Dear editor and reviewers,

Please find enclosed our revised manuscript entitled; “*XXXXXXXXX”*, which we would like you to consider for publication in *Oikos*. We appreciate the opportunity to resubmit the manuscript, and gratefully acknowledge the degree to which these reviews have improved it.

We carefully addressed and integrated the comments. We believe these changes help clarify our main messages and the significance of our contribution.

The following points summarize the major changes we have made to the manuscript:

- Following comments from R2, we clarified the methods, especially on how we measured lake connectivity and the how we sampled the different lakes.

- Following comments from R3, we addressed the ambiguities in our abstract and standardized our use of words throughout the manuscript.

- Following comments from R3, we clarified and expanded the justification of our second hypothesis that relies on active site selection for oviposition by semi-aquatic predators

- Following comments from R3, we re-did our analyses to address issues of non-independence in our data. The results remain qualitatively unchanged.

We provide below a point-by-point response to each comment.

We look forward to hearing back from you.

Sincerely,

Juliane Vigneault on behalf of all authors

EDITORIAL FORMATTING ISSUES

Authors contributions and affiliations have been corrected. Bibliography has been revised.

Figures format and quality have been adjusted. We also contributed a “Speculations” section at the end of the manuscript.

SUBJECT EDITOR

I got the recommendations and comments from two expert reviewers in the field. Both reviewers agree that the manuscript is technically sound and the data support the conclusions. However, "different families and utilise snail first intermediate hosts also belonging to different families" was suggested by both reviewers, and some limitations of this study. I share their comments. Therefore, I can invite you to submit a revised version of the manuscript that addresses the points raised by the reviewers.

As suggested by the Subject Editor and Reviewer 1, we addressed the issue of phylogenetic pooling by including this point as a caveat in the newly added “Speculations” section. We carefully addressed and integrated the other comments. The proposed changes greatly helped to clarify our main messages and the significance of our contribution.

REVIEWER 1

[1] I understand the rationale for lumping different parasite species together as black spots, however they belong to different families and utilise snail first intermediate hosts also belonging to different families. Parasites and snails may have different tolerances of abiotic conditions, etc. Pooling species (actually, pooling families) is bound to generate spurious patterns. Perhaps this needs to be discussed as a caveat.

We agree that pooling different parasite families is an unusual approach in the study of host-parasite ecology that comes with many challenges. We developed this issue in the new “Speculations” section acknowledging the overlook of phylogenetics and the need to explore it in future work.

[2] Similarly, focusing on fish community-level prevalence can mask interesting patterns at the fish species level. Would it be possible to repeat the same analyses by focusing only on a single fish species (the one found in highest numbers at most localities, probably pumpkinseeds)?

This is a good point, and indeed one that we had considered ourselves. After careful reflection, we decided to focus on fish community-level for several reasons: 1) black spot trematodes are difficult to identify to species level, and most species have a low host specificity, 2) community infection prevalence is easier to measure by wildlife managers/technicians since there is no need to target specific fishes. It also permits the inclusion of lakes where target species are not present or not abundant, and therefore greater coverage of the water bodies within a drainage basin.

Nonetheless, we agree that interesting patterns can be masked by grouping fishes at the level of the community. For this reason, we executed the three same analyses focusing on species-level metrics, using data on the pumpkinseed sunfish. This is the only species abundant and prevalent (in terms of localities) enough to perform all three analyses. We summarize the main findings from these analyses below, which largely reflect the patterns obtained at the community-scale. We decided not to incorporate these species-level results in the current manuscript for brevity’s sake but also because the results are qualitatively the same. However, we would be happy to include these results in the supplementary results if deemed necessary.

For the ***regional-scale analyses*** we can see that despite some quantitative discrepancies explained by population-level differences in infection prevalence, both plots show the same trends with decreasing prevalence value along the sampling gradient (see Figure R1 at the end of this document).

For the ***lake-scale analyses***, patterns of frequency distributions are different between population and community scales (Figure R2). However, we must consider that pumpkinseeds are only present in 11 lakes out of 15 lakes sampled meaning the community-level includes more data than the species-level plot. All distributions in Figure R2 are left-skewed because pumpkinseeds have a higher infection susceptibly, and potentially, a higher encounter rate with black spots than other species.

For the ***site-scale analyses***, some of the models previously significant at the community-level were not at the species-level (Figure R3). This is the case for pH, conductivity, macrophyte cover, perimeter, non-host abundance and Simpson’s diversity index models – suggesting that observed patterns at the community-level for these predictors do not rely on species-specific mechanisms. All models followed the same trends at both the community and population levels except for the turbidity model (Figure R3.Aa–R3.Ba). Only one model that was found significant at the species-level was not at the community level (total phosphorus; Figure R3.Ae)

[3] It is unfortunate that prevalence was the only infection metric considered here. The impact of trematode metacercariae on fish is intensity-dependent, therefore measures such as mean intensity or abundance would better reflect spatial variation in the parasites’ impact on host population dynamics. Even if counting the spots on each fish would have been impossible considering the time required, a rough visual estimate (relative scale from 0 to 5, i.e. few to very many black spots) could have been obtained quickly in the field and provided very useful data.

Indeed, other infection metrics such as abundance and intensity would better reflect spatial variation and risk assessment of host-population as well as individual-impact. However, as mentioned, these metrics would be impossible to accurately measure for all the different methods compared in our study. We did estimate the intensity of infection on a semi-quantitative scale from 1 to 5 for all fishes caught using the fishing methods (i.e., minnow trap and seine net). After reflection, we decided not to use this infection intensity data in the current manuscript because we could not compare this data among all sampling methods, which was one of the main goals of our study.

[4] The authors show that low sampling effort overestimates prevalence, with the estimates eventually stabilizing around the ‘true’ value as sampling effort increases (Figure 3). This is a really nice finding. Here, sampling effort is measured as the number of distinct host samples. However, sampling effort can also be measured as the number of individual hosts examined. In the Discussion, the authors allude to this and refer to Jovani & Tella (2006). They may also want to consider a more robust study, that of Gregory & Woolhouse (1993, Acta Tropica 54: 131-139), which suggests the opposite, i.e. that low sampling effort may lead to an underestimate of true prevalence.

A few studies have indeed examined the effect of sample size (i.e., the number of hosts) on parasitological parameters such as parasite abundance, infection prevalence or parasite aggregation. From what we understand, conclusions of Gregory and Woolhouse (1993) stipulate that estimates of parasite prevalence were close to the true values and that this did not vary with sample size. Their findings are nevertheless contrasting with our own conclusions. We think this is something worth mentioning

Thank you for the suggestion! We added the sentence (L462-463): “Gregory and Woolhouse (1993) found contrasting results where simple size has no effect on parasite prevalence values.”.

[5] Is it possible that the bimodal distribution of prevalence values across lakes (Figure 4) for the transect and combined methods is due to the fact that a snorkeler is more likely to detect infection when fish are heavily infected? So if you have two lakes, one with high intensity of infection and one with low intensity, but both with the same prevalence, visual detection by snorkelers might lead to high and low prevalence estimates, respectively. It’s easier to see 100 black spots on a fish than just two or three spots. Could the bimodal distribution reflect high variance in average intensity of infection among lakes, not in prevalence? The other two methods, where fish are held in the researchers’ hands, are much less likely to produce prevalence estimates affected by low intensity values.

This is an interesting thought that we had not considered. Indeed, this phenomenon might be possible since snorkeling observations have a higher infection assessment uncertainty. If we consider fishing prevalence as the “true” prevalence (i.e., prevalence measured with data from both minnow traps and sein net) and low-prevalence lakes as those reporting ≤ 25% of infection, we can compare the “true” prevalence of these lakes to the prevalence obtained through snorkeling observations. In our data, there are 4 lakes fitting these criteria that were sampled with all three methods (excluding Lake Triton because 0% prevalence of infection was estimated in all three methods). A table (R1) comparing the prevalence measured with either fishing method (“true prevalence”) or snorkeling transects shows that snorkeling observations underestimated infection prevalence in ¾ lakes with difference in prevalence varying between 8 and 15%. In Lake Echo, the snorkeling method led to a prevalence estimate 3% lower than the fishing methods. Therefore, we agree with you but our data doesn’t allow validation of the hypothesis.

For the second part of this comment relating to bimodal distribution reflecting variance in average intensity instead of prevalence, we compared frequency distributions of lake’s prevalence suggesting bimodality (transect and method combination) with frequency distributions of lake’s mean intensity class (mean intensity class estimated with data from both fishing methods – not used in this manuscript) (Figure R4). The data do not show a clear relationship between a lake’s parasite prevalence and intensity frequency distributions.

[6] Regarding spatial patterns in hotspots and coldspots of infection by trematodes in fish, the following may be relevant: Blasco-Costa et al (2015) Ecography 38: 301-310  
  
[7] Regarding the environmental (biotic and abiotic) drivers of prevalence, the following may be relevant: Martins & Poulin (2024) Proceedings B 291: 20241673  
  
[8] Lines 188-190: the same sentence is repeated twice.

Thank you for catching this typo. The repeated sentence has been deleted (L188-189).

[9] Lines 298-304: two of the 15 fish species were excluded from prevalence estimations because the literature suggested they were not suitable hosts. Why not repeat the analysis after excluding the other seven species (by my calculation) that were not found in your survey to be infected at all? Your own data provide some evidence that they may not be suitable hosts as well…

The problem with this suggestion is that cyprinids were not identified to the species level during the snorkeling transects. Nonetheless, we tested the effect of excluding uninfected species for the minnow trap and seine net methods while excluding all cyprinids for the transect method (cyprinids show none or very low infection levels). We saw minimal differences in the prevalence accumulation curve (conclusions stay the same). Frequency distributions were different when excluding these fishes from the prevalence estimation, probably because we are excluding a lot of individuals from the community. Only null, turbidity, dissolved oxygen, area:perimeter and fish abundance models were significant when excluding muskellunge, brown bullhead and all cyprinids.

[10] Lines 468-469: missing word? “in terms of area and time to ESTIMATE infection prevalence”?

This is correct. We added the missing word “estimate” in line 472.

REVIEWER 2

[11] The authors combine two contributions, the analysis of local determinants of parasites prevalence and the analysis of estimation biases. In a respectful perspective, my opinion is that this duality is not well articulated in the manuscript, with a result that dilutes these contributions rather than complementing and potentiating them. I wonder if the large development of the sampling bias issue should not be toned down here to support for the reported patterns and further develop a focus on the role of biotic and abiotic environments on parasites prevalence. Note that the effect of aggregation on sampling design and statistical analyses, as well as the inherent biases, were extensively considered elsewhere. Of course, the focus of the manuscript is the authors' responsibility, and here I just transmit a respectful reflection.

Thank you for these thoughts. We agree that the manuscript articulation and flow are unconventional relating, as mentioned, two different contributions (environmental drivers and sampling biases). We (the authors) discussed the format of the manuscript carefully for several months. We considered the current format or writing two different manuscripts (one on sampling biases and one on environmental drivers of prevalence). After discussion, we leaned towards the current format because it allows a better overview of the study system and variability involved in field studies. We believe that the sampling bias issues addressed here help fill a significant literature gap on the effect of sampling design on the estimation of infection prevalence. While the effects of spatial aggregation and sample size have been addressed before, especially in community ecology, the effect of the number of samplings has never been explored. In addition, an examination of sampling methods as related to infection prevalence estimates has not been extensively reported (but see Wilson et al., 1993; *Journal of Comparative Psychology*), and never in the context of community prevalence. We feel that these elements relate to the reality of field studies and surveys, with great contributions in the understanding of sampling design biases and in the planification of field work.

[12] The authors repeatedly highlight some general statements about the complexity of the study system, the dependence of scale on pattern and processes, and the role of different sources of heterogeneity. These properties are common to most, if not all, ecological systems. In this sense, they do not provide a particularly important context for the present contribution. In my opinion, a better framework could be considered.

These are indeed common to most system, but we believe that context-dependencies, while usually theoretically acknowledged by the scientific community, are often pratically overlooked or ignored, especially in parasitological ecology where these “common knowledge” are less developed than in other fields of ecology.

[13] In the abstract, the authors alternate between real patterns and their determinants and the sampling issue that does not adequately articulate the two focus. I think this section could be improved with a better order and clear ideas. It is now difficult to easily extract the main contributions of the study. In my opinion, the presentation of both the framework and the specific contribution could be improved.

The idea in the abstract was to follow the order in which analyses are presented in the manuscript (landscape- lake- site-scale). We moved the sentence: “Our results showed that infection prevalence is spatially heterogeneous across the landscape with evidence of infection hotspots and coldspots” to improve clarity (L13-15).

[14] The discussion devotes a great deal of attention to sampling issues, not providing a robust consideration of the theoretical framework and the contributions attained in this research. In addition, it is large and, in my opinion, fails to identify novel contributions over general issues such as individual aggregation and sampling design.

Our study was observational in nature, which means we could not directly falsify any theoretical hypothesis. We therefore tried to be cautions with our result interpretations.

[15] L114-126. I think this is not a good paragraph to close the introduction. It is not clear, it focuses on lateral issues about cold-hot areas and estimation biases, with general comments not fully rooted in statistical grounds. There is a broad and diverse literature on the effect of aggregation in sampling design in ecology and other areas. It is not evident that the study contributes something new in this framework.

We agree this an unusual paragraph to end the introduction. We decided not to present our predictions after all and removed this paragraph. We incorporated a new sentence at the end of the previous paragraph on general contributions of the present study (L119-121).

[16] L164 it is not evident that this represents a random design.

Our field survey was not based on a random design because of land and data accessibility limitations. However, the selected sites display the variability in morphometry, habitat types, and anthropogenic use picturing the regional variability of lakes.

[17] L281 a cubic spline with one variable at a time is producing an overfitting and suggesting more relationships than those that are actually supported by the study.

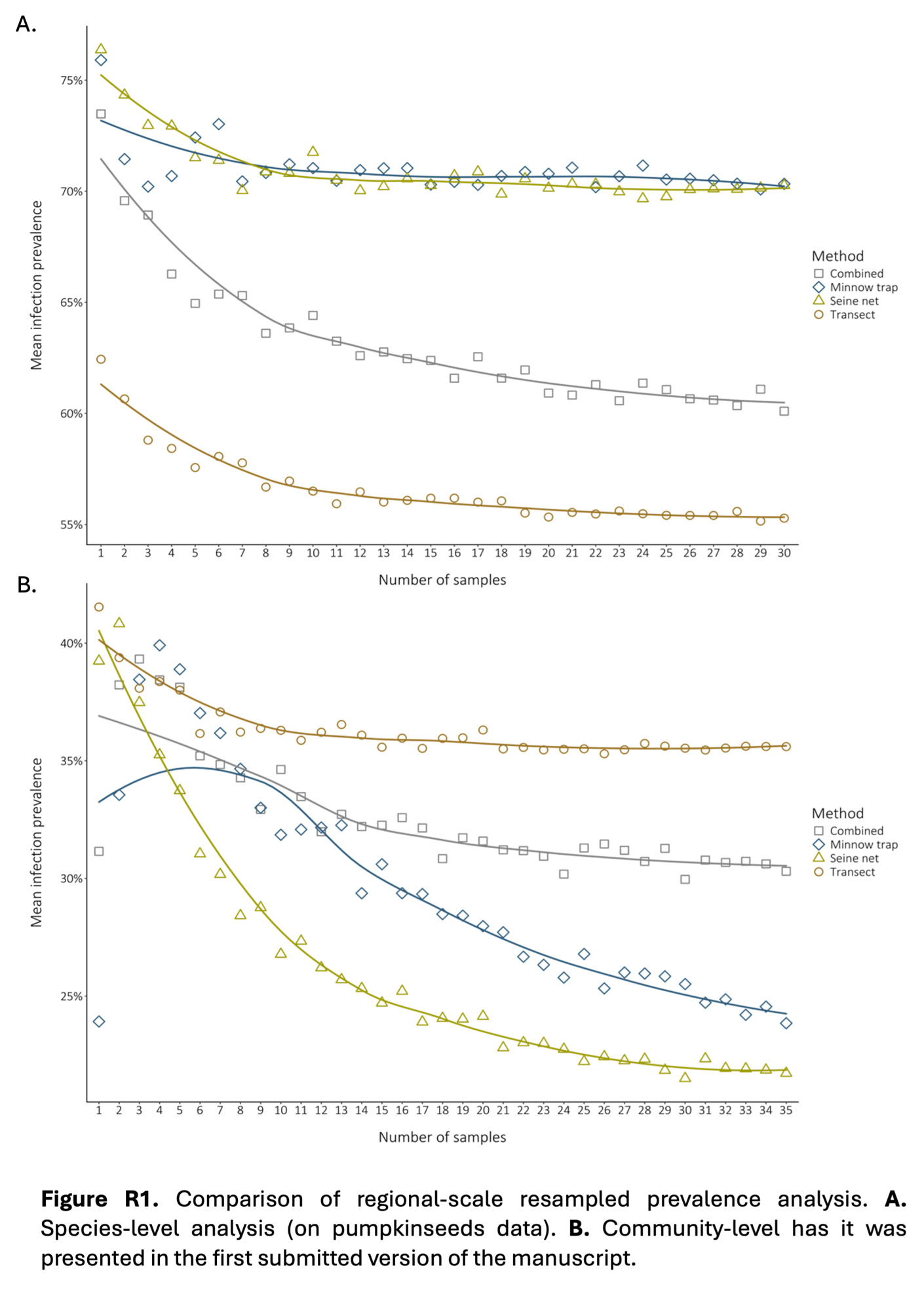
Using highly flexible models such as GAM(M)s, especially with small sample sizes, comes with this overfitting risk. Since we needed to include a random effect to account for nested data, we could not have many predictors to our models because it uses too many degrees of freedom. We chose a cubic spline instead of a cubic regression spline or a thin plate spline because i) we did not expect abrupt changes in the relationships between prevalence and environmental drivers, ii) cubic splines offer the smoothest option and offer no computational problems for univariate smoothing and, iii) their computational cost is lower than thin plate spine (that we want to minimize since we have a small sample size)(Wood, 2006; Wood & Augustin, 2002). Moreover, our aim here is not to use these models to make predictions but simply inspect data to better understand how prevalence estimates vary with local environment. We did not find any evidence that using a cubic spline with one variable actually increases the overfitting compared to other smoothers that are appropriate for our data. Nevertheless, we decided to reduce the number of basis function at 5 (k=5) to reduce the computational burden and chance of overfitting (Wood & Augustin, 2002).

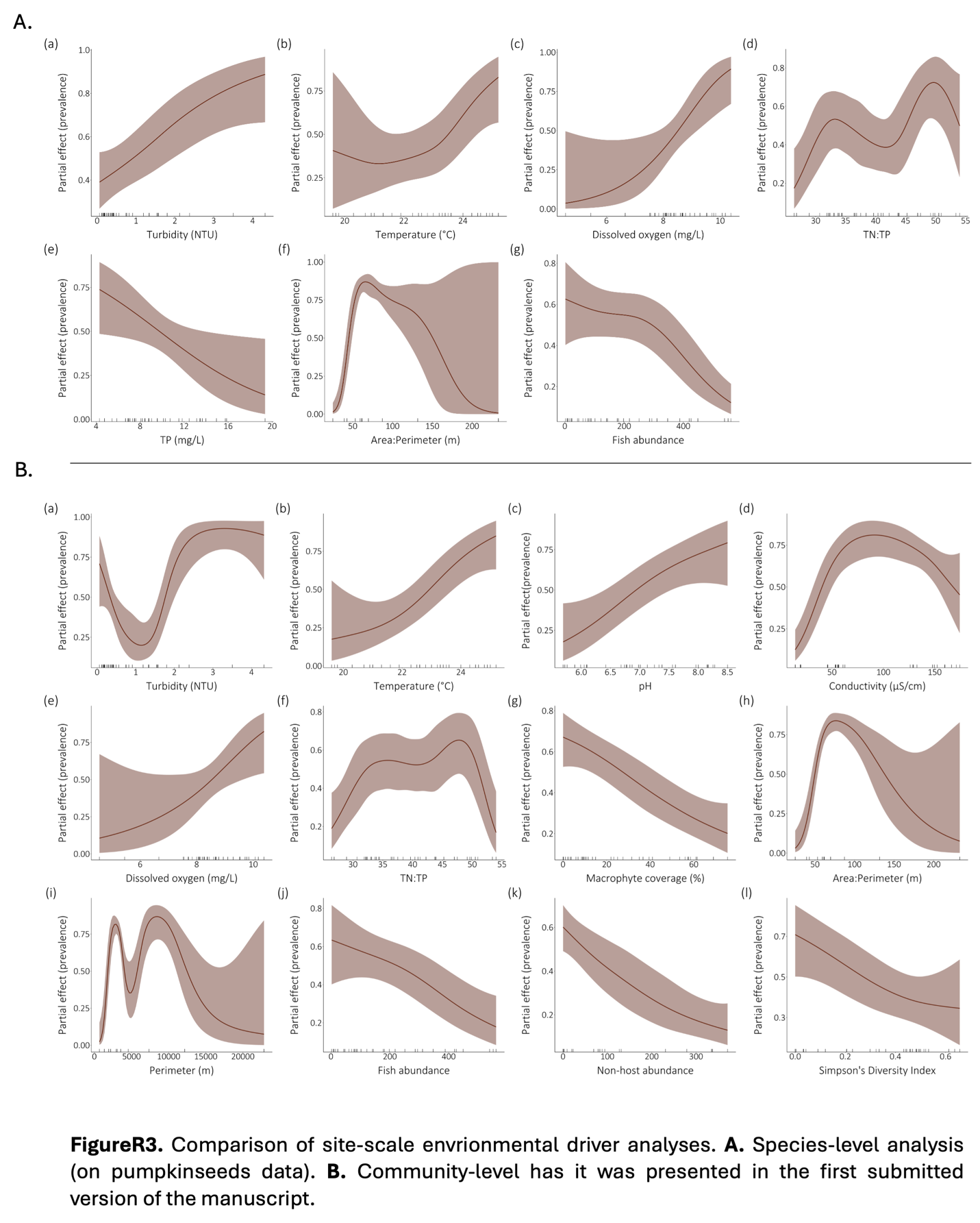
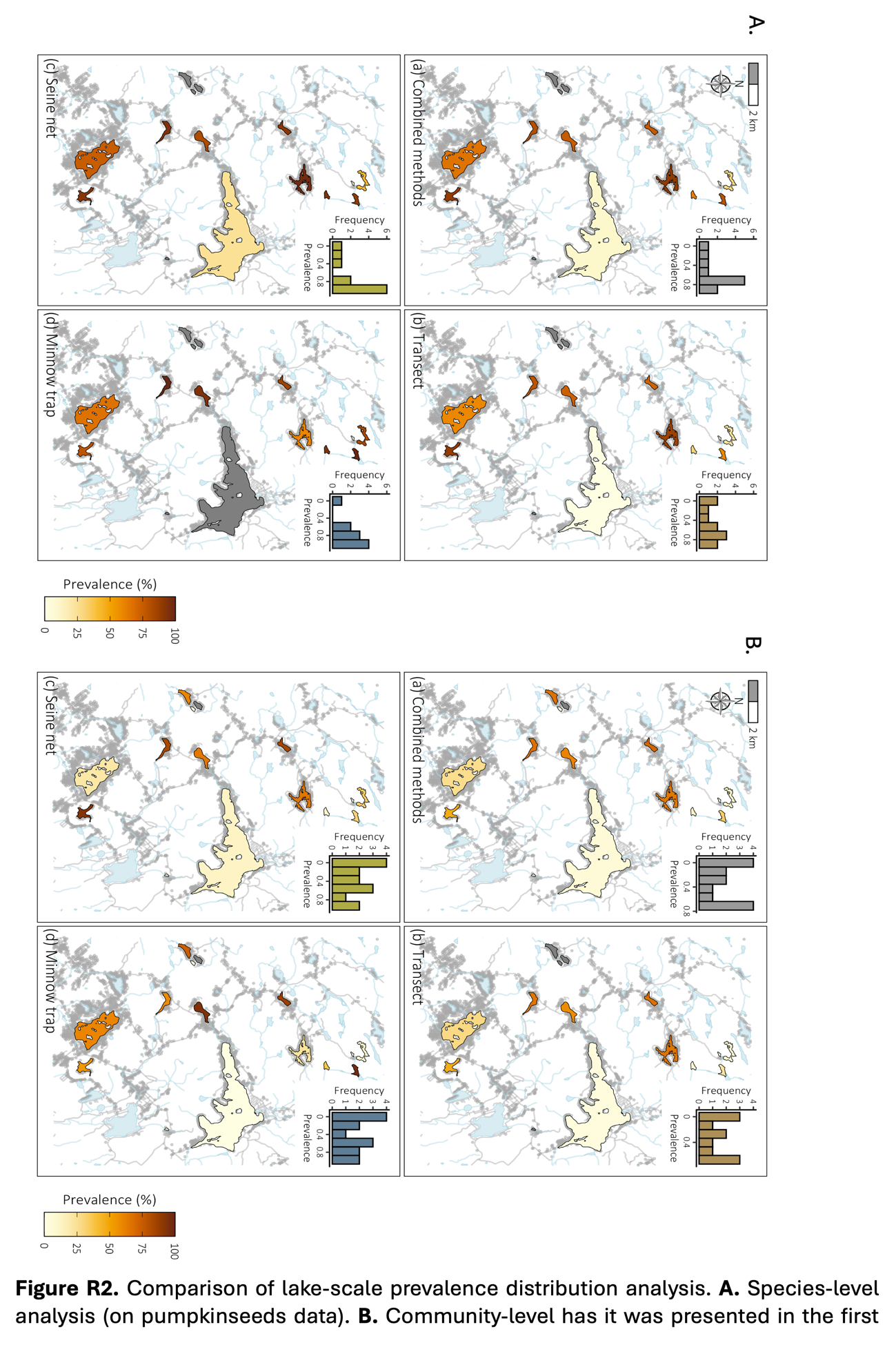
[18] Figure 2, this is a nice figure but not particularly informative for the space that will finally be in the published ms.

The figure is indeed quite large and not essential for readers with an understanding of parasite ecology. However, given the broad readership of Oikos across sub-disciplines in ecology, we believe that this figure is helpful to illustrate the complexity of the trematode life cycle and the environmental characteristics that might influence the observed patterns. That said, we are willing to remove it if the editor deems it necessary.

[19] Figure 3 Presenting the confidence interval for a Loess smooth makes no sense for the present analysis. The confidence interval for 999 resamples must be reported. Indeed, a large interval is expected at lower sampling levels and its reduction with the increase in sampling size. The trend in these intervals (from the true sampling error) is the subject of analysis for the sampling design.

Thank you for this comment. We have removed the confidence intervals for the loess smooths. We have now calculated confidence intervals for each N resampled 999 times. We report these in Appendix S2: Table S10, instead of the Figure 2 (new figure number) because adding the intervals to the figure made it more difficult to read and interpret.





|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Lakes | Achigan | Corriveau | Croche | Echo |
| Fishing prevalence | 12% | 20% | 18% | 25% |
| Snorkeling prevalence | 3% | 5% | 10% | 28% |
| Difference | 9% | 15% | 8% | -3% |

**Table R1.** Comparison of fishing “true” prevalence with snorkeling prevalence in low-prevalence lakes (≤ 25%).

