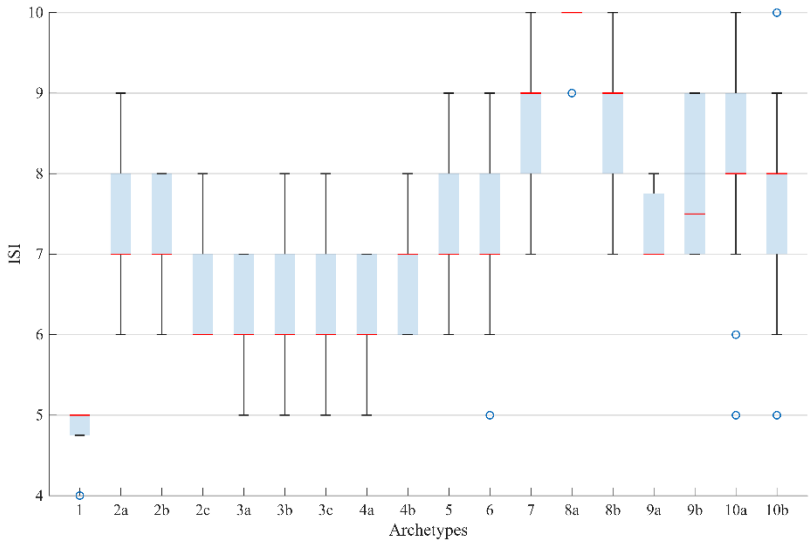
2. L744-747: If I am understanding correctly, the composite index of Sibilia et al. (2024) is not used by the authors. If so, I recommend removing the lines "In Sibilia et al… […] … and environmental", as they include too many details about an index that is irrelevant to what follows.

   Response: The lines citing the composite index from Sibilia et al. have been removed.

3. L767, Fig. 10: Several points here:

   a) The figure and discussion focus on the average ISI but no dispersion is reported. The dispersion is relevant to understand how different the results for each archetype are from one another, especially because so many of them have very similar mean ISI values. Please consider including some sort of measure of dispersion (whichever is appropriate to represent the numerical results) and include it in the discussion.

   Response: The following figure (10b) showing a boxplot of ISI values has been added:



In section 6.1 the following discussion, related to this figure, has been added:

"The box plot in Figure 10b illustrates the distribution of ISI values across the identified urban archetypes, providing insights into both central tendency and internal variability. Archetype 1 exhibits the lowest median ISI and minimal variability, suggesting a consistently low level of exposure and vulnerability across its municipalities. In contrast, Archetype 8a displays the highest median ISI with a very narrow spread, indicating strong internal homogeneity and high susceptibility to impacts. The few outliers with an ISI value of 9 highlight minor deviations but do not significantly affect the overall pattern. Archetype 9b shows the greatest dispersion, reflecting a high degree of internal heterogeneity. This wide variability suggests the presence of municipalities with both relatively low and high ISI values within the same archetype, potentially complicating uniform policy interventions. Archetypes 3a, 3b, and 3c present identical median, mean, and interquartile ranges, which aligns with their shared geographic and demographic features, as well as similar SoVI scores (see Table 3). However, their comparable SoVI outcomes result from distinct socio-economic compositions, as discussed in Section 5, underscoring the multidimensional nature of social vulnerability."

b) It is interesting that the three sub-clusters of cluster 3 have very similar ISI values. Their SoVI values are similar as well (Table 3). Are the ISI dispersion values similar as well? Are there any interesting insights to the method or the nature of these archetypes that are worth discussing in the text?

Response: Archetypes 3a, 3b, and 3c exhibit very similar ISI values due to their shared demographic and geographic characteristics, as well as comparable SoVI scores. However, these similar SoVI values result from different underlying socio-economic features. For example, Archetype 3a has a higher mean aging index, while Archetype 3b is characterized by a higher crowding index. A note on this has been added to the manuscript (see response to the previous comment). Additionally, the following sentence has been included in Section 5 to clarify the value of defining individual archetypes that may reflect specific social vulnerability factors: "*Understanding the influence of individual socio-economic indicators within each archetype can support the prioritization and tailoring of risk mitigation strategies and resilience policies*".

c) Following up from the previous point, looking at certain archetypes one gets the impression that the mean ISI values are highly correlated with the mean SoVI values (Table 3). For example, the

three subclusters if cluster 3, as mentioned above, the fact that cluster 1 has the lowest ISI and one of the lowest SoVI values (albeit not the lowest), the fact that subclusters 2a and 2b have very similar ISI values, larger than subcluster 2c, and present the same pattern in their SoVI values, the fact that 8a has both the highest ISI and SoVI values, etc. However, other archetypes suggest the opposite. For example, subclusters 9a and 9b have opposite trends of ISI vs SoVI, cluster 6 has one of the lowest SoVI values but not so much regarding ISI, etc. It would be interesting and insightful to explore these and similar observations and comment about their potential meaning and significance, if any.

Response: The following text has been added to the discussion to account for the influence of social vulnerability:

"In several cases, social vulnerability is the primary driver of high ISI values, as observed in Archetypes 8a, 10a, 7 and Archetypes 2a, 2b and 2c. The latter (i.e., Archetypes 2a, 2b, and 2c) share the same geographic and demographic profiles yet differ in ISI values - with 2a and 2b showing higher ISI than 2c - due solely to differences in their SoVI scores. In this context, SoVI emerges as the only influencing factor driving ISI variation among these archetypes. Conversely, for other archetypes, demographic and geographic characteristics play a more significant role in shaping ISI outcomes. For instance, Archetype 9b presents the lowest SoVI but a relatively high ISI, which can be attributed to its high population density. In contrast, Archetype 3b shows one of the highest mean SoVI scores, second only to Archetype 9b, but results in a relatively low ISI, primarily due to its low population density and geographic remoteness."

4. L880: In lines 824-825 you discuss as well the limitations of risk modelling for the purpose of validating the archetypes. I suggest adding this to the sentence "due to the lack of fully integrated social and institutional vulnerability data, as well as limitations of risk modelling".

   Response: The sentence in line 880 has been revised as suggested.

Edits

1. L104: Add a comma after "first" (i.e., "first, broad urban and rural…".

2. L167: Erase "acknowledging that".

3. L168: Add "the" between "factors of" and "built environment".

4. L175: "Greater population" instead of "Higher population".

5. L177: Consider starting a new paragraph at "The degree of urbanisation…".

6. L242: Should it say three categories, instead of four?

7. L262-264: "minimize", "prevent" and "enhance". These three items follow the sentence "in order to (infinitive verb)".

8. L404-411: The text says "proportion of under 15 aged" while Fig. 3 says "proportion of population under 14". Is it 14 or 15? Moreover, please change "proportion of under X aged" (this is grammatically incorrect) into "proportion of population under X" or "proportion of age X and under".

9. L418, Table 2: I infer the variable "Population class" refers to section 2.4. If so, consider rephrasing as "Residential population class" or similar, to increase consistency with the main text.

10. All throughout, "urban degree" does not read well. Replace with "degree of urbanisation".

11. All throughout, “urban centredness degree” does not read well. Replace with “degree of urban centredness”.

12. L667, Fig. 9: Introducing the map with the Italian regions is a useful addition. However, the resolution of the image is quite poor.

13. L877: “empirical and external validity were only partially addressed” (not “was”).

14. L877: Broken link…? “Fare clic…”?

Response: All suggested edits have been implemented.