**3 | RESULTS**

Our main objective was to test the effect of spatial scale on infection prevalence estimates across three different sampling methods (minnow trap, seine and transect). First, we built infection prevalence accumulation curves, then we compared the frequency distribution of the lake’s prevalence for each sampling method, and finally we identified predictors of prevalence patterns.

***3.1. Regional scale***

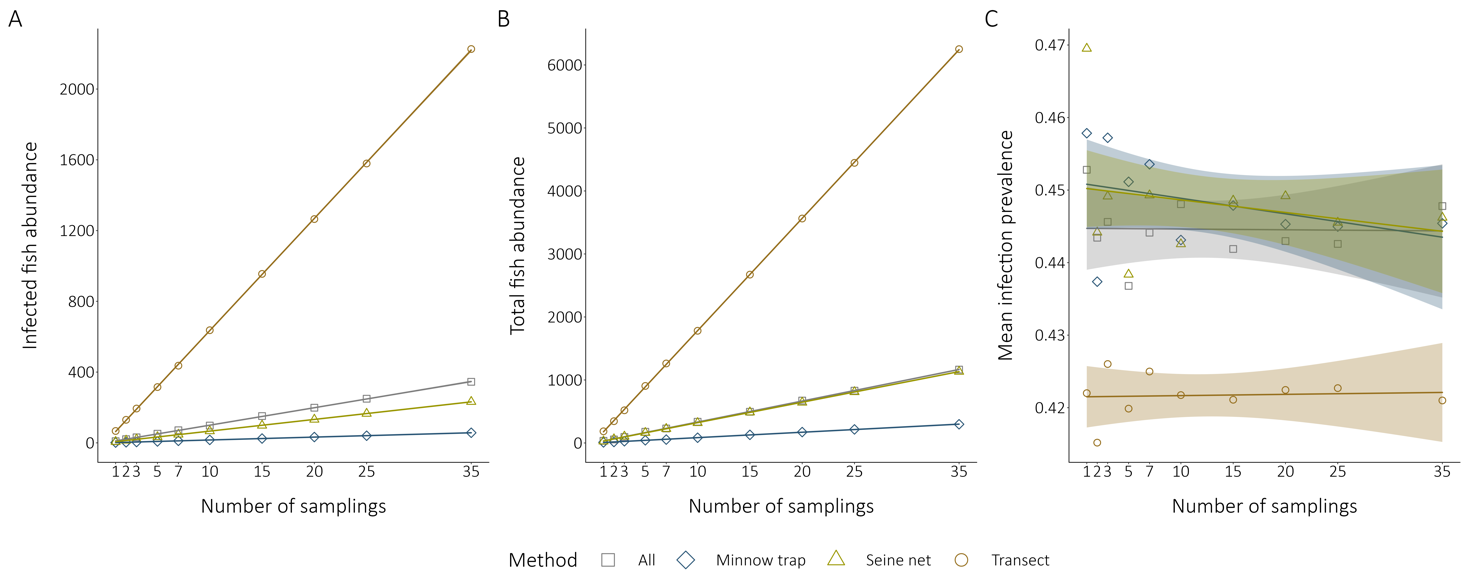
The number of infected individuals (Fig. 3A) and the total number of individuals (Fig.3B) increased linearly through the sampling gradient. The transect method sampled by far the most individuals while the minnow trap method caught the least individuals for the infected and total fish abundance. The rates (slope) of accumulation of individual and infected fishes for the same sampling method are constant (Fig.3A vs. 3B), which is then reflected in the prevalence accumulation curves (Fig.3C). The prevalence accumulation rate did not change between the number of samplings (slope did not differ from 0, T=, p-value=, see Table S16). The accumulation curves of the mean prevalence (Fig.3C) showed more variation than accumulation of infected individuals or total individuals. The minnow trap method leads to a lower regional prevalence estimation (XX%) while the transects displayed the highest estimation (XX%). Some data points remain outside the standard error interval suggesting that although there is a general tendency, extreme values are frequent, especially at lower sampling levels. However, relatively few random samples are necessary to estimate regional prevalence (at X samples, confidence intervals have already declined by 50%) The summary results of the linear models are shown in Table S16.

A graph of a number of dots

Description automatically generated with medium confidence

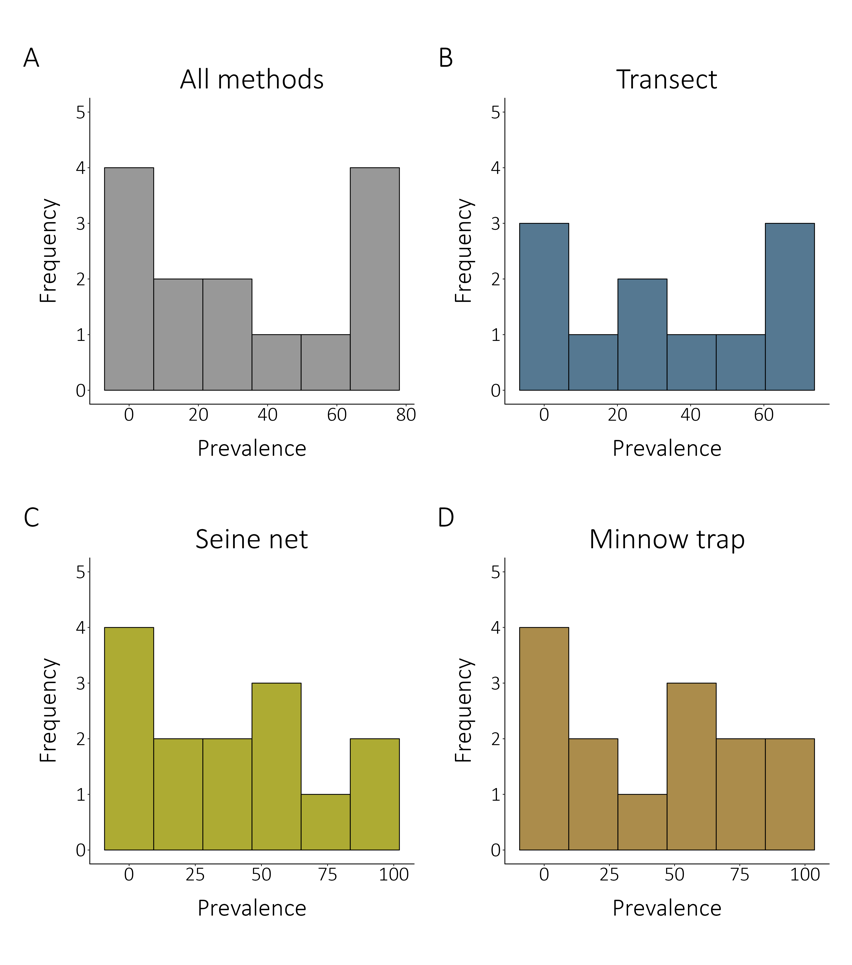
**FIGURE 3**

Accumulation curves of regional-scale sampling through an increasing random sampling intensity. The different sampling methods are compared for all three simulations. The data points indicate the mean prevalence for a given sampling intensity. Shadings indicate the 95% interval confidence from the 999 replicates of each sampling intensity. The pool data for all methods, minnow traps, seine nets and transects respectively contained 339, 225, 75 and 39 samplings. **(A)** The accumulation of the infected fish abundance. **(B)** The accumulation of the total fish abundance. **(C)** The accumulation of the regional mean community prevalence.

******

***3.2. Local scale***

We compared the frequency distribution of the lake’s prevalence for each sampling method to see how the regional prevalence is distributed at a local (among-lake) scale (Figure 4).

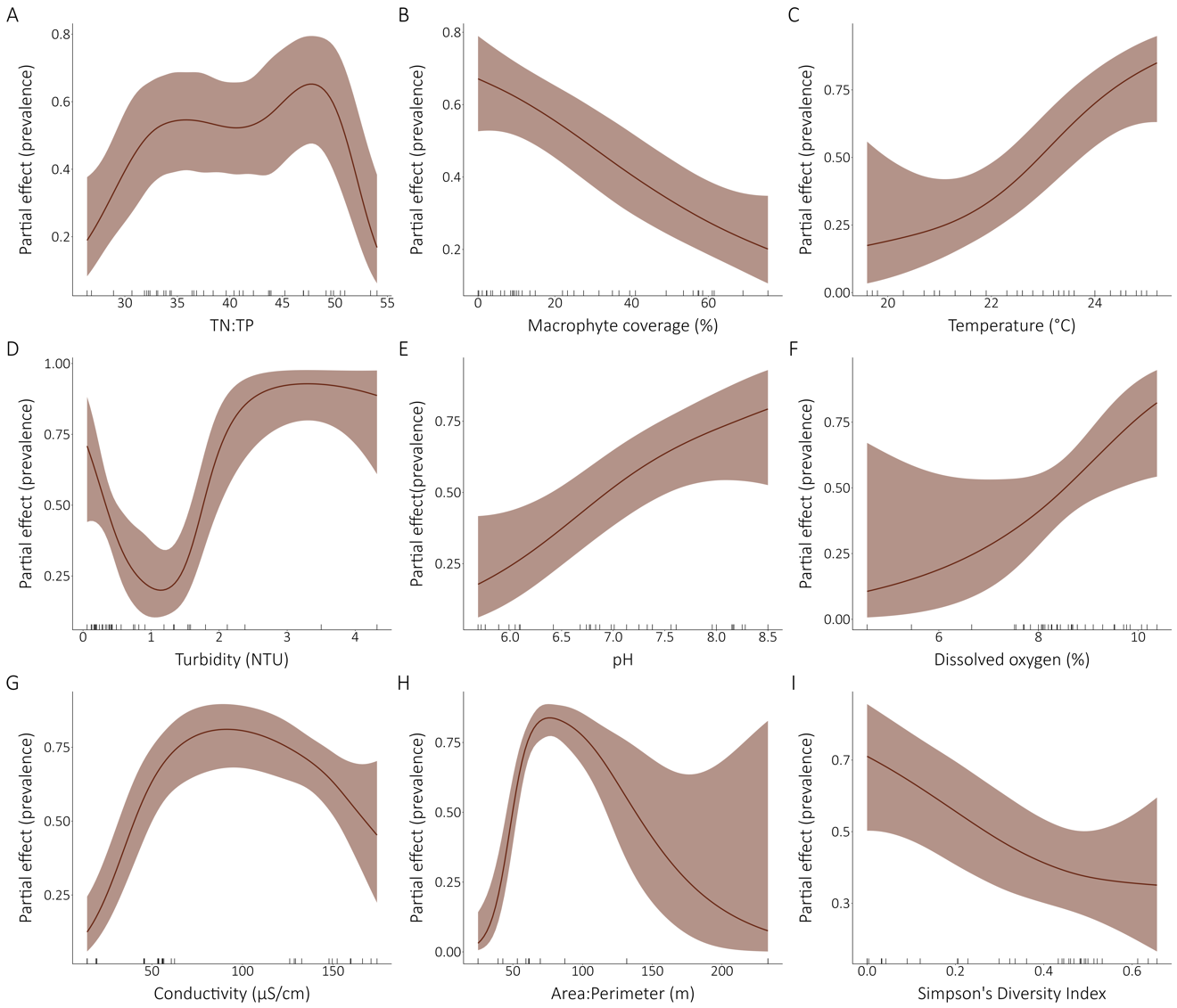


**FIGURE 4**

Comparison of the frequency distribution of the local community prevalence depending on the sampling method. All the frequency distributions considered 14 lakes, except for the transect method that survey 11 lakes. The figure shows frequency distribution among-lakes for **(A)** all the sampling methods combined ( weighted mean regional prevalence is XX%). **(B)** transect method only (weighted mean regional prevalence is XX%). **(C)** the seine nets only (weighted mean regional prevalence is XX%). **(D)** the minnow traps (weiged mean prevalence is X%).

The local community prevalence estimations are not constant across the regional scale. The local prevalence frequency distributions for the combined methods and for the transect method show a bimodal distribution, with the landscape composed of many low-prevalence and high-prevalence lakes (leading to regional prevalence estimates shown on Fig3C). The distribution patterns are more unclear for the two fishing methods (seine net and minnow trap), although they seem left-handed displaying more low-prevalence lakes. Interestingly, despite similar frequency distribution of infected lakes, seine and minnow traps lead to very different prevalence estimates at regional scale (see Fig.3C). Heavily infected and lightly infected lakes do not appear to be clustered in space at the regional scale (Figure 2). Moreover, close and connected lakes do not appear to follow a spatial infection gradient, suggesting that geographic attributes are not important drivers of the local infection prevalence (Fig.2 and Fig.5).

***3.2. Fine scale***



**FIGURE 5**

Relations between the fine-scale community infection prevalence and selected environmental drivers. All the models presented are univariate binomial generalized additive mixed models with the lake as a random effect on the intercept. The partial effects of the environmental variables on the prevalence are presented for the significant models. The ticks on the x-axes indicate a data point. TN:TP ratio, macrophyte coverage, temperature, turbidity, pH, dissolved oxygen, conductivity and Simpson’s Diversity Index are fine-scale measurements while area:perimeter ratio is a local attribute. The perimeter model is not presented because it was highly non-linear and not interpretable although significative (Figure S1).

The relations between the potential predictors and the fine-scale prevalence (transect-level prevalence) were assessed with generalized additive mixed effects models. The partial effects of environmental variable on prevalence are shown for all significant models in figure 5. The models for TN, TP, TOC, lake area, maximum depth, mean depth, water residence time, drainage area, elevation, distance to the nearest lake, centrarchid abundance and species richness were not significant (see Table S17 for models’ summary values). The turbidity model has the best fit (D2 = 88,71%) of all models (Table S17). The relation between TN:TP ratio and prevalence is highly non-linear showing some kind of unimodal relation. For the macrophyte coverage, we found a decreasing relation between the amount of macrophyte cover and the prevalence infection, meaning that low macrophyte cover correlates with high prevalence of fine-scale communities. The temperature has an increasingly proportional relation with the infection prevalence. The pattern shown for the turbidity model is mostly non-linear. The relation evidence a plateau for high-prevalence values, indicating that a prevalence saturation is reached for turbidity values above 2 NTU. That said, this model must be carefully interpreted as we sampled only a few high-turbidity sites. The pH increased proportionally with prevalence estimate. The same pattern is observed for the dissolved oxygen. However, we must take under consideration that the variation interval is very large for low-concentration oxygen values because of only three values below 7%. The conductivity and prevalence have a non-linear relation following a unimodal relation peaking around 80 (μS/cm). The relation must be carefully interpreted because of some gaps in the conductivity values. A parabolic curve is also observed in the Area:Perimeter model although, high ratio values are more uncommon increasing the variation interval. The relation between the prevalence and the diversity index is the only significant model related to the fish community structure per se and shows a decreasing tendency. Fine-scale communities are slightly to moderately diverse (Simpson’s diversity index between 0 and 0.64) indicating the dominance of some species.