Response letter on the manuscript Classification and characterization of anthropogenic plant communities at the ecoregion level

Co-ordinating Editor Comments to Author

Dear authors, I found this study interesting, even if limited to a relatively short geographical area. I believe this study can also be relevant at the European scale; therefore, I think it is suitable to be published in AVS. However, I agree with the referees that the manuscript can be significantly improved before publication.

We would like to thank the editor and the reviewers for the time invested in reviewing our manuscript and providing constructive comments. We have revised the manuscript to incorporate all the suggestions. Please find below a point-by-point response to each comment.

Referee 1

In general, the survey is interesting and has been conducted correctly in the analysis procedure and the results yielded are consistent and respond to the initial questions. For that reason I consider it can be published in AVS. However, in the comments below are suggested some changes, mostly in the introduction, and other changes are highlighted in the notes over the manuscript. I recommend to accept the paper with major changes.

Thank you for your positive evaluation and your useful comments, which we have integrated into a revised version of the manuscript.

The title does not have any indication of the geographical frame of the survey and it is relevant to indicate it to potential readers. It is in the keywords but it should be in the title.

We agree and we have changed the title to *Classification and characterization of anthropogenic plant communities in the northwestern Iberian Peninsula.*

The entire article is built up using the frame of the concept of ecoregion (Bailey 2004, 1989; Olson et al. 2001). About this there are two related questions: 1 What is the benefit of using such biogeographic frame in terms of how the results or the interpretations could be altered or deteriorated if not used? Is it a merely geographical description of the study area? If so, then, 2 What is the difference of considering the ecoregions instead the classical biogeographic units such as regions, in the sense of Good or Takhtatajan? It is important to quote that biogeographic regions also include vegetation units living within them as was established early in the 20th century by Flauhaut and Braun-Blanquet (Loidi 2021)

We would argue that the manuscript is built up based on the idea that this type of studies should be conducted using biogeographical units, regardless of the rank or classification system of the unit. In our specific example, focusing on the Iberian Atlantic territories (which rank would be "province", but see below for a discussion of the definition of the Iberian Atlantic territories and its correspondence with different biogeographical classifications) allows to study a territory that shares (1) a similar macroclimate and (2) a shared biogeographical history, including a history of biological invasions. This is especially important when studying anthropogenic vegetation, to define lists of regional neophytes and archaeophytes, which would be not possible if considering a larger area (for example, most "archaeophytes" and some "neophytes" would be considered "native" at the European level) or a smaller area or political unit (for example, species native to some parts of the Iberian Atlantic territories are some times considered "alien" in some other parts, e.g. *Pinus pinaster*). This can allow to make comparisons between invasion levels in different vegetation types in different regions, see for example Chytrý et al. (2008).

We have revised the usage of the term "ecoregion" throughout the manuscript, to avoid giving the false impression that we are specially interested in the ecoregion concept and classification by Olson et al. (2001), rather than in studying the Iberian Atlantic territories as a biogeographically-coherent unit.

It seems that the plots used are original of NW Iberia, as shown in Figure 1. The boundaries of the grey area, supposedly taken from the map of Olson et al. 2001, are quite coincident with what has been profusely named as Cantabro-Atlantic subprovince by Rivas-Martínez (Rivas-Martínez et al. 2017) and related authors such as Fernández Prieto et al. 2020, 2023. I would suggest to use these units and names to prevent from nomenclatural inflation and cite the corresponding papers (see below). If we comment the names, the name "Iberian Atlantic" is not so precise as "Cantabro-Atlantic". The former could include southern Portugal (Algarve or Vicentine Coast), which is not at all the case. To use the term Cantabrian or Cantabrico is particularly precise to indicate the geographic position of this territory. Additionally, it would be necessary to make a more formal description of the "Iberian Atlantic ecoregion" and discuss the differences from this ecoregion with the "Cantabro-Atlantic subprovince".

We agree that the definition and name of the studied biogeographical unit requires further clarification. First, we must make it clear how we define our study region: "the territories with a temperate climate in the north-western part of the Iberian Peninsula". In other words, we are following precisely the definition by Fernández Prieto et al. (2020) in *Naturalia Cantabricae* 8(2): 30-37. More specifically, we are following the geographical limits presented in Figure 1 of Fernández Prieto et al. (2020).

The concept of this biogeographical unit as presented by Fernández Prieto et al. (2020) is essentially equivalent to the *Cantabrian Mixed Forests ecoregion* by Olson et al. (2001). The major difference between the two concepts is in the southern borders of the region, which are much more precisely drawn in Fernández Prieto et al. (2020).

However, the biogeographical unit does not correspond with Rivas-Martínez's *Cantabro-Atlantic subprovince*, as there are two deviations from that concept. First, our study region also includes Rivas-Martínez's *Orocantabrian subprovince*. Second, Rivas-Martínez's *Cantabro-Atlantic subprovince* includes a larger portion of French territory than our study region. Moreover, the *Cantabro-Atlantic subprovince* was not a stable entity in the third versions of Rivas-Martínez's map. For example, in the first edition, it extended in France from the French Basque country in the south to Calais in the north. In the second edition, the extent in the north was reduced to the Cotentin Peninsula in Normandy. Regarding the third edition, we have been unable to check it, since it is not available in digital format. In sum, the more precise correspondence between our study region and Rivas-Martínez's biogeographical classification would be "the Iberian part of the *European Atlantic province*". We include this equivalence in our methods section, in the subsection "Biogeographical territory under study".

Finally, we agree that, from a geographical point of view, "Iberian Atlantic" can refer to the Iberian coast down to Gibraltar, but the term "Atlantic" is used here in a bioclimatic sense, in accordance with the definition by Fernández Prieto et al. (2020) and the biogeographical unit names by Rivas-Martínez (cf. *European Atlantic province* in Rivas-Martínez et al. (2014)) and the European Union's Natura 2000 framework (cf. *Atlantic biogeographical region*, https://www.eea.europa.eu/en/analysis/maps-and-charts/biogeographical-regions-in-europe-2).

In any case, to avoid confusion, we have revised the manuscript's section "Biogeographical territory under study" [page 7, lines 8-17] to better explain the study area and its equivalence in different classifications. We have also revised the whole manuscript to avoid using the term *Iberian Atlantic ecoregion*, which mixes the concepts *Cantabrian Mixed Forests ecoregion* by Olson et al. (2001) and *Iberian Atlantic territories* by Fernández Prieto et al. (2020). Now, we use *Iberian Atlantic territories* in accordance with Fernández Prieto et al. (2020). We have also added the reference to Fernández Prieto et al. (2023) which further supports the definition of the studied biogeographical unit.

One of the main results of this survey is the selection of 28 alliances (appendix S4) from an initial set of 38 that could be present in the Iberian Atlantic ecoregion. So, I understand that there are 10 alliances have not been detected. It would be interesting to list them somewhere (in appendix S2, for instance) and make clear that no data from them have been found in the datasets.

We list them in the methods [from page 10, line 21 to page 11, line 4] and the discussion [page 19, lines 14-22]. These alliances were present in the dataset in the sense that the relevés had been assigned to syntaxa belonging to these alliances by the original authors of the relevé. However, the alliances were not supported by our numerical classification, i.e. their relevés were assigned to another alliance, or to the noise group.

Concerning the weakly represented alliances. This is inevitable if we start with a definition of the study area prior to surveying its plant communities. There will be always some types which are found in the extreme or border of its geographical area or ecological amplitude and are, thus, weakly represented. It is an inherent problem of such a territorial planning and has to be conveniently considered. Other possibility is to choose the vegetation alliances and study them in all their geographic amplitude. This would entail that their geographical territory would be diverse but the survey would bring more accurate results about their character and differences.

We agree that the risk exist, although we believe that this risk is lessened by focusing on a biogeographical unit rather than in e.g. a political division. Furthermore, there were alliances that are only found in the extremes (i.e. the contact with the Mediterranean region) of the study area, but still they were supported by the numerical classification. We believe that the fact that the missing alliances were not supported by the numerical classification indicates that the relevés that had been assigned to these alliances are poor

representatives of the vegetation types, and more likely transitional communities at the extreme of the alliance distributions.

In S2 author's names should be corrected and homogenized e.g. Lence Paz is C. or Carmen, all the patronimic names such as Pérez, Fernández, etc„ have an accent but have instead a strange signal in the involved letter. It is Rivas-Martínez, not Rivas Martínez; Is it Romero or Romero Buján?, etc.

In Appendix S2, author names are provided exactly as they are stored in the SIVIM database. This is necessary to ensure cross-referencing between our dataset and the SIVIM database. We have specified this in the heading of the Appendix S2.

In the figure 1 the area does not include any French terrotory [. . .] In the maps of figure 1 there is no French territory included

The figure includes a portion of French territory in the Pyrénées-Atlantiques department, following the biogeographical borders shown by Fernández Prieto et al. (2020) in their Figure 1, although our figure may be to small to fully appreciate this. We have revised the mention in the abstract to make clear that we are speaking about SW France.

Anthropogenic syntaxa have a larger biogeographical scope than other vegetation types, particularly grasslands and scrub, but they have narrower distribution than others such as aquatic vegetation. On the other hand, they have been described by traditional phytsociology after fine ecological variations. If there is inflation it should be somhow documented.

We have removed this sentence from the revised version.

There are several papers on this topic published in the area by Loidi and other authors, they are quoted in the adjointed sheet

Thank you for these references, we have incorporated them into our manuscript.

classification at the alliance level ..

We have updated the text as suggested.

In this paper only forest and mantle communities are dealt, no nitrophilous vegetation.

We cite this reference as the basis for defining the borders of our study biogeographical territory, we have revised the text to make this clear.

In these disturbed communities, geophytes, with some exceptions, use to be scarce. Do you think neccesary to highlight their proportion more than other life-forms? It seems to me that the main point is to describe the relative importance of hemicriptophytes in relation to therophytes as they become dominant in their respective alliances.

We agree that, in most of the alliances, only the variation of hemicryptophytes vs. therophytes is relevant. However, the presence of geophytes is an important part of the description of some of the alliances (e.g. see the description of *Allion triquetri* by Mucina et al. (2016): *Mesic nitrophilous geophyte-rich fringe vegetation of the Western Mediterranean*). Furthermore, the variation in geophytes explained a significant part of the

overall variance in an exploratory PCA. For these two reasons, we decided to highlight the proportion of geophytes in our manuscript.

This figure is difficult to read. Point clouds overlap and the information is partially not visible. I suggest to use two figures for the same PCA, each one representing a half of the alliances selected in form that overlpings are avoided. This could also allow to use larger dots which would be more easily identified by the color.

Thank you for your suggestions to improve Figure 7. We have provided a new version of the figure, where each vegetation class is represented separately.

It should be cited as: Castroviejo, S. (coord. gen.). 1986-2020. Flora iberica. Real Jardín Botánico, CSIC, Madrid [and other comments regarding the references]

Thank you for checking our references, we have corrected them as needed.

Referee 2

This paper presents a classification and characterisation of the anthropogenic vegetation alliances of the Iberian Atlantic ecoregion in Europe. The study area is not very large, but I think that the paper is a valuable contribution to the knowledge of this vegetation type in Europe, and it can be considered suitable for AVS after some revision.

Thank you for your positive evaluation and suggestions.

The authors claim that (1) there is a need for a numerical classification and synthesis of the full spectrum of anthropogenic vegetation, and (2) an appropriate scale to conduct such study is the ecological region. While I strongly agree with point 1, I'm less convinced about point 2. Why should a revision at the ecoregion level be superior to a continental pan-European synthesis? I think, both is necessary, and the authors should acknowledge this fact.

We fully agree that a pan-European synthesis would be a worthwhile goal, and we never intended to suggest the contrary. In fact, we would be very happy of any research efforts on continental studies of anthropogenic vegetation. We have added a new final sentence to our conclusions: *We hope that our synthesis of the anthropogenic communities of the Iberian Atlantic territories serves as a contribution towards a pan-European synthesis of human-made vegetation, a synthesis that could shed light on the coexistence between humans and plants.* We also added a similar closing statement to the abstract. We have also modified our text to avoid giving the impression that we are suggesting that the ecoregion (or a similarly ranked biogeographical unit) is the *only* appropriate scale at which one can produce a synthesis of anthropogenic vegetation.

The title is a bit too general. Please consider to add a subtitle such as "a case-study from SW Europe", or something similar.

We agree and have changed the title to *Classification and characterization of anthropogenic plant communities in the northwestern Iberian Peninsula.*

The merits of the applied two-step approach (first a Twinspan classification, then a "semi-supervised" noise clustering) remain unclear to me. More details are needed to fully understand what the authors have done. How many Twinspan clusters did you choose, and how did you determine the optimal number of clusters? How did you match the individual Twinspan clusters with EVC alliances? A crosstab showing the Twinspan clusters and the original alliance assignment should be presented. In the noise clustering, no additional clusters were detected, but plots were allowed to be re-assigned, and a considerable part of the data (17%) remained unclassified in a "noise group". Still, the "semi-supervised" classification is a numerical cluster algorithm. In my opinion, a pure unsupervised classification with subsequent interpretation of the result would have been more straightforward. Re-assignment of plots should be done on the basis of well-defined diagnostic species or using formal definitions, not by applying two numerical algorithms one after another. Or did I miss something important?**

We agree that we may have done a poor job of explaining our classification approach. The main aim of our classification was for it to be consistent with regional phytosociological expertise, and to match that expertise with current EuroVegChecklist alliances (Mucina et al. 2016), not to create a new classification. To achieve this, one advantage of using the SIVIM database is that a majority of the plots had been assigned to associations or alliances by the original authors (Font et al. 2012). However, there could be misclassification due to unclear concepts or human error. Therefore, we used TWINSPAN to revise the classification by the original authors, to define the number of clusters that better aligned with our syntaxonomical checklist, and to remove outliers or misclassified plots. The obtained clusters therefore provided an agreement between phytosociological and numerical classification, and contained only those plots that better represented each alliance. These clusters were then used to assign other plots to the most similar cluster. We think semi-supervised classification is an optimal procedure to reach an agreement between existing and revised classifications (following Tichý et al. (2014), the aim is "preserving the good old units and searching for new ones").

We found that this important point was not clearly explained in the text or the abstract, so we revised the text accordingly [from page 9, line 19 to page 11, line 9]. In the same lines, we have also provided more details on the TWINSPAN procedure as requested by the reviewer.

While revising these aspects, we have repeated all the classification analysis. During the new classification, we found and fixed some errors in the syntaxonomical checklist, as well as in the assignation of TWINSPAN clusters. This has resulted in an improved classification into 25 alliances.

The authors conclude that ten alliances previously reported from the study area do not occur there. This seems a bit hasty. Maybe, these alliances were only represented by a few plots, so they were not detected by the numerical methods.

We note that some alliances were supported by the methods, despite being extreme (ecologically or geographically) for the study region and/or being represented by a few plots. Therefore, we believe that the fact that the missing alliances were not supported by the numerical classification supports that the relevés that had been assigned to these alliances are poor representatives of the vegetation types, and more likely transitional communities at the extreme of the alliance distributions. Nonetheless, we have modified the text to be less conclusive about their absence, by adding the following sentence to the discussion: *Of course, the fact that these nine alliances are not supported by our numerical methods only indicates that they are poorly represented in our dataset, and should not be taken as absolute proof that they are absent from the Iberian Atlantic territories.*

Most importantly, a synoptic table of the alliances is missing. I consider a table showing the constancy of all species in the 28 alliance, with the diagnostic species highlighted, an absolute prerequisite for the paper to be acceptable. A shorted table in the main text would be nice, but there should be a least the full table as supplementary information.

We have included the full synoptic table as Appendix S5.

p5, l11: "non-overlapping definitions" ... different, but certainly overlapping definitions!

We agree, and we have changed it to *"not fully equivalent"*.

p9, l15: Why did you use 15% as cut level, and not 5%? The latter would correspond to the border between Braun-Blanquet class 1 and 2, while 15 cuts the Braun-Blanquet class 2 in the middle. Which cut level then corresponds to Braun-Blanquet class 2?**

Our transformation of Braun-Blanquet values follows the standard of assigning intermediate cover values. Thus, the Burble value 2 (representing 5-25% cover) is transformed to 15% and the value 3 (25-50%) is transformed to 37.5%. Although we could use more approximate pseudospecies levels, the results of TWINSPAN will be the same when the transformation values are selected within the original cover ranges.

p10, l5: Please clarify if the same TwinspanR settings were used as before.

Yes, and we have added this information to the manuscript.

p20, l10: "which are mostly differentiated by dominant vegetation height" ... Is this a new result of your study, or has this been mentioned in the literature (if so, please provide a reference)?

Yes, this is a traditional part of the alliance descriptions, but in any case we have removed the sentence from the manuscript.

p21, l12: "neophytes tended to be in higher proportions in vegetation types associated to wet conditions (Bidentetea)" ... I would have expected an even higher portion of neophytes in the Senecionion fluviatilis. Maybe, plots of alluvial forb communities dominated by neophytes are underrepresented in your dataset?

Our *Senecionion fluviatilis* plots are certainly rich in alien species, specially some communities invaded by *Arundo donax*, in which this species has a mean cover of 74%. However, in these communities, *Arundo donax* is accompanied by many native species (mean = 12 native species per plot) and so the plot-level proportion of neophytes seems low (as it is just the proportion, not weighted by species cover). That being said, it might be possible that plots with neophytes are generally underrepresented across the dataset, because of a sampling bias towards "natural" communities. But this underrepresentation is not necessarily higher in *Senecionion fluviatilis* vs. other alliances.

References: In several references the page numbers are missing!

We have revised the references.

Appendix S3: What is the value for dominant species? Constancy? Mean cover value? Something else? Please clarify.

We have added to the Appendix header the information on how we calculated the values for diagnostic, constant and dominant species.

References

Chytrý, M., Maskell, L.C., Pino, J., Pyšek, P., Vilà, M., Font, X., & Smart, S.M. 2008. Habitat invasions by alien plants: a quantitative comparison among Mediterranean, subcontinental and oceanic regions of Europe. *Journal of Applied Ecology* 45: 448–458.

Fernández Prieto, J.A., Amigo, J., Bueno, A., Herrera, M., Rodríguez-Guitián, M.A., & Loidi, J. 2023. Bosques y orlas forestales de los territorios atlánticos del Noroeste Ibérico. *Guineana* 23: 1–240.

Fernández Prieto, J.A., Amigo, J., Bueno, Á., Herrera, M., Rodríguez-Guitián, M.A., & Loidi, J. 2020. Justificación de una nueva delimitación de los territorios iberoatlánticos peninsulares. *Naturalia Cantabricae* 8: 18–24.

Font, X., Pérez-García, N., Biurrun, I., Fernández-González, F., & Lence, C. 2012. The Iberian and Macaronesian Vegetation Information System (SIVIM, www. sivim. info), five years of online vegetation's data publishing. *Plant Sociology* 49: 89–95.

Mucina, L., Bültmann, H., Dierßen, K., Theurillat, J.-P., Raus, T., Čarni, A., Šumberová, K., Willner, W., Dengler, J., García, R.G., Chytrý, M., Hájek, M., Di Pietro, R., Iakushenko, D., Pallas, J., Daniëls, F.J.A., Bergmeier, E., Santos Guerra, A., Ermakov, N., Valachovič, M., Schaminée, J.H.J., Lysenko, T., Didukh, Y.P., Pignatti, S., Rodwell, J.S., Capelo, J., Weber, H.E., Solomeshch, A., Dimopoulos, P., Aguiar, C., Hennekens, S.M., & Tichý, L. 2016. Vegetation of Europe: hierarchical floristic classification system of vascular plant, bryophyte, lichen, and algal communities. *Applied Vegetation Science* 19: 3–264.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V., Underwood, E.C., D'amico, J.A., Itoua, I., Strand, H.E., & Morrison, J.C. 2001. Terrestrial ecoregions of the world: a new map of life on earth: a new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience* 51: 933–938.

Rivas-Martínez, S., Penas, Á., Del Río, S., Sánchez Mata, D., & Díaz González, T.E. 2014. Biogeographic map of Europe. Third concise advance. *XXIV Jornadas Internacionales de Fitosociología y I Congreso de la Sociedad Española de Geobotánica*

Tichý, L., Chytrý, M., & Botta-Dukát, Z. 2014. Semi-supervised classification of vegetation: preserving the good old units and searching for new ones. *Journal of Vegetation Science* 25: 1504–1512.