**3 | RESULTS**

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

***3.1. Landscape-scale***

We used a resampling approach to compare how different sampling methods change the mean prevalence estimate through an increasing sampling effort across the landscape (Figure 3).

The estimate of landscape prevalence differed between the sampling methods. After 35 samples, the transect method led to the highest mean prevalence (36 %) followed by the methods combination (31 %), the minnow trap method (24 %) and the seine method (21 %). Although, the minnow trap curve did not stabilize after the 35 random samples explaining that the prevalence value obtained is higher than the actual estimate (landscape infection prevalence measured by the minnow trap is 19%). Some data points remain outside the confidence interval suggesting whereas there is a general tendency, extreme values are frequent, especially at lower sampling levels. Each method overestimated the mean prevalence of low N values (number of samples). However, relatively few random samples are necessary to estimate a landscape prevalence for the transect method (approximately 10 samples) while accurate prevalence estimate only occurs after 30 samples. Even if the method combination curve (in gray) stabilizes around 20 samples, it displays some of variation around the smooth because of the variability among the different method used.

***3.2. Lake-scale***

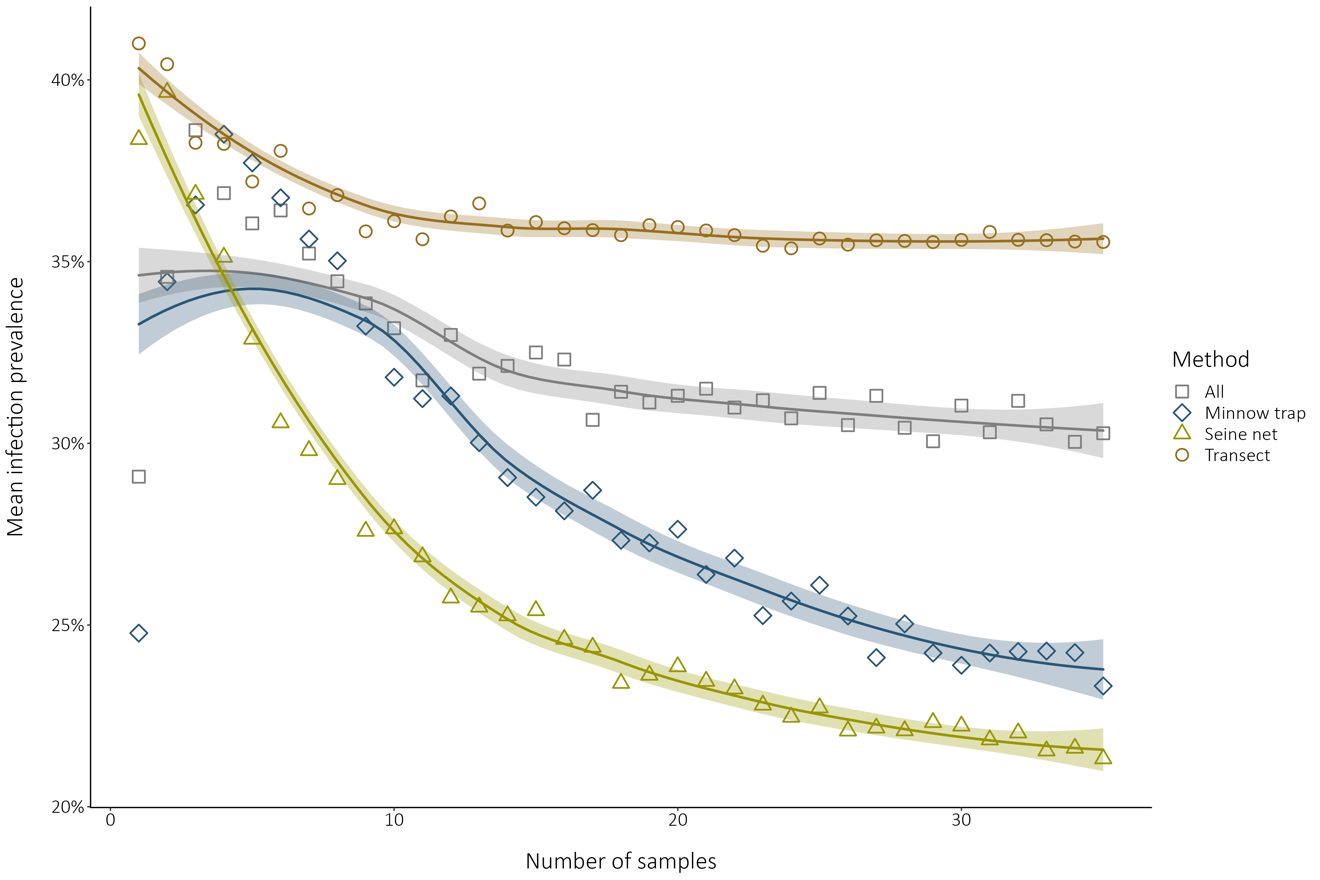
We compared the frequency distribution of the lake’s prevalence for each sampling method to see how the landscape prevalence is distributed among lakes (Figure 4).

The lake community prevalence estimates are not constant across the landscape. The lake 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. However, neither method led to prevalence estimate over 80%. 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. Accordingly, seine and minnow traps lead to very similar prevalence estimates at the landscape-scale (20.4% and 19.2% respectively). 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 (Figure 2).

***3.2. Site-scale***

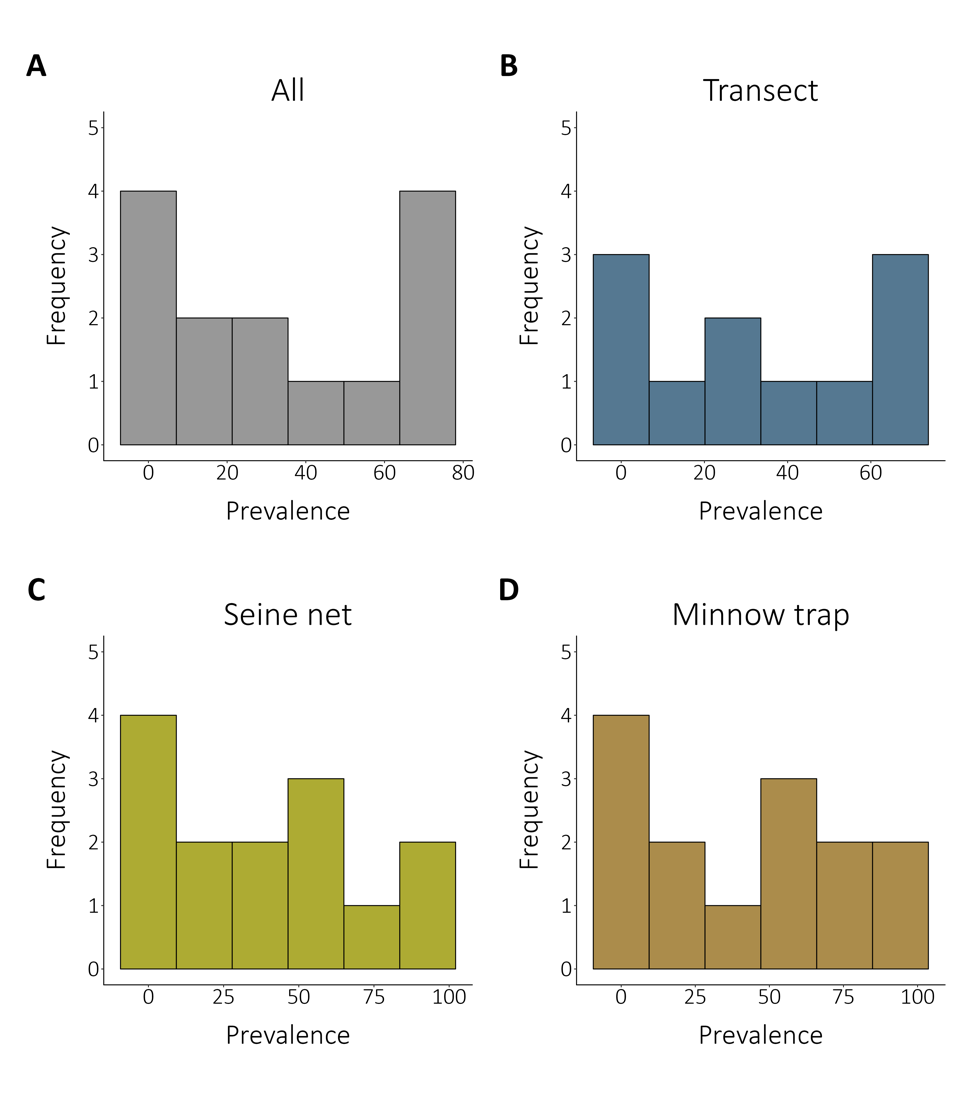
The relations between the potential predictors and the site-scale prevalence (transect-level prevalence) were assessed with generalized additive mixed effects models.

The partial effects of environmental variable on prevalence are shown for all significant models in figure 5. The turbidity model has the best fit (D2 = 88,71%) of all models (see 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 site-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 mg/L. 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. Site-scale communities are slightly to moderately diverse (Simpson’s diversity index between 0 and 0.64) indicating the dominance of some species. 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.

******

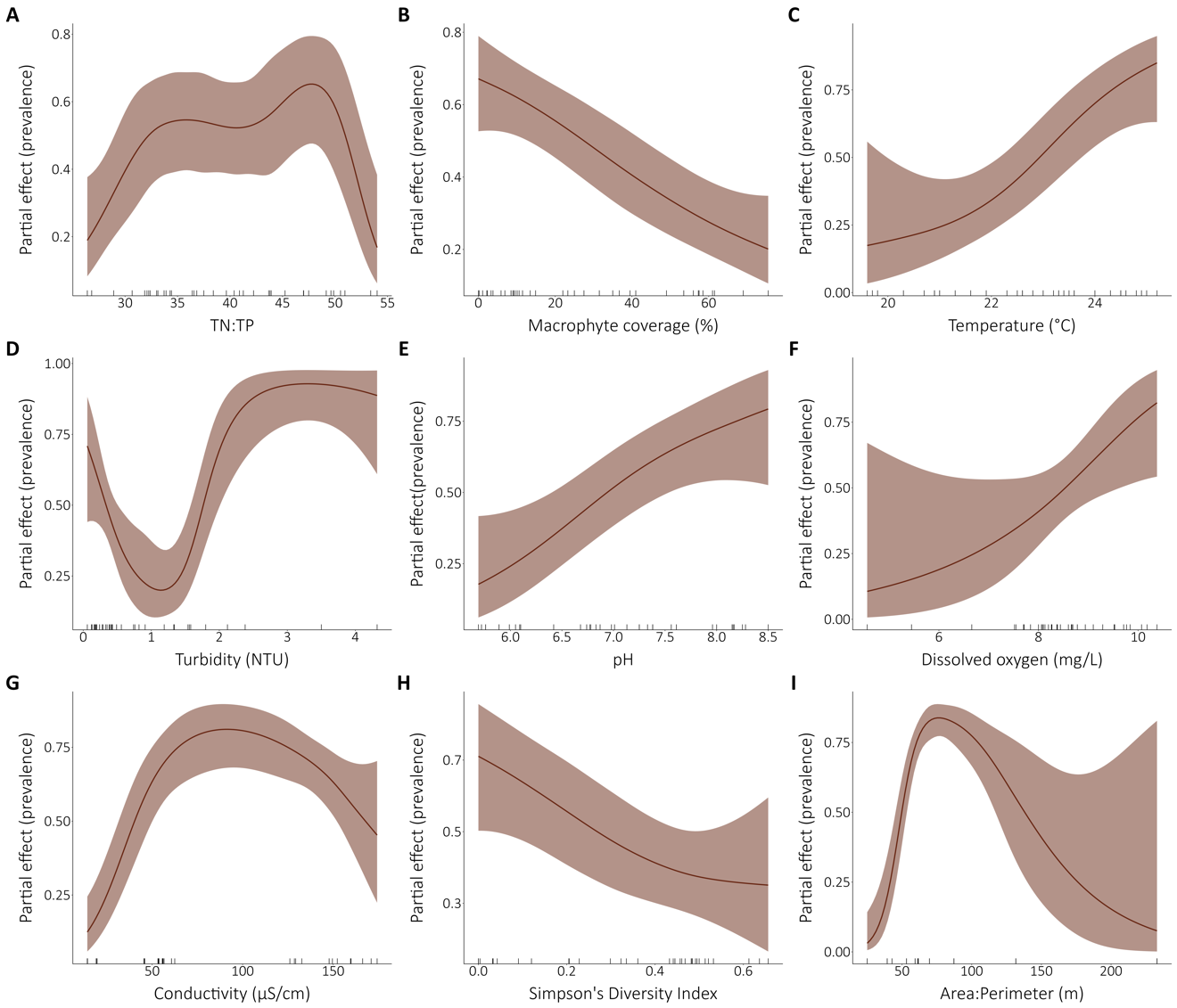
**FIGURE 3**

Simulated method comparison of landscape prevalence estimates sampling through an increasing random sampling effort. The data points indicate the mean prevalence for a given number of samples (N) randomly drafted. The shadings indicate the 95% confidence interval on the loess smooth. The pool data used for the simulation with the minnow traps, the seine nets, the transects and all methods combined respectively contained 225, 75, 39 and 339 samples.



**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 29.55 %), **(B)** the transect method only (weighted mean regional prevalence is 35.55 %), **(C)** the seine nets only (weighted mean regional prevalence is 20.44 %), **(D)** the minnow traps (weighted mean regional prevalence is 19.20 %).



**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. **(A)** TN:TP ratio, **(B)** macrophyte coverage, **(C)** temperature, **(D)** turbidity, **(E)** pH, **(F)** dissolved oxygen, **(G)** conductivity and **(H)** Simpson’s Diversity Index are site-scale measurements while **(I)** area:perimeter ratio is a lake attribute. The perimeter model is not presented because it was highly non-linear and not interpretable although the model was significative (Figure S1).