**Compositional novelty of plant, fungal and bacterial communities across urban habitats**

## **Editor**

I note that Reviewer 1 and Reviewer 2 have conflicting assessments on your analyses. I agree with Reviewer 1 in that there are issues to address in your revision, and R1 provides detailed ways to move forward. In addition to addressing Reviewer comments, please also include 1-2 sentences in your Abstract that provides the context of the study. These first sentences usually summarize the Introduction. Starting the abstract with the hypotheses does not provide the background of the study.

Thank you for reviewing our manuscript and for providing directions to improve it. We have prepared a new version based on the suggestions of both reviewers. Among other things, we have conducted additional tests to assess the issues raised by reviewer 1. We note that, overall, the additional tests did not change the output and the interpretation of the results. We have also revised the abstract following your advice. During resubmission, we have uploaded a track-changes version of the manuscript as a “supporting file”.

## **Reviewer 1**

This paper offers an intriguing study on assessing community novelty for multiple taxa in urban habitats. The authors use ordination methods to evaluate the differences between communities in urban habitats and those in natural and semi-natural settings, aiming to assess the novelty of urban communities. Their findings indicate that novelty is highest for plant communities and lowest for soil bacteria, with industrial vacant lots exhibiting the highest novelty among urban habitats.

Thank you for your positive evaluation and your suggestions for improving the manuscript.

The NMDS (Non-Metric Multidimensional Scaling) method should be carefully considered for ordination analysis in your case. Since NMDS employs an iterative algorithm to seek optimal solutions, the results may vary across different computational runs, leading to inconsistencies in ordination outcomes. Therefore, it is advisable to consider alternative ordination methods based on eigenvalue decomposition.

The referee is right in his description of NMDS. Following his advice, we tested the alternative of using a method based on eigenvalue decomposition: Principal Component Analysis (PCA) with Hellinger transformation. Using one method or the other did not change the outcome and interpretation of the analysis (neither the pattern of novelty across taxa, urban habitats and reference habitats; nor the statistical significance of the LMM model). Therefore, we decided to use NMDS for the final presentation in the manuscript, since NMDS makes less assumptions about data and can be applied to a wider variety of datasets. We explain this in the methods section (lines 197-216). We also note that, although in the manuscript and figures we are reporting the output of a specific set of NMDS runs, during analysis we repeated the runs many times and the results never changed significantly; this supports that the results are robust despite the iterative nature of NMDS.

Additionally, the community data used for different biological groups are incompatible: plant data are based on abundance, while fungal and bacterial data use presence/absence metrics, resulting in a lack of comparability across taxa.

The referee is right, and thus we have changed the analysis by transforming plant cover into presence/absence, so the three taxa can be compared. We note that the output and interpretation of results have not been altered by this change.

Furthermore, critical details of the NMDS analysis—such as the number of axes used, the specific distance metric applied (e.g., Bray-Curtis), and the stress value (a key indicator of model fit)—were not adequately documented in the text.

We agree, and we have expanded the methods to provide a full report of NMDS including distance metric, number of axes and stress values (lines 208-216).

The methodology mentions the use of a Generalized Mixed Model, but the corresponding results are not reported. If the authors indeed used this method, they need to provide a more detailed description of the analysis process, such as the total sample size, the distribution family used, and whether model selection was conducted.

We thank the reviewer for spotting a mistake in our text. We did not use a GLM, but a Linear Mixed Model (i.e. Gaussian family). We did this after checking that LMM assumptions were met, as described in the methods section. We did not conduct model selection. We report model results (F and p values) in the text of the results section (lines 335-350). We have included a sentence about sample size in the methods (lines 221-223).

The paper employs PCA to reduce the dimensionality of soil physicochemical variables, but subsequently selects only three factors with the highest loadings on the first three axes for further analysis. Why not use the first three axes directly for subsequent analysis? Additionally, the explanatory rates of the first three PCA axes are not described.

Following the reviewer’s advice, we now use the three axes for subsequent analysis. The only change this has caused in the results is that the first axis (corresponding to soil heavy metals) now has a significant fit with the NMDS ordination of the plant communities. We are not certain of what is meant by “explanatory rates of the first three PCA axes”. We describe the contribution of each axis to the variability in the results section (lines 235, 241, 247).

The limited sample size of 30 plots undermines the robustness and reliability of the study’s findings.

We believe that model significance indicates that the sample size was robust enough to provide statistical power to test the hypotheses of this study.

Furthermore, while natural forests were surveyed in the surrounding areas, the urban habitat sampling primarily concentrated on grassland vegetation, neglecting forest vegetation within urban contexts.

In the study city, there is no urban forest vegetation, except for natural forest remnants in the city’s fringes. The level of management does not allow for succession of urban lots to proceed to any stage dominated by trees.

The Discussion section lacks depth, merely reiterating results without offering substantive interpretation. Additionally, the section titled “Widespread Soil Contamination Across Urban Habitats” strays from the central focus of the study.

We have revised the discussion and removed the section on soil contamination, following the reviewer’s suggestion. In the discussion, there is a brief recap of the results for the benefit of the reader, but this is followed by their interpretation in light of the original hypotheses and the literature. Specifically, we believe that the discussion provides substantive interpretation regarding how our results should translate into management actions for each habitat type (see for example lines 414-419).

Several minor comments are detailed in the annotated manuscript.

We have addressed the minor comments included in the annotated PDF.

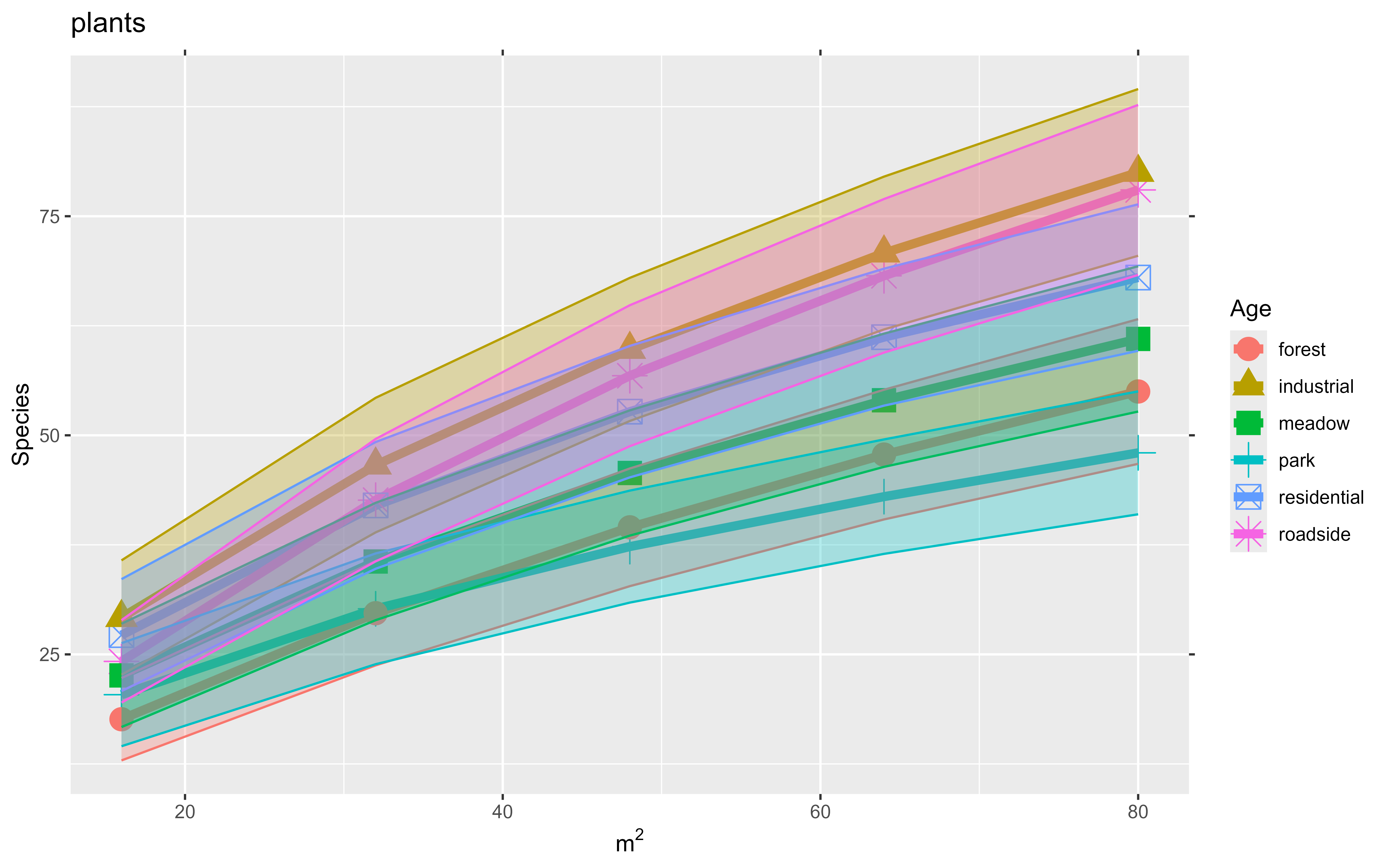
## **Reviewer 2**

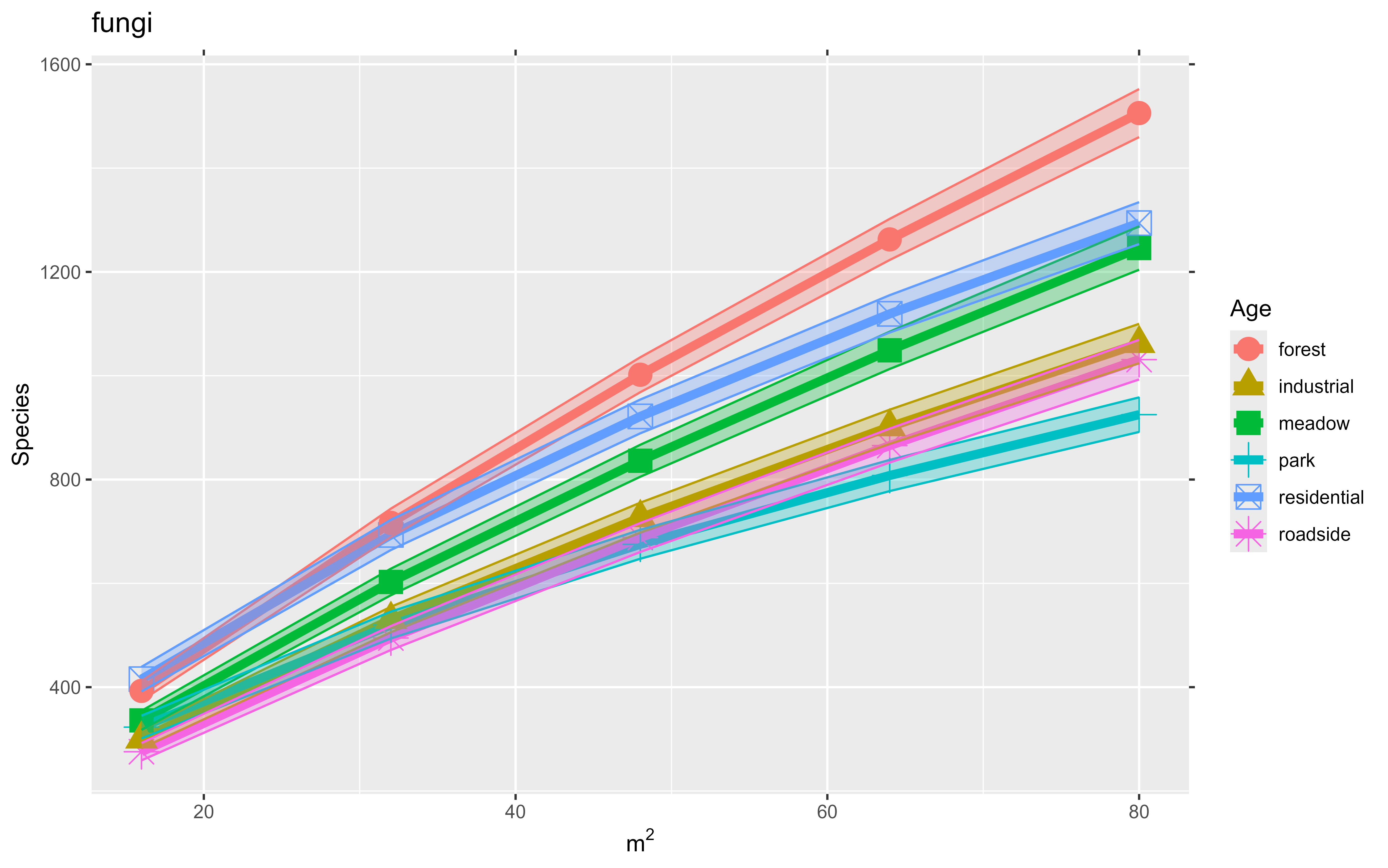
The paper describes the outcome of research conducted to characterize and better understand the diversity and compositional novelty of urban habitats with respect to their plant, fungal and bacterial communities. Although many studies have explored plant diversity associated with different habitats, this study is original in its investigation of the composition, diversity and relative novelty of not only plant communities, but also the soil microbiome in a range of urban habitats. The composition of these urban habitats is compared with remnant forest and meadow sites in the study region. Although there are relatively few investigations of bacterial and fungal community composition, the authors do cite other relevant work in the field. The authors also noted the motivation and significance of the work to urban restoration efforts in the introduction and returned to discuss their results in the light of restoration and rewilding in the discussion section. Most of the methods used in data acquisition and statistical analyses are appropriate to the investigation and results are reported clearly in the text and in the figures. The one exception is the lack of description of methods used to determine “Naturalness”, as noted below. The manuscript is quite well written. I found very few errors and had only a few questions after reading the article, as noted below.

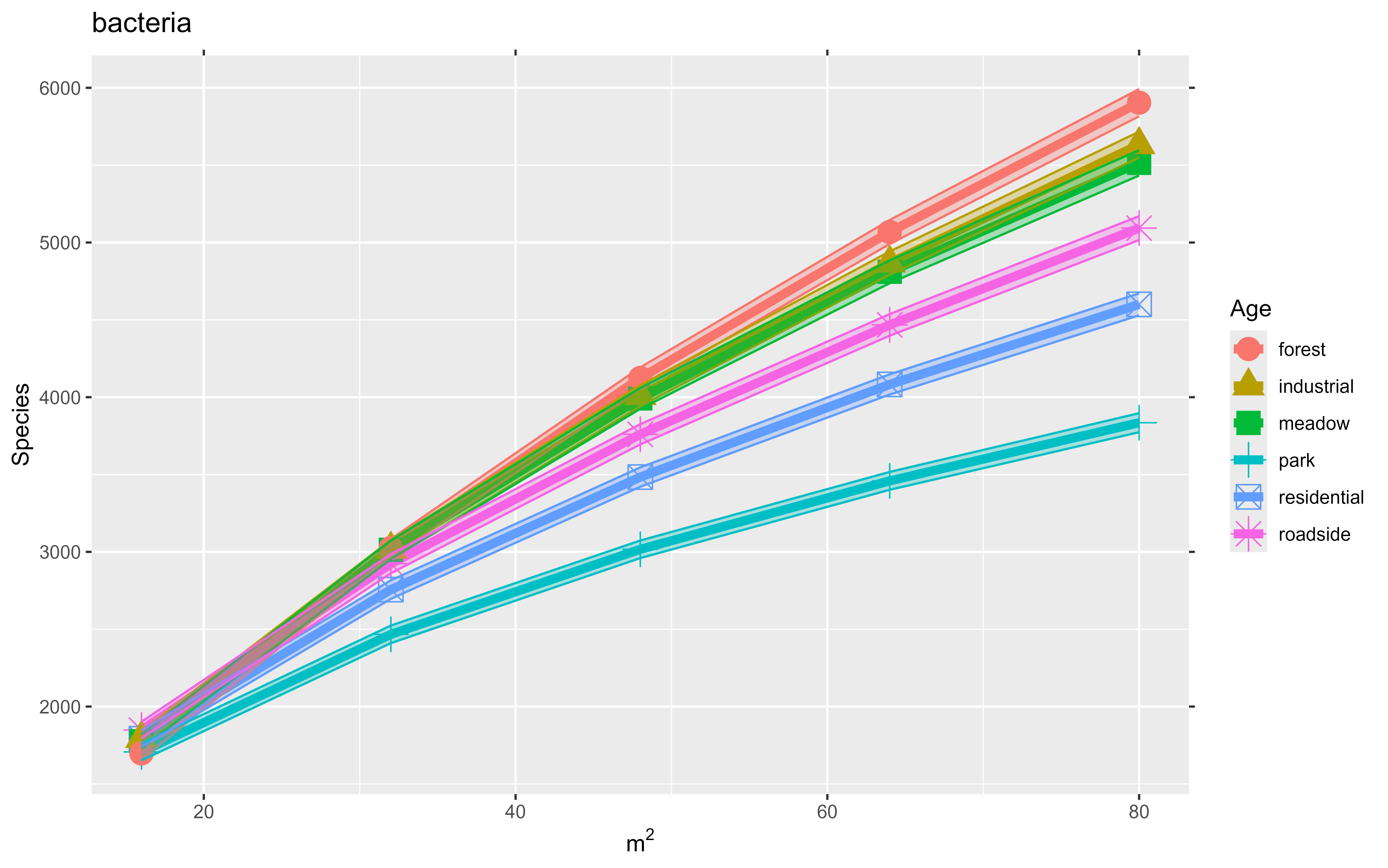
Thank you for your positive consideration of our work.

Do you think that you were able to sample a sufficiently large number of replicates for each habitat type that you are approaching an understanding of the actual alpha diversity of these habitats (i.e., where are you on the species-area curve with respect to sampling the urban habitats you have chosen to study? What might you be able to discern from your data about the potential number of species that might be found if additional sampling had been conducted?

We have checked species-area curves (see below) and our sampling is not sufficient to cover the whole species diversity of the studied habitats. The are indeed additional data to support this: an iNaturalist project on the flora of the city currently has > 2,500 research-grade observations of c. 460 vascular plant species; our survey detected 216 species. Nonetheless, we note that it was not our goal to survey the entire biodiversity of the habitats, but to compare them. With this goal in mind, we have the same number of plots per habitat, of the same size.







The methods used to determine the level of “Naturalness” of each habitat are not described. In Figure 1, I am confused by the assignment of Industrial habitats to “high naturalness” as well as “high novelty”. While these habitats were species rich, they were rich in exotic and invasive species as well as native species. This seems to me to be a state of lower naturalness, despite the high diversity. Please clarify the methods used to estimate the “degree of ecosystem self-organization” , at the very least, cite other works whose methods you are using to determine “naturalness”.

We are following the terminology by Kowarik (see for example <https://doi.org/10.1016/j.ufug.2017.05.017>). According to this terminology, a community can be rich in introduced species (“novel”) but also “natural” in the sense that the species in the assemblage have arrived with little human intervention. Of course, the species have arrived to the regional species pool by human means, but their occurrence in the local community is spontaneous. For example, the community in a park lawn is less “natural” than the community in an abandoned industrial site, because the community in the park lawn is periodically affected by seeding, watering, fertilization, etc.; while the community in the industrial site is left to follow ecological succession for several years (until the site is “cleared” out). We describe this and provide the reference in the introduction, paragraph 3 (lines 61-71). We note however that this is a conceptualization which we use for an a priori characterization of the habitats, rather than a methodology.

Page 18 line 6 – Use “As” rather than “Like” to start the sentence.

Changed as indicated.

I was taught that sentences should not begin with an acronym, such as sentences beginning with “NMDS…” or “DNA …” . However, I realize that in recent times, fewer journals seem to be adhering to this grammatical rule. I will defer to the editor to determine whether the authors should revise their sentences to avoid having to begin sentences with an acronym.

We have revised the text to avoid starting sentences with acronyms.

Figure 4 A. The color for the Residential box in the box plot seems to be incorrect, at least it does not match the color scheme in the other figures- shouldn’t it be teal blue and not purple?

The same color scheme is followed in all figures, however, in the case of Figure 4A, the residential box is so narrow that the black line which marks the median value hides the blue color of the box, and gives the impression of being purple when rendered at the final resolution.

Line 5-6 on page 14- To improve clarity, revise this sentence to : “As previously indicated by PCA, organic matter was the only soil property with a clear pattern of variation by habitat: a decrease from pre-urban to urban habitats (Figure 2B).

Changed as suggested.