Dear Pierre Labadie,

Please find attached the manuscript that my co-authors and I would like to publish in *PCI Ecotoxicology and Environmental Toxicology*. We would like to thank you for the opportunity to resubmit our manuscript and are grateful to you and the reviewers for the helpful comments.

We modified the manuscript in line with all the reviewers' comments. Please find below a point-by-point response to these comments along with a summary of the corresponding (Lange, 1969)revisions made to the manuscript. Our responses are highlighted in grey as are our changes to the text.

The data and the scripts are available on a depository (In.Do.RES: https://doi.org/10.48579/PRO/TX0PU9) and they are under embargo until publication. The DOI will be made public on BioRXiv when the article is accepted. The article and the supplementary material are available on BioRXiv. The conflict of interest disclosure and funding are given at the end of the article.

We hope that we have fully addressed all the changes requested by the reviewers. We look forward to hearing from you in the near future.

Yours sincerely,

Florence D. Hulot, on behalf on my co-authors

Review by anonymous reviewer 1, 03 Jan 2024 12:09

Our sincere thanks to reviewer 1 for his proofreading and precise comments, which helped improve the manuscript.

General comments:

The authors aim at examining the effects of various classes of micropollutants (e.g. trace elements, pharmaceuticals, pesticides, PAH) from urban and/or agricultural watersheds on the abundance and taxonomic structure of benthic macroinvertebrate assemblages of peri-urban ponds based on a 2 year-survey.

For this purpose, the authors set up various inferential and descriptive statistical analyses relating the presence/absence or abundance of taxa and taxonomic diversity metrics (i.e. alpha & beta diversities) to contaminant concentrations in the water and sediment and/or to a description of land use in the immediate vicinity (i.e. a radius of 100 m) of 12 ponds of various size (64 – 5,108 m2).

This article is interesting for the scientific community. The authors put a lot of effort into the manuscript, which is clearly written and well-structured. The core objectives are attractive; tested hypotheses are clearly defined (lines 110-114) and the information flow makes the paper a quite attractive piece of work. The literature review is comprehensive and may be useful to both ecologists and ecotoxicologists. The “material and methods” section is presented in a rather concise and reproducible way, but some major methodological choices – and their limits - need to be more deeply discussed (e.g. the definition of morphotaxa, the different sampling strategies adopted). The statistical design seems globally sound (even if not explicit enough for several steps; see my “main comments” below). The results are rather clearly described. Perspectives could be more developed; e.g. why not exploring a more functional approach based on “species traits”? Note also that a biotic index, based on macroinvertebrates (BECOME-I#15) has been recently and specifically developed for small standing waters in mainland France, and could be used as supplementary variable in some of the analyses (Labat & Usseglio-Polatera, 2023; table 4, p. 10; reference provided below).

I have detailed my main concerns in the following "main comments".

Main comments:

Abstract:

The abstract is rather concise and clearly present the main objectives (exploring local and regional macroinvertebrate spatiotemporal diversity and understanding the effects of contaminants on the abundance and structure of pond invertebrate assemblages) and the main findings of the study.

Following changes made to the text, we have adapted the summary (lines 26 – 33 and 38 & 40).

Introduction:

The introduction is quite interesting and correctly structured.

Lines 49-58: The authors wisely emphasise the importance of pond size as an element of vulnerability to anthropogenic impacts, but curiously do not include this size among the environmental variables that could explain between-ponds variations in the response of taxonomic structure or taxonomic biodiversity of invertebrate assemblages to anthropogenic pressures (e.g. in the MFA). Why not?

We have added the surfaces in the Materials and Methods section (lines 132-133). These surfaces were estimated on a single date and do not correspond to the actual surface areas of the ponds during sampling, which can be quite variable. For this reason, it is not relevant to include them in analyses.

Lines 64-66: But probably largely depending on their diversity in aquatic (meso)habitats?Has the within-pond diversity of (meso)habitats (*sensu* Armitage & Pardo, 1995) been considered in the study to potentially explain variation in invertebrate assemblage in relation to anthropogenic impacts?

We agree on the importance of meso-habitat diversity, which is why we sampled two points in each pond. However, our aim was to study the effects of contaminants and to monitor a large number of them, taking into account their diversity. Given the large number of analyses, we pooled samples within each pond. We did not design a protocol to study meso-habitat diversity, as this was not our objective.

<u>Cited reference</u>:

Armitage P.D. & Pardo I. 1995. Impact assessment of regulation at the reach level using macroinvertebrate information from mesohabitats. *Regulated Rivers: Research & Management*, 10, 147-158. hDps://doi.org/10.1002/rrr.3450100210

Lines 67-81: While contaminants are clearly one of the main threats to small stagnant waters, the physical degradation of littoral habitats can also be a major source of alteration to invertebrate assemblages (cf. some of the sources of degradation of ponds – e.g. trampling, tractor washing directly in the pond, etc - described in the "characteristics of the peri-urban environment" sub-section of the discussion; pp. 531-548).

We do agree on the importance of littoral quality but it is not easy to measure and monitor. Furthermore, it’s not the objective of our study.

Material & methods:

Lines 127-129: Why a buffer zone of 100 m radius has been uniformly considered to describe the proportion of agricultural and urbanised land around each pond, whatever the size of the pond, given the wide range of sizes of the ponds selected? This point should be justified a little more. First, the effects of adjacent land use on both water quality and biotic assemblages have already been considered as strongest at larger distances (e.g. between 200 m and 4000 m; see Houlalan & Findlay, 2003; 2004; Houlalan et al., 2006). Second, would it not have been more appropriate to adapt the size of the buffer zone considered, taking into account the size of the pond, as was done - for example - by Labat & Usseglio-Polatera (2023)?

To justify our choice, we added the following sentence (lines 135-136): “As the aim was to study the effect on water quality of land use in the vicinity of ponds in a fairly fragmented landscape, we used a short-radius buffer.”

Cited references:

Houlahan J.E., Findlay C.S. 2003. The effects of adjacent land use on wetland amphibian species richness and community composition. Canadian Journal of Fisheries and Aquatic Sciences, 60, 1078–1094. https://doi.org/10.1139/f03-095

Houlahan J.E., Findlay C.S. 2004. Estimating the 'critical' distance at which adjacent landuse degrades wetland water and sediment quality. Landscape Ecology 19, 677–690. https://doi.org/10.1023/B:LAND.0000042912.87067.35

Houlahan J.E., Keddy P.A., Makkay K., Findlay C.S. 2006. The effects of adjacent land use on wetland species richness and community composition. Wetlands 26, 79–96. https://doi.org/10.1672/0277-5212(2006)26[79:TEOALU]2.0.CO;2

Labat F., Usseglio-Polatera P. 2023. A new bioassessment multimetric index (BECOME) and diagnostic tool (BECOMEd) for small standing waters. Ecological Indicators, 154, 110831. https://doi.org/10.1016/j.ecolind.2023.110831

Lines 145-147: The field campaigns appear to have been considered homogeneously. Why didn't the authors try to deconvolute the seasonal effect (June vs. September) from the interannual effect (2016 vs. 2017)?

Our focus is on contaminants and we measured a lot of them twice a year during two years. The number of repetition (2) is not high enough to meaningfully deconvolute seasons and years.

Lines 159-166: It is rather strange to change the invertebrate sorting strategy so radically in the middle of the sampling plan. Have the authors examined the potential impact of such a modification on the resulting faunal list, insofar as it may constitute a potential confounding factor for robust identification of interannual variation in assemblage structure.

As specified in the manuscript (lines 180-181), we changed the protocol to speed up sampling. *A posteriori*, we agree that this was not a good idea. The few tests we have carried out show that there is no strong year effect. With a few exceptions, the distribution of ponds in the RDA remains clustered.

Lines 166-168: There is something inconsistent in the presentation of the number of samples collected: "12 ponds x 2 points x 2 years x 2 seasons, minus 6 cases" equals 90 samples (and not 44 as indicated in the MS). Even if the invertebrates collected at the two sampling points of a given pond have been pooled (as indicated in lines 223-224), the resulting number of samples does not correspond to 44. Please, explain more clearly what has been done.

It's just a miscalculation! Corrected (lines 185-187).

Lines 180-183 and tables S4/S5: I am really embarrassed by the definition of "morphotaxa" in this document. As indicated by the authors, organisms that could not be identified at the species level were identified at a higher taxonomic level, while adding a numerical suffix when more than one species was present (e.g., Microvelia sp.2)." I understand the process and validate it if applied at given – homogeneous taxonomic level (e.g. the genus level, as for the example provided). However, this process has been applied by the authors at different nested levels of the systematics. For example (cf. Table S4), for the dipteran family “Chironomidae”, nine morphotaxa have been defined at the “family” level (Chironomidae_01 to Chironomidae-11 in table S4), six morphotaxa have been defined at the Chironomini or Tanytarsini “tribe” level (two tribes within the Chironomidae family; Chironomini_01 to Chironomini_04 and Tanytarsini_01 to Tanytarsini_02) and six morphotaxa at the Chironomus “genus” level (one of the genera of the Chironomini tribe; Chironomus_sp._01 to Chironomus_sp._06). Moreover, in this family, additional morphotaxa have been also defined at the “sub-family” level (e.g. Tanypodinae_01 to Tanypodinae_11; Orthocladiinae_01 to Orthocladiinae_04). How be sure that organisms belonging to the same taxon have not been included in many different (nested) morphotaxa of the faunal list, especially if difficult to identify at species level?

What could be the impact of this series of methodological decisions on the quantification of pond assemblage diversity defined on taxonomic criteria (e.g. richness)?

Moreover, the authors have indicated that “when it was not possible to link the different stages (larvae, nymph, adult) to the same species, they were assigned to different morphotaxa”. In my opinion, this second major decision adds a further level of uncertainty to the quantification of invertebrate assemblage diversity. An indirect effect is also that the authors implicitly give more weight in the data - and therefore in the analyses - to "poorly identified" taxa, insofar as they may constitute several morphotaxa (corresponding to several stages of development or because they may contribute to taxa of different systematic levels - e.g. Diptera/Chironomidae /Chironomini/Chironomus) to the detriment of taxa “reliably identified” at the specific level (e.g. Hydrometra stagnorum, Ochthebius minutus or Limnoxenus niger in Table S4 and S5), which is not ecologically justified. What impact did these authors' choices have on their ability to interpret "taxonomic richness" or "taxonomic diversity" based on such morphotaxa in ecological terms? This crucial point needs at least to be deeply discussed.

We agree that this has an impact on results and analyses. As our objective is to compare ponds at 4 dates, we believe that this way of considering specimens does not alter our analysis, as we have always kept the same methodology. We do not compare diversity indices with ponds other than those in our study. In addition, we have kept the presence-absence analyses in the main text and discuss them. Analyses with abundances, other than diversity indices, may contribute to increased bias due to over-represented groups; they are now in the appendices.

We have added a warning lines in the Material and methods section (lines 264 – 266).

Last point on the taxonomic list: I’m not sure that Daphniidae (three morphotaxa in table S4), ostracods (six morphotaxa of the Podocopida order in table S4) and copepods (two morphotaxa) can be considered as “macro” invertebrates, and have been efficiently sampled with a 1 mm mesh size pond-net.

Indeed, they are not macro-invertebrates and we redid all the analyses without considering taxa (Daphniidae, ostracods and copepods).

Lines 230-231 and table 1: I'm not sure to clearly understand which kind of ANOVA has been performed: Are the results provided in table 1 correspond to (i) "one-way ANOVAs" independently testing the "pond" effect and the "field session" effect (but two numbers of degrees of freedom should have been provided by ANOVA) or (ii) "twoway ANOVAs without replication" simultaneously testing both effects (but three numbers of degrees of freedom should have been provided by ANOVA).

We did two way ANOVAs with additional effects of "pond" and "field session". It is now specified on line 270.

Lines 239-241: The authors have explored two different analysis strategies for measuring pond dissimilarity based either on taxon abundances (with Hellinger coefficient) or occurrences (with Jaccard coefficient). After reading the MS, I wonder whether it is really necessary to present both approaches in the main manuscript. Examine the possibility to move the presentation of one approach in the Supplementary Material.

We moved the approach based on abundances in Supplementary section.

Lines 250-255 and 265-276: Using both (i) a Multi Factorial Analysis (MFA) to explore the relationship between the "environmental parameters" (in fact, land-use descriptors only) and macroinvertebrate-based diversity metrics and (ii) a Redundancy Analysis (RDA) to identify environmental parameters (in this case, physico-chemical parameters measured in water or sediment) best explaining variation in taxonomic structure of assemblages should be briefly justified. Why not applying only one of these approaches to the whole set of environmental variables (i.e. both land use and physico-chemical descriptors)?

MFA typically deals with datasets where variables from different sources (here, taxonomic and land use) are organized in groups, highlighting a common structure of all the groups, and the specificity of each one. In your case, the three variables describing the land use (the second group of variables; i.e. "%urbanised areas", "%agricultural areas" and "%grassland and forest") do not make it possible to distinguish the surveys carried out on the same station over the 4 campaigns (since their description is certainly similar), which undoubtedly explains the particular position (aligned) of the 4 samples carried out on each pond in the first factorial plane of the MFA (fig. 4c), given the levels of correlation of the corresponding variables with the first two axes (fig. 4b) and also the particular shape of the associated confidence ellipses (completely flat).

Consequently, is MFA objectively adapted to the analysis of the data at the scale of the sampling unit (i.e. pond x field session)?

For MFA, interesting information can be found in Abdi et al. (2013).

Cited reference:

Abdi H., Williams L.J., Valentin D. 2013. Multiple factor analysis: principal component analysis for multitable and multiblock data sets. WIREs Computational Statistics, 5, 149-179. 2013. https://doi.org/10.1002/wics.1246

Thank you for your comment, as we couldn't understand why the ellipses were flattened! As the land use did not change between sessions, we preferred to withdraw the analysis. We ran a linear mixed-effect model as specified on lines 297-299.

Results:

Line 290: More probably “morphotaxa richness” and not “species diversity”.

Corrected.

Lines 360-363 and figure 3: How may be ecologically interpreted morphotaxa gains and losses if morphotaxa can be different instars in a given taxon?

TBI analyses are based on morphotaxa abundances. As explained above, we deleted this analysis.

The X-axis and Y-axis labels (i.e. “species” gains or losses) are somewhat abusive.

The “solid line” is not green but red and the “dashed line” is not red but green. As a result, it becomes difficult to interpret the figure based on its legend. Adapt the figure caption.

The authors have indicated in the figure legend: “The position of the red line below the green line in all comparisons shows that morphotaxa losses dominated the gains”. Can the lower panels (C1-C2 and C3-C4 sessions) be considered as illustrations of the seasonal effect and the upper panels (C1-C3 and C2-C4 sessions), with greater differences, as illustrations of the interannual effect (including the change in sampling strategy)?

Lines 380-392 and Figure 4: See my comment above (on lines 250-255 and 265-276).

TBI analyses are based on morphotaxa abundances. As explained above, we deleted this analysis and the figure.

Lines 399-434 and figure 5:

Why the variable arrows on the biplots (i.e. the standardized canonical coefficients?) don’t have exactly the same length on the panels a (morphotaxa) and b (ponds) – and similarly on the panels c and d – even if these panels correspond to the same analysis?

For readability, it would have been more judicious to use the same scale for the abscissas on the one hand, for the ordinates on the other, in the two figures corresponding to the same analysis.

To make the figures, we used two different projection systems. We've redone the figures using the same projection system, and the arrows are now the same length (Figure 3).

Discussion:

Lines 444-448: It is difficult to know the relative contribution of the effect of the change in sampling method and the effect of inter-annual variability (e.g. related to difference in anthropogenic pressures or weather conditions).

We discuss this point lines 402 – 406.

Lines 475-482: It seems difficult for authors to avoid discussing the relevance of their assessment of alpha and beta diversity, given the choices made in defining morphotaxa (see my comments on Lines 180-183 and tables S4/S5).

Our comparisons focus on ponds as a function of pollutant concentrations. Our morphotaxa definitions (lines 203-204) are the same for the 4 sampling sessions. So it does not interfere with the analyses.

Non-capture of pollution-sensitive or rare taxa is also worth discussing. Which taxa were expected in these ponds? Was the sampling method (without scraping the bottom substrate) really adapted to their capture?

We did not expect specific taxa and we observed benthic organisms. The sampling method, with infinite signs, resuspended sediments and benthic individuals that were collected in the sampling net.

Lines 531-560: Given the variety of pond uses, the size of the pond was undoubtedly an important factor to consider (local uses are likely to have more or less impact depending on their size).

We did not analyse the effect of pond size as we have only one measure, i.e. we did not measure the size of the pond at each sampling session. Moreover, we have one small pond, one big pond and the others are of roughly the same size.

Probably, the metrics used to describe diversity, only based on morphological/taxonomic criteria, were not optimal to highlight significant relations between land use and morphotaxa diversity variations. Why did the authors base the analysis solely on taxonomic criteria and not (also) on functional criteria, e.g. via the analysis of biological traits in macrobenthic communities? When only "common" taxa are captured, there may be more interesting information to draw from examining variations in the selection of biological adaptations within invertebrate communities.

We first wanted to analyse diversity. The analysis of traits could be the subject of another work, bearing in mind that the article already contains a lot of data and that it is difficult not to give results on diversity indices.

Minor points:

- Line 106: “different” or “differed”?

Corrected.

- Lines 132-133: Is there a form of gradient from A to L or have these codes been allocated randomly?

Letter allocation is random. This is now specified on line 143.

- Line 150: “net pond” or “pond-net”?

Corrected (pond-net) (line 164).

- Line 177: Change “Poisson” in “Poisson, 1957”.

Corrected (line 197)

- Line 179: “Taxonomic rank” or “taxonomic level”?

Corrected (“taxonomic level”) (line 201).

- Supplementary material; Table S1 (and lines 215-220 in the main text): This table isintended to provide an exhaustive list of pesticides and pharmaceuticals measured in water and sediments, in which only compounds in bold have been included in the analysis … but it seems that none of the 47 compounds listed in the last column are in bold type. Please, harmonise the table with its legend and the last paragraph of the sub-section “Determination of water and sediment quality parameters, including trace elements and organic pollutants”, which exhaustively provides the micropollutants included in the analyses.

The pollutants studied are now highlighted in bold.

- Line 482: Change “Antipodarum” in “antipodarum”.

  Corrected (line 476).

- Line 519: Change “Proasselus sp.” in “Proasellus sp.”.

  Corrected (line 497).

- Lines 599-722: All the references are not provided in a homogeneous format. Some ofthem need marginal edits:

- Line 654: Change “Forest Ecosystems 2019; 6” in “Forest Ecosystems 2019; 6: 7”.

- Line 660: Change “Science of the Total Environment 2020; 740” in “Science of theTotal Environment 2020; 740: 140029”.

- Line 663: Change “Environ Sci Pollut Res 2020” in “Environ Sci Pollut Res 2020; 28,59256–59267”.

- Line 667: Change “Ecosphere 2019; 10” in Ecosphere 2019; 10: e02810”.

- Line 674: Change “Science of the Total Environment 2021; 773: 12” in “Science of theTotal Environment 2021; 773: 145467”.

- Lines 690: Change “Sun ZH” in “Sun Z”.

- Line 710: Change “Environmental Monitoring and Assessment 2021; 193: 11” in“Environmental Monitoring and Assessment 2021; 193: 694”.

We corrected the references.

Review by Aurélie GOUTTE, 18 Jan 2024 16:24

> The manuscript entitled “Do macroinvertebrate abundance and community structure depend on the quality of ponds located in peri-urban areas?” deals with spatial (12 sites) and temporal (spring and fall in two consecutive years) diversity of macroinvertebrate communities and the effects of physicochemical parameters and contaminants (trace elements, pharmaceuticals, pesticides and polycyclic aromatic hydrocarbons) on abundance and community structure.
>
> This is an interesting and original study, the manuscript is well-written and the scientific approach is clearly presented.
>
> I only have a few comment and questions:
>
> - I am more familiar with techniques used in lotic systems than in ponds for macroinvertebrates inventories. I am a bit surprised to see no reference in the methods (lines 159-169). Do you think that the method you have developed here can be extended at a larger scale and used as a reference? Can you include a critical opinion, recommendations and improvements in the discussion?
>
> This method is widely used by experts for sampling (Hanot, pers. Comm.). Moreover, it is justified by pond physionomy (presence of trees, brambles, stones, topology, etc.) that prevent linear sampling.
>
> - For I2M2, some species characteristics and functional traits are considered, such as the relative abundance of polyvoltine taxa or ovoviviparous taxa, which can be pertinent indicators of environmental perturbations. Do you think that these traits can be helpful in ponds and that they could be considered in your study, in addition to Shannon index, evenness, etc?
>
> As previously mentioned, we have chosen to focus on diversity indices in this article, which already contains a great deal of data.
>
> - Lines 230-231: “We tested the effect of individual ponds and field campaign on these parameters with an analysis of variance followed by a pairwise comparison with Tukey’s HSD test” Since each pond was sampled 4 times, mixed models with repeated measures can be more appropriate.

We replaced the analysis by a linear mixed-effect model with the sampling session as a random effect.

Several environmental parameters were measured and included tested. I wonder if the number of ponds (N=12) was sufficient for testing all these parameters.

For ANOVAs and the linear models, the number of ponds was sufficient. For the RDA, we consider the ponds for each field campaign as “individuals”.

- I have noticed a few grammatical errors, for instance line 106-107 “the water contamination profiles of these ponds different depending on their location”

Corrected.

Review by anonymous reviewer 2, 04 Feb 2024 20:43

Title: Do macroinvertebrate abundance and community structure depend on the quality of ponds located in peri-urban areas?

General comments: The methods are clearly stated. The manuscript " Do macroinvertebrate abundance and community structure depend on the quality of ponds located in peri-urban areas? " is an interesting study. It is well written and has a massive amount of data. I have a few comments.

Introduction: The last part of the introduction is a little too long, giving too much information that prevents a quick and clear overview of the objectives. Perhaps the authors could delete the sentences from line 102 to line 109 (from "... Depending on their chemical... to... mainly due to pesticides (Nélieu 2020)"). Some of these sentences could be moved in material and methods if needed. Moreover, line 109 add "et al" in the citation (Nélieu 2020).
We moved in the Material and methods, section 2.4, the sentences "The choice of contaminants assessed here was based on the local activities: cereals, maize, rapeseed, sunflower, orchard and vegetable crops for pesticides; nearby roads for TE and PAH; and the presence of humans, farming, and domestic pets for pharmaceuticals."
We corrected for Nélieu et al. 2020.

Material and Methods: The methods are clearly stated.
Line 167: I'm not sure to understand the number of invertebrate samples. Could you check your calculation? (44 invertebrate samples: 12 ponds x 2 points x 2 years x 2 seasons, minus6 cases).
Corrected.
Line 168: What's the reference of « Section 2.3 »? Please check this part as there is no numbered section through the text.
Corrected.

Results: The results section is a very dense. There's a lot of information and it's sometimes hard to follow. Maybe the authors could find a way to shorten it a little?
We moved the RDA results with abundances in the Supplementary material.
Table 1 is not very relevant as a single table. For greater clarity and to make it easier to read the results, the values presented in table 1 could be placed on each corresponding panel of fig 1 as an insert, or write the corresponding values in the legend of Figure 1.
We moved the table in Supplementary material.
Figure 3: please check the figure. The letters to identify each panel are missing. It seems that solid and hatched curves are inverted. Figure legend should be moved below the figure
We deleted this figure.

Discussion: The discussion is long as well, but it helps the reader to better understand the results. Maybe the manuscript will gain if it could be summarized a little without losing any information.
Line 538: Citation Nélieu et al. 2020 has to be corrected (remove "and al")
Line 539: Check the sentence it's seems that a word is missing?

Corrected.

Lange RT. The piosphere: sheep track and dung patterns. Journal of Range Management 1969; 22: 396-400.