

---

## Landscape models can predict the distribution of aquatic insects across agricultural areas

Gerber Rémi <sup>1,2,4</sup>, Piscart Christophe <sup>1,4</sup>, Roussel Jean--Marc <sup>2</sup>, Georges Romain <sup>1,4</sup>, Houet Thomas <sup>3</sup>, Royer Justine <sup>1,2</sup>, Bergerot Benjamin <sup>1,4,\*</sup>

<sup>1</sup> Univ Rennes, CNRS, ECOBIO - UMR 6553, 35042, Rennes, France

<sup>2</sup> DECOD (Ecosystem Dynamics and Sustainability), INRAE, Institut Agro, IFREMER, Rennes, France

<sup>3</sup> Univ Rennes 2, CNRS, LETG - UMR 6554, 35042, Rennes, France

<sup>4</sup> LTSER ZA Armorique, 35000, Rennes, France

\* Corresponding author : Benjamin Bergerot, email address : [benjamin.bergerot@univ-rennes.fr](mailto:benjamin.bergerot@univ-rennes.fr)

---

### Abstract :

#### Context

Adult aquatic insects are major vectors of aquatic subsidies to terrestrial ecosystems, providing substantial ecosystem services to agriculture. Distance to water is reported as the prime factor to model their spatial distribution, but the role of local drivers, either terrestrial or aquatic habitat features, is not well established, notably in highly heterogeneous landscapes.

#### Objectives

We assess the distribution and prioritise explanatory factors of aquatic insects in agricultural landscapes.

#### Methods

Adult aquatic insects (Ephemeroptera, Plecoptera, Trichoptera) were caught with sticky traps on a regular grid-pattern set across a heterogeneous agricultural landscape during two years. Using innovative landscape variables related to water, woodland, elevation and insolation (derived from drone data), classical random forest models were run to predict the distribution of insects.

#### Results

Variables that significantly explained insect distribution differed between insect orders, but the distance to water was consistently amongst the three most informative variables in models. Ephemeroptera occurrence was influenced by woodland density, while Plecoptera and Trichoptera were more influenced by potential insolation. Predictions revealed a low distribution pattern for Ephemeroptera and a high probability of occurrence for Plecoptera across terrestrial ecosystems. Predictions were poor for Trichoptera, perhaps reflecting various dispersal patterns among species that emerged earlier or later in spring and from a wide range of aquatic habitats.

---

## Conclusions

This work is the first to predict the probability of occurrence of aquatic insects across a spatially heterogeneous agricultural landscape. Prediction maps suggest that the aquatic footprint, defined as the occurrence of aquatic insects across a terrestrial landscape, is higher than previously thought.

**Keywords** : Stoneflies, Mayflies, Caddisflies, Random forest, Merolimnic insect

## 48 **Introduction**

49 Aquatic insects constitute major biological flows from aquatic to terrestrial ecosystems (Polis et al.  
50 1997; Nakano and Murakami 2001). After their aquatic larval stage, numerous insect species emerge  
51 as winged adults and disperse across the terrestrial ecosystems. Characterizing the spatial and temporal  
52 magnitude of such movements is central to understand the associated ecological functions for receiving  
53 ecosystems, as well as potential services to agriculture (Raitif et al. 2019). Notably, various aquatic  
54 insect species can feed on pollen during their adult stage (e.g., Winterbourn 2005; Hass et al. 2018),  
55 and thus have the potential to contribute to pollination while agroecosystems suffer a global decline in  
56 pollinators (Potts et al. 2010). These aquatic insects also provide nutritional subsidies to terrestrial  
57 consumers in the riparian zone and farther inland from banksides (Baxter et al. 2005; Uno 2016). They  
58 disseminate essential polyunsaturated fatty acids from aquatic environments to terrestrial ecosystems  
59 (Martin-Creuzburg et al. 2017; Mathieu-Resuge et al. 2021), and they are recognized as key elements  
60 of terrestrial food web functioning (Wesner 2010; Lafage et al. 2019). Many terrestrial predators are  
61 natural enemies of pest outbreaks in cropland (e.g., spiders and carabids, Riechert and Lockley 2003;  
62 Paetzold et al. 2005; Symondson et al. 2006; Stenroth et al. 2015), and can benefit from aquatic prey  
63 that usually emerges when terrestrial prey is scarce. Significant quantities of winged insects can emerge  
64 from lakes, reaching 1,200 to 2,500 kg.ha<sup>-1</sup>.yr<sup>-1</sup> (Gratton et al. 2008) and generally exceeding that of  
65 rivers (Williams et al. 2004; Bartrons et al. 2013). In agricultural landscapes however, 14 to 74 kg.ha<sup>-1</sup>  
66 yr<sup>-1</sup> of insect dry mass can emerge from streams (Raitif et al. 2018), among which 12.5 kg.ha<sup>-1</sup>.yr<sup>-1</sup>,  
67 on average, ultimately fell to the ground (Raitif et al. 2022). This is a substantial source of fertilization  
68 for soils (Jefferies et al. 2004). Notably, stream-derived N deposition on land can exceed the current  
69 amounts of atmospheric N deposition in temperate regions (Raitif et al. 2019). These study cases on  
70 winged aquatic insects highlight that the mechanisms governing their dispersal and distribution must  
71 be understood before appropriate prediction of their distribution across terrestrial areas, and latent  
72 beneficial impacts to agriculture, can be made.

73 To date, the most studied factor to explain the distribution of winged aquatic insects is the  
74 distance to water, their abundance decreasing rapidly farther inland from aquatic edges after a few  
75 meters (Svensson 1974; Miller et al. 2002; Carlson et al. 2016). Dispersal distances likely vary between  
76 species (Muehlbauer et al. 2014; Peredo Arce et al. 2021), but also depend on the aquatic habitat  
77 considered (Gratton and Vander Zanden 2009; Schindler and Smits 2017). To date, a majority of  
78 studies have focused on a single habitat type, mostly permanent streams or lakes (Gratton and Vander  
79 Zanden 2009; Carlson et al. 2016; Mathieu-Resuge et al. 2021). Yet many invertebrate species live in  
80 other temporary or permanent environments such as ponds, ditches or irrigation canals (Williams 1997;  
81 Labat et al. 2022). These habitats may represent a substantial proportion of surface water in these  
82 landscapes (Davies et al. 2010), more particularly in agricultural areas (Williams et al. 2004). In a  
83 recent study, Gerber et al. (2022a) raised the issue of defining aquatic habitats that serve as potential  
84 sources of emerging insects, highlighting the need for fine and thorough hydrographic mapping for  
85 studying aquatic insects' dispersal and distribution at the landscape scale. Additionally, several  
86 landscape features likely modulate the distribution of winged aquatic insects dispersing across  
87 terrestrial ecosystems, but again, underlying mechanisms are poorly characterised. It has been  
88 reported that riparian woodland density can hamper the dispersal of emerging chironomids and  
89 caddisflies (Delettre and Morvan 2000; Greenwood 2014). Temperature, woodland edges and  
90 topography are other variables that control the flight and movements of terrestrial insects across the  
91 landscape (Mattila 2015; Tonkin et al. 2018; Sweaney et al. 2022); however, they have been rarely  
92 investigated for winged aquatic insects. Some studies have shown the importance of temperature for  
93 aquatic insects. In the case of specific Diptera such as the crane fly, *Tipula maxima* (Jourdan et al.  
94 2019), dispersal-related traits vary depending on ambient environmental conditions (temperature  
95 regimes, discharge patterns and biotic interactions during individual development). For adult  
96 Plecoptera, air temperature acts as a factor that potentially influences longevity (Collier and Smith

97 2010). This result suggests that maintenance of appropriate microclimatic conditions should be a  
98 consideration in riparian management (Petersen et al. 2004). Consequently, the biophysical  
99 characteristics of riparian zones and, more broadly, terrestrial areas adjacent to aquatic environments,  
100 are an important factor to consider when studying the distribution of aquatic insects. Accordingly,  
101 much research attention has focused on how terrestrial environmental features regulate subsidy  
102 dispersal and the responses of terrestrial consumers (e.g., forest cover, complexity of ground habitats,  
103 availability of terrestrial resources), and on how human activities alter these linkages (Petersen et al.  
104 2004). For example, small aquatic insects (Nematocera) were most abundant under agricultural land  
105 use, whereas larger bodied aquatic insects (Plecoptera and Trichoptera) were more associated with  
106 forest land use (Stenroth et al. 2015).

107 In this work, we aim at (1) identifying landscapes features, defined hereafter as explanatory  
108 factors, that explain the occurrence of some aquatic insects (i.e., Ephemeroptera, Plecoptera and  
109 Trichoptera, namely EPTs) and (2) developing a spatially explicit statistical model to predict their  
110 potential distribution across an agricultural landscape. Investigations were performed in a 140ha  
111 agricultural area including meadows, crops, woodlands and a variety of aquatic ecosystems (stream,  
112 ponds and ditches). Winged EPTs were collected during their aerial dispersal across the area in spring  
113 and summer for two consecutive years. Random forest models were run on high-resolution landscape  
114 variables (e.g., distance to water, potential insolation, elevation and woodland density) extracted from  
115 drone data to predict the distribution of EPTs across the experimental area. Notably, we hypothesized  
116 that distance from water edge negatively correlates with the probability of occurrence of aquatic  
117 insects, while high densities of riparian trees along streams could hamper inland distribution. We also  
118 investigated differences among species owing to their ecological preference for woody and shaded or  
119 open and insolated habitats.

120

## 121 **Methods**

### 122 ***Study area***

123 The study site was located at the northeast of Brittany (48°36' N, 1° 32' W, France), is part of the  
124 European LTSER (Long-Term Socio-Ecological Research) network and occurs within the “Zone  
125 Atelier Armorique” (ZAAR). The ZAAR is a temperate region with no dry season and warm summers.  
126 The area is representative of an intensive agricultural area dominated by dairy production, composed  
127 of small fields separated by a hedgerow network called “bocage” (Baudry et al. 2000). Annual crops  
128 in this area are winter cereals and maize (Fig. 1A). The study area was selected based on the  
129 representativeness of the different landscape elements within it, based on the Shannon Diversity Index  
130 (SHDI) using Chloé software (Boussard and Baudry 2017). The average SHDI index for the ZAAR is  
131  $0.29 \pm 0.10$  (calculated at a resolution of 1 km by 1 km) and that for the study area is  $0.39 \pm 0.12$ . The  
132 140ha study area consisted of a patchwork landscape of woodland, hedgerows, pastures, croplands,  
133 and agricultural areas (Fig. 1B), and is crossed by a 33.2 km long first-order (Strahler 1952) stream  
134 named Le Guyoult.

### 135 ***Sampling***

136 Emerging EPTs were caught using interception traps called sticky traps (Raitif et al. 2022).  
137 Sticky traps consisted of a transparent A3 plastic cover (42 x 29.7 cm, 2 mm thick) on which was  
138 spread a thin layer of non-drying glue (Tanglefoot®, Tangle-Trap® Sticky Coatings). The plastic cover  
139 sprayed with glue was fixed cylindrically around a stake at 1 m height to catch insects coming from all  
140 possible directions (Wesner 2010). Hoops placed inside the cylinder reinforced the structure (Fig. 1C).  
141 The sticky traps were set in a grid pattern spaced out by at least 50 m to cover the entire study site  
142 homogeneously.

143 Two sampling campaigns were carried out, in spring and summer in 2020 and 2021, from April  
144 to June, which is the period of emergence of most aquatic insects in Brittany (Raitif et al. 2018). In

145 2020, 56 sticky traps were placed. In 2021, 64 sticky traps were placed, among which 20 were co-  
146 localized with those in 2020 (Fig. 1B). The plastic covers spread with glue remained in place for 15  
147 consecutive days for each sampling session before being retrieved and frozen in the laboratory.  
148 Monitoring of the traps throughout the emergence period has made it possible to visualize the  
149 emergence peaks of the insects, which are dependent on the weather. Therefore, two sampling sessions  
150 were conducted in 2020, from May 18<sup>th</sup> to June 2<sup>nd</sup> and from June 2<sup>nd</sup> to June 16<sup>th</sup>. In 2021, three  
151 sampling sessions were conducted, from April 1<sup>st</sup> to April 15<sup>th</sup>, from April 26<sup>th</sup> to May 10<sup>th</sup>, and from  
152 May 20<sup>th</sup> to June 3<sup>rd</sup>. For each sampling session, only the plastic covers were replaced.

153 In the laboratory, EPTs were extracted from the glue using D-limonene terpene and stored in  
154 96°C alcohol. All specimens were identified using a stereomicroscope (Leica M205 C) equipped with  
155 a binocular camera (Leica DMC4500) at the species or genus level using several identification keys  
156 (Ephemeroptera: Elliott 1983; Bauernfeind and Soldán 2013, Plecoptera: Despax 1951; Lubini et al.  
157 2012, Trichoptera: Malicky 2010; Barnard and Ross 2012).

### 158 ***Selected landscape features as explanatory factors***

159 *Water maps.* We characterized the full hydrographic networks including all water bodies (i.e., the  
160 mainstream, the tributaries, the ditches, and the water-full ponds at the time of the survey). For 2020,  
161 the hydrographic networks used were directly extracted from Gerber et al. (2022a). For 2021,  
162 hydrographic networks were characterized at four successive dates (April 12<sup>th</sup>, May 5<sup>th</sup>, May 27<sup>th</sup>, and  
163 June 29<sup>th</sup>). We divided the water bodies into five aquatic habitats: the mainstream, the tributaries,  
164 ditches with running water, ditches with standing water, and ponds. From the aquatic habitats and the  
165 four dates of mapping, we defined 14 hydrographic networks. For each year and each order, we defined  
166 the hydrographic network that best explains the distribution of the aquatic insects following Gerber et  
167 al. (2022a). All the details are provided in the Supplementary information (S1).

168        *Woodland maps.* On May 25<sup>th</sup> in 2020, a drone survey was performed to characterize the landscape  
169 features of the study site. A fixed-wing (sensefly eBee X) was used to survey the whole area at once.  
170 The flight altitude was 120 m above ground level. Using a RGB camera (sensefly soda 3D, 20M pixels)  
171 and a photogrammetric approach, a RGB orthorectified map and a digital surface model (DSM) of 0.04  
172 m and a mean XYZ geometrical horizontal and vertical accuracy of 0.016 m were obtained (Houet  
173 2022). Using eCognition software, homogenous polygons according to their color and height were  
174 segmented. All woodland polygons were then manually classified from RGB visual interpretation and  
175 rasterized at 0.05 m. Resolution was checked using a trimble geo7x Dgpps, which recorded the absolute  
176 position of 10 targets covered by the drone's flight.

177        *Elevation maps.* We used a digital elevation model (DEM) with 1 m resolution obtained from the  
178 French National Institute of Geographic and Forest Information (“Institut national de l'information  
179 géographique et forestière”, <https://ign.fr/>)

180        *Potential insolation maps.* The potential insolation mapping (Potential Incoming Solar Radiation  
181 kWh.m<sup>-2</sup>, PISR) was calculated using the SAGA software v.7.9 (Conrad et al. 2015) for each sampling  
182 session. PISR represents the sum of the solar energy per square meter and is calculated every hour  
183 based on the position of the sun and the DSM. Shadowed areas near woodlands will exhibit a lower  
184 PISR value than an area in the middle of a field. The PISR values are calculated for all the points of  
185 the DSM, thus on the canopy. The sticky traps are sometimes located under trees, and therefore receive  
186 much lower insolation than that received by the canopy. To take this into account, we removed the  
187 wooded areas from the PISR. Then the missing values were interpolated from edges using the GDAL  
188 translator library v. 3.5, *fillnodata* function.

189        *Explanatory factors.* We derived 19 biophysical factors from the environmental maps described  
190 (Supplementary information, Table S2). Movement distances may vary according to the insect orders,  
191 land uses, and models used to estimate the abundance decrease with distance from water. For example,



192 previous studies differed in the distance over which Chironomidae abundance decreased by 50%, from  
193 13 m to 25 m, and of Trichoptera from 1.3 m to 5 m (Muehlbauer et al. 2014; Raitif et al. 2022). This  
194 potentially modifies the extent of the variables effect. Hence, we used three buffer sizes because the  
195 scale at which environmental factors may influence aquatic insect distribution is not well known. The  
196 woodland density (the number of wooded pixels divided by the area) and the mean and variance of the  
197 elevation and potential insolation were calculated for each sticky trap within three scale buffers: a 5 m  
198 radius buffer, a 5 m to 25 m radius ring buffer (a 25 m buffer without the 5 m radius buffer) and a 25  
199 m to 50 m radius ring buffer (a 50 buffer without the 25 m radius buffer) using *zonal statistics* function  
200 GIS software v.3.22 (QGIS Association 2022). Then, for each sticky trap, the distance from the nearest  
201 wooded area was calculated using the *st\_distance* function from the *sf* package (Pebesma 2018). The  
202 nearest aquatic habitat (i.e., mainstream, tributary, ditch with running water, ditch with standing water,  
203 or pond) was determined using the *st\_join* function from the *sf* package.

204 To help with the interpretation of the models, we explored the relationships between the factors  
205 with Pearson coefficients (correlation matrix available in the Supplementary information, Fig. S3)  
206 because we suspect some factors to be correlated (e.g., insolation and woodland).

207

208

### 209 ***Statistical analyses***

210 To identify explanatory factors for predictive models, we calibrated the models by insect order  
211 (i.e., Ephemeroptera, Trichoptera and Plecoptera) with the random forest algorithm using the  
212 *randomForest* package (Liaw and Wiener 2002). Random forest, based on recursive partitioning, does  
213 not rely on assumptions about relationship between the explanatory variable and the response variable  
214 (as in GLM or GAM models). This approach is entirely data-driven and reduces the associated variance

215 of the prediction (Breiman 2001). Therefore, within the limit of the data, random forest which  
216 maximizes the reduction of deviance is more suitable for prediction than GLM or GAM (Guisan et al.  
217 2017).

218 The response variables were the occurrences of EPT orders and the explanatory variables were  
219 the environmental factors and the sampling sessions. We used the occurrence of different orders  
220 because it allows us to validate our prediction, and to compare our results with other studies. We  
221 followed the procedure described by Genuer and Poggi (2020). According to this procedure, the  
222 number of trees was fixed at 10,000. The number of factors tested at each division minimizing the out-  
223 of-bag error (OOB error estimates of the prediction error using the unselected observations in a  
224 bootstrap sample as test data, Genuer and Poggi 2020) was chosen. For each order, we kept a session  
225 when the occurrence rate was higher than 10% to have enough occurrences to calibrate the models  
226 (Fig. 2). We used the dataset from 2020 to calibrate the model because the predictions were better  
227 compared to 2021 (Supplementary information, Fig. S4).

228 For predictions, we used the *VSURF* function from the *VSURF* package (Genuer et al. 2015). It  
229 is a three-step variable selection procedure based on random forests. The first step is dedicated to  
230 eliminating irrelevant variables from the dataset according to the Mean Decrease Accuracy, a measure  
231 that expresses how much accuracy the model losses by excluding each variable. The second step aims  
232 to select all variables related to the response for interpretation purposes. The model is built with  
233 variables selected in the first step and variables with the lowest OOB errors. The third step refines the  
234 selection by eliminating redundancy in the set of variables selected by the second step, for prediction  
235 purposes. Variables are kept if the OOB error decreases more than the average variation, thus limiting  
236 the correlations between the variables. Finally, the remaining variables are added in order of increasing  
237 importance. Model validation was made using the area under curve (AUC) criteria. AUC is a standard  
238 measure often applied to random forest (Guisan et al. 2017). It measures the discrimination (i.e., the

239 ability to distinguish between occupied and unoccupied sites, Guisan et al. 2017) by comparing the  
240 order of occurrences observed in 2021 with the predictions of the model based on environmental  
241 variables in 2021. AUC ranges in value from 0 to 1. A model whose predictions are 100% wrong has  
242 an AUC of 0, and one whose predictions are 100% correct has an AUC of 1. An AUC of 0.5 suggests  
243 no discrimination, 0.7 to 0.8 is considered acceptable and 0.8 to 0.9 is considered excellent (Guisan et  
244 al. 2017).

245 Finally, for the whole area, we created predictive maps of the probability of ETPs occurrences  
246 based on the explanatory factors selected by the *VSURF* procedure. To obtain predictive maps of the  
247 whole area, we computed for each selected factor its map in 2021. Distance maps (for woodland and  
248 water) were calculated using Chloé software v. 4.1 (Boussard and Baudry 2017). Maps of density,  
249 mean and variance (for woodland, potential insolation, and elevation) were calculated using moving  
250 windows. A moving window assigns to each pixel the value of a function (e.g., mean, variance) in a  
251 window centered on each of these pixels, using the *focal* function from the *terra* package (Hijmans et  
252 al. 2023). Moving windows sizes and shapes were identical to the three scale buffers so that the pixel  
253 values are comparable to the corresponding factor values for each of the sticky traps. All the maps  
254 were resampled at 5 m resolution using the nearest neighbor method, *aggregate* function, in the *terra*  
255 package.

256 All statistical analyses were performed using the R software v. 4.2.1 (R Core Team 2021).

257

258

## 259 **Results**

260 A total of 748 adult aquatic insects (EPTs) were sampled. Among Ephemeroptera, 91  
261 individuals were counted and 6 species identified. The most represented species was *Habrophlebia*  
262 *fusca* (80.2%). The distances at which the abundance of Ephemeroptera was at 50% and 10% of its

263 near-water source levels were 1.1 and 6.5 m, respectively. Among Plecoptera, 484 individuals were  
264 counted and 6 species identified. The three most abundant species were *Nemoura cinerea* (60.33%),  
265 *Nemoura lacustris* (15.49%) and *Amphinemoura standfussi* (13.42%). The distances at which the  
266 abundance of Plecoptera was at 50% and 10% of its near-water source levels were 2.5 and 109 m,  
267 respectively. Among Trichoptera, 128 individuals were counted and 9 species identified. The four main  
268 species were *Stenophylax lateralis* (26.56%), *Beraea pullata* (25%), *Limnephilus centralis* (20.31%)  
269 and *Limnephilus sparsus* (10.15%). The distances at which the abundance of Trichoptera was at 50%  
270 and 10% of its near-water source levels were 3.4 and 304 m, respectively.

271

### 272 ***Factors explaining occurrences of aquatic insects***

273 Distance from water consistently ranked within the top three variables controlling the  
274 distribution of EPTs (Fig. 3). For Ephemeroptera, two correlated variables (Supplementary  
275 information, Fig. S3) stood out in addition to distance from water: woodland density and the mean of  
276 PISR in a 5 to 25 m ring. For Plecoptera, the top three variable controlling the distribution were the  
277 distance from water and average PISR at 5 m and to a lesser extent the woodland density in a 5 to 25  
278 m ring. For Trichoptera, the session was the main variable followed by the variance of PISR in a 25 to  
279 50 m ring and the distance from water.

280

### 281 ***Prediction***

282 For Trichoptera, the variance of PISR in a 25 to 50 m ring was retained for prediction through  
283 the *VSURF* procedure. For Ephemeroptera, the distance from water and the woodland density in 5 to  
284 25 m were retained. For Plecoptera, the distance from water and the mean of PISR in a 5 m buffer were  
285 retained. Predictions were good for Ephemeroptera (AUC = 0.86, Fig. 4) and nearly acceptable for

286 Plecoptera (AUC = 0.68, Fig. 4). For Trichoptera, the model gave poor predictions (AUC = 0.43, Fig.  
287 4). Thus, prediction maps were made for Ephemeroptera and Plecoptera (Fig. 5). The prediction map  
288 of Ephemeroptera showed visually three levels of probability of occurrence. The highest one represents  
289 areas close to the water with riparian forests (from 0.75 to 1) and occupies 16 % of the study area,  
290 including the hydrographic network (0.4 % of the area). Then, the probability is medium for wooded  
291 patches without or with small water bodies or for an area close to the water without riparian forests  
292 (from 0.25 to 0.75, occupying 16 % of the study area). Finally, the probability of occurrence is very  
293 low in the fields (from 0 to 0.25) which occupy 68 % of the study area. The prediction of Plecoptera  
294 exhibited a very high probability of occurrence in an approximately 12 m buffer around water bodies  
295 (> 0.9 of occurrence probability, 13 % of the study area, including the hydrographic network, i.e., 1.1%  
296 of the area). Most of the values from 0.75 to 0.9 (33 % of the study area) are in a 12 m to 50-60 m  
297 buffer from the water bodies.

298

## 299 **Discussion**

300 In this study, we highlighted the distribution extent of EPTs in terrestrial environments. Aquatic  
301 ecosystems are often considered physically well-defined. However the biological width of a stream or  
302 stream signature (which encompasses the aquatic insect subsidies on the terrestrial ecosystem) has  
303 been recently acknowledged (Muehlbauer et al. 2014; Gurnell et al. 2016; Kopp and Allen 2019). In  
304 this study, we considered not only streams but all the water bodies which can produce aquatic insects  
305 (e.g., pond, ditch, temporary tributaries) within a landscape mosaic. We included both the aquatic  
306 habitat for insect production and the terrestrial environment where the insects spread. We showed that  
307 aquatic insect fluxes to terrestrial environments cover the entire surrounding land, even for site located  
308 relatively far from water (several hundred meters). Furthermore, random forest analyses confirmed  
309 that the spatial distribution of adult EPTs is explained differently for each order.

310 Our results confirm that distance from water is a key factor explaining the inland distribution of  
311 aquatic insects and is one of the top three factors for all orders. Indeed, EPTs are more abundant near  
312 water bodies (Carlson et al. 2016; Peredo Arce et al. 2021). However, our analysis also highlights that  
313 while distance to water is a key parameter, other landscape factors also help explain the predictions for  
314 each order. For Ephemeroptera, the density of woodland within 25 m stands out. One explanation  
315 would be that the larvae of *Habrophlebia fusca*, the main species in our study, are shredders that  
316 consume leaf litter and are therefore common in wooded areas (Bauernfeind and Soldán 2013). The  
317 scale of 25 m for woodlands could therefore be related to larval habitat, as this species is known for its  
318 poor dispersal capability (Sarremejane et al. 2020). Ephemeroptera distribution also appears sensitive  
319 to potential insolation (PISR) between 5 and 25 m. This result is consistent with the hypothesis that an  
320 open environment promotes inland dispersal of aquatic insects. In an open environment, the absence  
321 of tall vegetation and exposure to higher wind speeds could lead to farther inland dispersal than that in  
322 forested landscapes (Delettre and Morvan 2000; Carlson et al. 2016). Potential insolation is also  
323 important for Trichoptera and Plecoptera, suggesting that open environments may also favor inland  
324 dispersal for these orders. This factor, which has not been examined in previous studies, may provide  
325 an interesting explanation for the inland movement of aquatic insects, probably because of its influence  
326 on flight. The amount of potential insolation may be stronger in open environments as is it highly  
327 negatively correlated with woodland density ( $\rho = -0.7$ , Supplementary information, Fig. S3). In  
328 forest, insolation values are very low and homogeneous due to the tree cover, which reduces its impact  
329 on insect distribution. For Trichoptera, we noticed a significant impact of the sampling session during  
330 the study. This result is probably due to the high diversity of species with very different biological  
331 traits and emergence times (Singh et al. 1984). The emergence of different species between each  
332 session may induce temporal and spatial heterogeneity between sessions which induces noise in our  
333 analyses.

334 To our knowledge, for the first time, we have provided a predictive map of the occurrence of  
335 some aquatic insects (i.e., Ephemeroptera and Plecoptera) at the landscape scale, thereby emphasizing  
336 the extent of stream influence (Gurnell et al. 2016; Gounand et al. 2018). According to the meta-  
337 analysis from Muehlbauer et al. (2014), these two orders have a similar extent (10 % of abundances at  
338 140 m for Plecoptera and 160 m for Ephemeroptera), but Ephemeroptera tends to stay closer than  
339 Plecoptera (50 % of abundances at 2.7 m for Plecoptera and 0.9 m for Ephemeroptera). However, our  
340 models predict a highly divergent distribution of adult Plecoptera and Ephemeroptera at the landscape  
341 scale. The probability of occurrence decreases rapidly for Ephemeroptera but does not stabilize at low  
342 values to zero, in contrast to other studies (Muehlbauer et al. 2014; Raitif et al. 2022), while the extent  
343 for Plecoptera is large and the probability of occurrence remains high even away from water. This  
344 difference may be explained by several factors. For Ephemeroptera, almost all species included in the  
345 Muehlbauer et al. (2014) meta-analysis are good dispersers and are more likely to be found far from  
346 water, unlike *Habrophlebia fusca* (Sarremejane et al. 2020), which dominated the Ephemeroptera  
347 assemblages in our study. This low dispersal capability of this species is also congruent to its body  
348 stores consisting primarily of triglycerides for reproduction and less glycogen and proteins for long-  
349 distance flight (Gerber et al. 2022b). Another possibility is the unsuitable sampling methods used to  
350 study dispersal and distribution at the landscape scale. Indeed, most emergence studies have sampled  
351 along transects from stream to terrestrial habitats. This protocol is efficient to study the linear dispersal  
352 from the stream but does not account for i) alternative sources of insects (pool, ditch, temporary  
353 running waters) (Gerber et al. 2022a) and ii) erratic insect movements due to wind and other landscape  
354 elements (e.g., woodland density). For example, in Muehlbauer et al. (2014), most studies were  
355 conducted in forests with homogeneous conditions that may favor dispersal over a longer distance from  
356 streams. Raitif et al. (2022) conducted a similar study but their furthest sampling point from the aquatic  
357 source was 50 m. They cannot therefore conclude anything about the real dispersal distance of

358 individuals beyond 50 m. Our sampling strategy was chosen to be suitable for studying dispersal at the  
359 landscape scale (i.e., beyond 50 m).

360 For Plecoptera, we found a very large extent of inland occurrence. The probability of occurrence  
361 could be relatively high up to 60 m around each water body. This result is surprising because most of  
362 Plecoptera species tend to be weak dispersers (Sarremejane et al. 2020). This result suggests that  
363 Plecoptera may disperse farther from the water than previously thought. As with Ephemeroptera, this  
364 difference with the literature may be explained by our sampling protocol, which is more adapted to  
365 studying spatial distribution at the landscape scale. Furthermore, this result is also in line with the  
366 higher fuels for flight in Plecoptera than in Ephemeroptera found by Gerber et al. (2022b).

367 The quality of the predictions differed strongly between Trichoptera, Plecoptera, and  
368 Ephemeroptera. Predictions for Plecoptera and Ephemeroptera were strong but remained weak for  
369 Trichoptera. We can make three suggestions to explain this. First, we observed that between the two  
370 sampling periods (i.e., 2020 and 2021), the taxonomical composition changed significantly for  
371 Trichoptera (11 species belonging to 9 genera in 2020 and only 4 belonging to 3 genera in 2021) and  
372 to a lesser extent for Plecoptera (the proportion of *Amphinemoura standfussi* strongly increased in 2021  
373 while the proportion of *Nemoura lacustris* decreased in samples). Within an order, different species  
374 may vary in flight ability (Goldsworthy and Wheeler 1989), behavior (Steyn et al. 2016), biological  
375 traits (Gerber et al. 2022b) and larval habitats (Tachet et al. 2000). These differences could explain the  
376 poor quality of predictions for Trichoptera. Second, the quality of the prediction could depend on the  
377 movement ability of the organisms. Indeed, most of the Trichoptera are considered good dispersers  
378 (Sarremejane et al. 2020), good flyers (Graham et al. 2017) and are frequently found far from water  
379 (Peredo Arce et al. 2021). In contrast, when organisms like Ephemeroptera and Plecoptera remain close  
380 to water (Muehlbauer et al. 2014), their distribution can be more easily modeled. Third, some  
381 explanatory variables could still be missing or need to be refined. For instance, hedgerow permeability,



382 defined by vegetation structure and density, could limit insect dispersal differently from one species to  
383 another. The use of other sensors, such as Lidar, would be highly contributive to provide additional  
384 explanatory variables. In the same way, measurements of surface temperature using a thermal camera  
385 may be interesting to explore. PISR could also be refined from one period to another by accounting for  
386 irradiance measurements available from weather station near the study site. Indeed, as shown,  
387 insolation seems to influence EPTs distribution although we did not account for real insolation during  
388 each of the sampling sessions. One sunny/cloudy period may positively or negatively affect the  
389 magnitude of EPTs distribution compared to another. One way to improve our predictions would be to  
390 adjust the resolution of our models, the sampling strategy (e.g., by considering later emergence periods)  
391 or even the drone flight surveys according to the flight ability of the species, as we did with the best  
392 hydrographic network for each sampling session. Indeed, in highly mobile organisms, the choice of  
393 spatial and temporal resolution greatly influences distribution models (Guisan et al. 2017). However,  
394 to our knowledge, no studies have attempted to model inland distribution of aquatic insects, limiting  
395 the information available on which resolution to use.

396 In conclusion, our results show that the spatial extent of emergent aquatic EPTs is much wider  
397 than expected in the literature, likely because sampling along linear transects is not suitable for  
398 studying the distribution of aquatic insect at the landscape scale with multiple aquatic sources.  
399 Moreover, the analysis of new landscape elements in our study confirms the role of distance to the  
400 water, independent of aquatic habitat type and of woodland for some species. It also highlights the role  
401 of new and untested variables such as the PISR. Finally, our models have been successfully used to  
402 predict the distribution of Ephemeroptera and Plecoptera at the landscape scale. This opens the door to  
403 studying their impacts in terrestrial ecosystems in terms of fertilization, pollination, or as food source  
404 for terrestrial predators at the landscape scale.

405

406

407 **Acknowledgements**

408           We would like to thank the two reviewers and the editor who helped us to improve this work.

409 We gratefully acknowledge all the farmers who allowed us to set up the sticky traps on their fields.

410 **Author contributions**

411           RGer, CP, JMR and BB conceived the ideas and designed methodology; RGer, CP, JMR, RGeo,

412 JR and BB collected the data; TH performed UAV flights; RGer, RGeo, TH, and BB analyzed the

413 data; Rger, CP, JMR and BB led the writing of the manuscript. All authors contributed critically to the

414 drafts and gave final approval for publication.

415

416 **Funding**

417           This research was funded by two grants from the LTSER Zone Atelier Armorique and Earth

418 Sciences and Astronomy Observatory (OSU) Rennes, OSUR (DIPATE).

419 **Declarations**

420 **Competing interests**

421           The authors have no relevant financial or non-financial interests to disclose.

422

423 **Data availability**

424           The data are available from the corresponding author on reasonable request.

425

426 **References**

427

428 Barnard PC, Ross E (2012) The Adult Trichoptera (Caddisflies) of Britain and Ireland. Royal

429 Entomological Society, St Albans

- 430 Bartrons M, Papeş M, Diebel MW, et al (2013) Regional-level inputs of emergent aquatic insects from  
431 water to land. *Ecosystems* 16:1353–1363. <https://doi.org/10.1007/s10021-013-9688-6>
- 432 Baudry J, Bunce RGH, Burel F (2000) Hedgerows: An international perspective on their origin,  
433 function and management. *J Environ Manage* 60:7–22.  
434 <https://doi.org/10.1006/jema.2000.0358>
- 435 Bauernfeind E, Soldán T (2013) *The Mayflies of Europe (Ephemeroptera)*. Brill, Leiden
- 436 Baxter CV, Fausch KD, Saunders WC (2005) Tangled webs: reciprocal flows of invertebrate prey link  
437 streams and riparian zones. *Freshw Biol* 50:201–220. <https://doi.org/10.1111/j.1365-2427.2004.01328.x>
- 439 Boussard H, Baudry J (2017) *Chloe4.0: A software for landscape pattern analysis*
- 440 Breiman L (2001) Random Forests. *Mach Learn* 45:5–32. <https://doi.org/10.1023/A:1010933404324>
- 441 Carlson PE, McKie BG, Sandin L, Johnson RK (2016) Strong land-use effects on the dispersal patterns  
442 of adult stream insects: implications for transfers of aquatic subsidies to terrestrial consumers.  
443 *Freshwater Biology* 61:848–861. <https://doi.org/10.1111/fwb.12745>
- 444 Collier KJ, Smith BJ (2010) Interactions of adult stoneflies (Plecoptera) with riparian zones I. Effects  
445 of air temperature and humidity on longevity. *Aquat Insects* 22:275–284.  
446 [https://doi.org/10.1076/0165-0424\(200010\)22:4;1-Y;FT275](https://doi.org/10.1076/0165-0424(200010)22:4;1-Y;FT275)
- 447 Conrad O, Bechtel B, Bock M, et al (2015) System for automated geoscientific analyses (SAGA) v.  
448 2.1.4. *Geosci Model Dev* 8:1991–2007. <https://doi.org/10.5194/gmd-8-1991-2015>
- 449 Davies BR, Biggs J, Williams PJ, et al (2010) A comparison of the catchment sizes of rivers, streams,  
450 ponds, ditches and lakes: implications for protecting aquatic biodiversity in an agricultural  
451 landscape. In: Oertli B, Céréghino R, Biggs J, et al. (eds) *Pond Conservation in Europe*.  
452 Springer Netherlands, Dordrecht, pp 7–17
- 453 Delettre YR, Morvan N (2000) Dispersal of adult aquatic Chironomidae (Diptera) in agricultural  
454 landscapes. *Freshw Biol* 44:399–411. <https://doi.org/10.1046/j.1365-2427.2000.00578.x>
- 455 Despax R (1951) *Faune de France: Plécoptère*, Fédération Française des Sociétés de Sciences  
456 Naturelles
- 457 Elliott JM (1983) *A key to the adults of the British Ephemeroptera with note on their ecology*, 1st  
458 edition. Freshwater Biological Association, Ambleside, Cumbria
- 459 Genuer R, Poggi J-M (2020) *Random forests with R*. Springer, Cham
- 460 Genuer R, Poggi J-M, Tuleau-Malot C (2015) VSURF: An R Package for Variable Selection Using  
461 Random Forests. *The R Journal* 7:19. <https://doi.org/10.32614/RJ-2015-018>
- 462 Gerber R, Bergerot B, Georges R, et al (2022a) The Role of Hydrographic Mapping in the Study of  
463 Emerging Aquatic Insects on the Landscape Scale. *Water* 14:3019.  
464 <https://doi.org/10.3390/w14193019>

- 465 Gerber R, Cabon L, Piscart C, et al (2022b) Body stores of emergent aquatic insects are associated  
466 with body size, sex, swarming behaviour, and dispersal strategies. *Freshw Biol* 67:2161–2175
- 467 Goldsworthy GJ, Wheeler CH (eds) (1989) *Insect Flight*. CRC Press, Boca Raton
- 468 Gounand I, Harvey E, Little CJ, Altermatt F (2018) Meta-Ecosystems 2.0: Rooting the theory into the  
469 field. *Trends Ecol Evol* 33:36–46. <https://doi.org/10.1016/j.tree.2017.10.006>
- 470 Graham SE, Storey R, Smith B (2017) Dispersal distances of aquatic insects: upstream crawling by  
471 benthic EPT larvae and flight of adult Trichoptera along valley floors. *N Z J Mar Freshwater*  
472 *Res* 51:146–164. <https://doi.org/10.1080/00288330.2016.1268175>
- 473 Gratton C, Donaldson J, Zanden Vander MJ (2008) Ecosystem Linkages between Lakes and the  
474 Surrounding Terrestrial Landscape in Northeast Iceland. *Ecosystems* 11:764–774
- 475 Gratton C, Vander Zanden MJ (2009) Flux of aquatic insect productivity to land: comparison of lentic  
476 and lotic ecosystems. *Ecology* 90:2689–2699. <https://doi.org/10.1890/08-1546.1>
- 477 Greenwood MJ (2014) More than a barrier: The complex effects of ecotone vegetation type on  
478 terrestrial consumer consumption of an aquatic prey resource. *Austral Ecol* 39:941–951.  
479 <https://doi.org/10.1111/aec.12159>
- 480 Guisan A, Thuiller W, Zimmermann NE (2017) *Habitat Suitability and Distribution Models, with*  
481 *Applications in R*. Cambridge University Press, Cambridge
- 482 Gurnell M, Bertoldi W, Tockner K, et al (2016) How large is a river? Conceptualizing river landscape  
483 signatures and envelopes in four dimensions. *Wiley Interdisciplinary Reviews: Water* 313–  
484 321. <https://doi.org/10.1002/wat2.1143>
- 485 Hass AL, Kormann UG, Tschardt T, et al (2018) Landscape configurational heterogeneity by small-  
486 scale agriculture, not crop diversity, maintains pollinators and plant reproduction in western  
487 Europe. *Proc R Soc B-Biol Sci* 285:20172242. <https://doi.org/10.1098/rspb.2017.2242>
- 488 Hijmans RJ, Bivand R, Pebesma E, Sumner MD (2023) *terra: Spatial Data Analysis*
- 489 Houet T (2022) Jeu de données Drone de la Timbertière (Zone Atelier Armorique).  
490 [https://portail.indigeo.fr/geonetwork/srv/api/records/57833b75-9ed7-4e03-8673-](https://portail.indigeo.fr/geonetwork/srv/api/records/57833b75-9ed7-4e03-8673-4c48ea7d2159)  
491 [4c48ea7d2159](https://portail.indigeo.fr/geonetwork/srv/api/records/57833b75-9ed7-4e03-8673-4c48ea7d2159). In: Catalogue de métadonnées d'indigeo.  
492 [https://portail.indigeo.fr/geonetwork/srv/api/records/57833b75-9ed7-4e03-8673-](https://portail.indigeo.fr/geonetwork/srv/api/records/57833b75-9ed7-4e03-8673-4c48ea7d2159)  
493 [4c48ea7d2159](https://portail.indigeo.fr/geonetwork/srv/api/records/57833b75-9ed7-4e03-8673-4c48ea7d2159). Accessed 14 Sep 2022
- 494 Jourdan J, Baranov V, Wagner R, et al (2019) Elevated temperatures translate into reduced dispersal  
495 abilities in a natural population of an aquatic insect. *J Anim Ecol* 88:1498–1509.  
496 <https://doi.org/10.1111/1365-2656.13054>
- 497 Kopp DA, Allen DC (2019) Stream network geometry and the spatial influence of aquatic insect  
498 subsidies across the contiguous United States. *Ecosphere* 10:e02926.  
499 <https://doi.org/10.1002/ecs2.2926>

- 500 Labat F, Piscart C, Thiébaud G (2022) Invertebrates in small shallow lakes and ponds: a new sampling  
501 method to study the influence of environmental factors on their communities. *Aquat Ecol*  
502 56:585–603. <https://doi.org/10.1007/s10452-021-09939-1>
- 503 Lafage D, Bergman E, Eckstein RL, et al (2019) Local and landscape drivers of aquatic-to-terrestrial  
504 subsidies in riparian ecosystems: a worldwide meta-analysis. *Ecosphere* 10:e02697.  
505 <https://doi.org/10.1002/ecs2.2697>
- 506 Liaw A, Wiener M (2002) Classification and regression by random forest. 2:.  
507 <https://cogns.northwestern.edu/cbmg/LiawAndWiener2002.pdf>
- 508 Lubini V, Knispel S, Vinçon G (2012) Die Steinfliegen der Schweiz: Bestimmung und Verbreitung /  
509 Les pléoptères de Suisse. Identification et distribution, Fauna Helvetica
- 510 Malicky H (2010) Atlas of European Trichoptera/ Atlas Der Europäischen Kockerfliegen/ Atlas Des  
511 Trichopteres D'europe, 2nd ed. Softcover of orig. ed. 2004 édition. Springer
- 512 Martin-Creuzburg D, Kowarik C, Straile D (2017) Cross-ecosystem fluxes: Export of polyunsaturated  
513 fatty acids from aquatic to terrestrial ecosystems via emerging insects. *Sci Total Environ*  
514 577:174–182. <https://doi.org/10.1016/j.scitotenv.2016.10.156>
- 515 Mathieu-Resuge M, Martin-Creuzburg D, Twining CW, et al (2021) Taxonomic composition and lake  
516 bathymetry influence fatty acid export via emergent insects. *Freshw Biol* 66:2199–2209.  
517 <https://doi.org/10.1111/fwb.13819>
- 518 Mattila ALK (2015) Thermal biology of flight in a butterfly: genotype, flight metabolism, and  
519 environmental conditions. *Ecol Evol* 5:5539–5551. <https://doi.org/10.1002/ece3.1758>
- 520 Miller MP, Blinn DW, Keim P (2002) Correlations between observed dispersal capabilities and  
521 patterns of genetic differentiation in populations of four aquatic insect species from the Arizona  
522 White Mountains, U.S.A. *Freshw Biol* 47:1660–1673. <https://doi.org/10.1046/j.1365-2427.2002.00911.x>
- 524 Muehlbauer JD, Collins SF, Doyle MW, Tockner K (2014) How wide is a stream? Spatial extent of  
525 the potential “stream signature” in terrestrial food webs using meta-analysis. *Ecology* 95:44–  
526 55. <https://doi.org/10.1890/12-1628.1>
- 527 Nakano S, Murakami M (2001) Reciprocal Subsidies: Dynamic Interdependence between Terrestrial  
528 and Aquatic Food Webs. *Proceedings of the National Academy of Sciences of the United States*  
529 *of America* 98:166–170. <http://www.jstor.org/stable/3054647>
- 530 Paetzold A, Schubert CJ, Tockner K (2005) Aquatic terrestrial linkages along a braided-river: Riparian  
531 arthropods feeding on aquatic insects. *Ecosystems* 8:748–759. <https://doi.org/10.1007/s10021-005-0004-y>
- 533 Pebesma E (2018) Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*  
534 10:439. <https://doi.org/10.32614/RJ-2018-009>

535 Peredo Arce A, Hörren T, Schletterer M, Kail J (2021) How far can EPTs fly? A comparison of  
536 empirical flying distances of riverine invertebrates and existing dispersal metrics. *Ecol Indic*  
537 125:107465. <https://doi.org/10.1016/j.ecolind.2021.107465>

538 Petersen I, Masters Z, Hildrew AG, Ormerod SJ (2004) Dispersal of adult aquatic insects in catchments  
539 of differing land use. *J Appl Ecol* 41:934–950. [https://doi.org/10.1111/j.0021-](https://doi.org/10.1111/j.0021-8901.2004.00942.x)  
540 8901.2004.00942.x

541 Polis GA, Anderson WB, Holt RD (1997) Toward an Integration of Landscape and Food Web Ecology:  
542 The Dynamics of Spatially Subsidized Food Webs. *Annu Rev Ecol Syst* 28:289–316.  
543 <https://doi.org/10.1146/annurev.ecolsys.28.1.289>

544 Potts SG, Biesmeijer JC, Kremen C, et al (2010) Global pollinator declines: trends, impacts and drivers.  
545 *Trends Ecol Evol* 25:345–353. <https://doi.org/10.1016/j.tree.2010.01.007>

546 QGIS Association (2022) QGIS Geographic Information System

547 R Core Team (2021) R: A language and environment for statistical computing

548 Raitif J, Plantegenest M, Agator O, et al (2018) Seasonal and spatial variations of stream insect  
549 emergence in an intensive agricultural landscape. *Sci Total Environ* 644:594–601.  
550 <https://doi.org/10.1016/j.scitotenv.2018.07.021>

551 Raitif J, Plantegenest M, Roussel J-M (2019) From stream to land: Ecosystem services provided by  
552 stream insects to agriculture. *Agric Ecosyst Environ* 270–271:32–40.  
553 <https://doi.org/10.1016/j.agee.2018.10.013>

554 Raitif J, Roussel J-M, Olmos M, et al (2022) Assessing spatial deposition of aquatic subsidies by  
555 insects emerging from agricultural streams. *Sci Total Environ* 837:155686.  
556 <https://doi.org/10.1016/j.scitotenv.2022.155686>

557 Riechert SE, Lockley T (2003) Spiders as Biological Control Agents. *Annu Rev Entomol* 29:299–320.  
558 <https://doi.org/10.1146/annurev.en.29.010184.001503>

559 Sarremejane R, Cid N, Stubbington R, et al (2020) DISPERSE, a trait database to assess the dispersal  
560 potential of European aquatic macroinvertebrates. *Sci Data* 7:386.  
561 <https://doi.org/10.1038/s41597-020-00732-7>

562 Schindler DE, Smits AP (2017) Subsidies of Aquatic Resources in Terrestrial Ecosystems. *Ecosystems*  
563 20:78–93. <https://doi.org/10.1007/s10021-016-0050-7>

564 Singh MP, Smith SM, Harrison AD (1984) Emergence of some caddisflies (Trichoptera) from a  
565 wooded stream in southern Ontario. *Hydrobiologia* 112:223–232.  
566 <https://doi.org/10.1007/BF00008088>

567 Stenroth K, Polvi LE, Fältström E, Jonsson M (2015) Land-use effects on terrestrial consumers through  
568 changed size structure of aquatic insects. *Freshw Biol* 60:136–149.  
569 <https://doi.org/10.1111/fw.12476>

570 Steyn VM, Mitchell KA, Terblanche JS (2016) Dispersal propensity, but not flight performance,  
571 explains variation in dispersal ability. *Proc R Soc B-Biol Sci* 283:20160905.  
572 <https://doi.org/10.1098/rspb.2016.0905>

573 Strahler AN (1952) Hypsometric (area-altitude) analysis of erosional topography. *GSA Bulletin*  
574 63:1117–1142. [https://doi.org/10.1130/0016-7606\(1952\)63\[1117:HAAOET\]2.0.CO;2](https://doi.org/10.1130/0016-7606(1952)63[1117:HAAOET]2.0.CO;2)

575 Svensson BW (1974) Population Movements of Adult Trichoptera at a South Swedish Stream. *Oikos*  
576 25:157–175. <https://doi.org/10.2307/3543638>

577 Sweaney N, Lindenmayer DB, Driscoll DA (2022) Movement across woodland edges suggests  
578 plantations and farmland are barriers to dispersal. *Landsc Ecol* 37:175–189.  
579 <https://doi.org/10.1007/s10980-021-01340-5>

580 Symondson WOC, Cesarini S, Dodd PW, et al (2006) Biodiversity vs. biocontrol: positive and negative  
581 effects of alternative prey on control of slugs by carabid beetles. *Bull Entomol Res* 96:637–  
582 645. <https://doi.org/10.1017/ber2006467>

583 Tachet H, Richoux P, Bournaud M, Usseglio-Polatera P (2000) *Invertébrés d'eau douce: systématique,*  
584 *biologie, écologie.* CNRS Editions, Paris

585 Tonkin JD, Altermatt F, Finn DS, et al (2018) The role of dispersal in river network metacommunities:  
586 Patterns, processes, and pathways. *Freshw Biol* 63:141–163.  
587 <https://doi.org/10.1111/fwb.13037>

588 Uno H (2016) Stream thermal heterogeneity prolongs aquatic-terrestrial subsidy and enhances riparian  
589 spider growth. *Ecology* 97:2547–2553. <https://doi.org/10.1002/ecy.1552>

590 Wesner JS (2010) Seasonal variation in the trophic structure of a spatial prey subsidy linking aquatic  
591 and terrestrial food webs: adult aquatic insects. *Oikos* 119:170–178.  
592 <https://doi.org/10.1111/j.1600-0706.2009.17687.x>

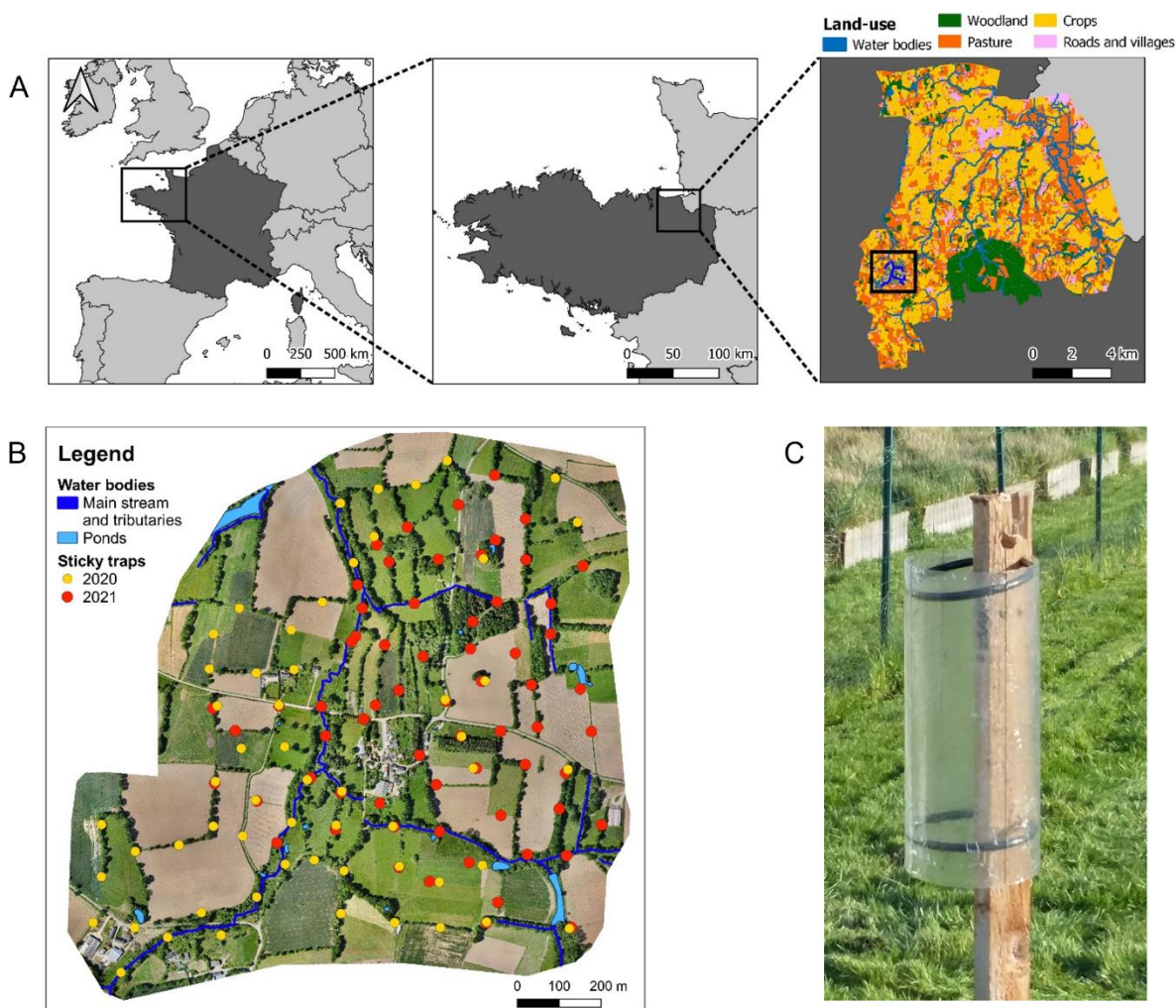
593 Williams DD (1997) Temporary ponds and their invertebrate communities. *Aquat Conserv-Mar*  
594 *Freshw Ecosyst* 7:105–117. [https://doi.org/10.1002/\(SICI\)1099-0755\(199706\)7:2<105::AID-  
595 AQC222>3.0.CO;2-K](https://doi.org/10.1002/(SICI)1099-0755(199706)7:2<105::AID-AQC222>3.0.CO;2-K)

596 Williams P, Whitfield M, Biggs J, et al (2004) Comparative biodiversity of rivers, streams, ditches and  
597 ponds in an agricultural landscape in Southern England. *Biol Conserv* 115:329–341.  
598 [https://doi.org/10.1016/S0006-3207\(03\)00153-8](https://doi.org/10.1016/S0006-3207(03)00153-8)

599 Winterbourn MJ (2005) Dispersal, feeding and parasitism of adult stoneflies (Plecoptera) at a New  
600 Zealand forest stream. *Aquat Insects* 27:155–166.  
601 <https://doi.org/10.1080/01650420500062840>

602

603



605

606 Fig. 1 Sampling site. A: Location of Brittany, the Zone Atelier Armorique (ZAAr) and the study site.

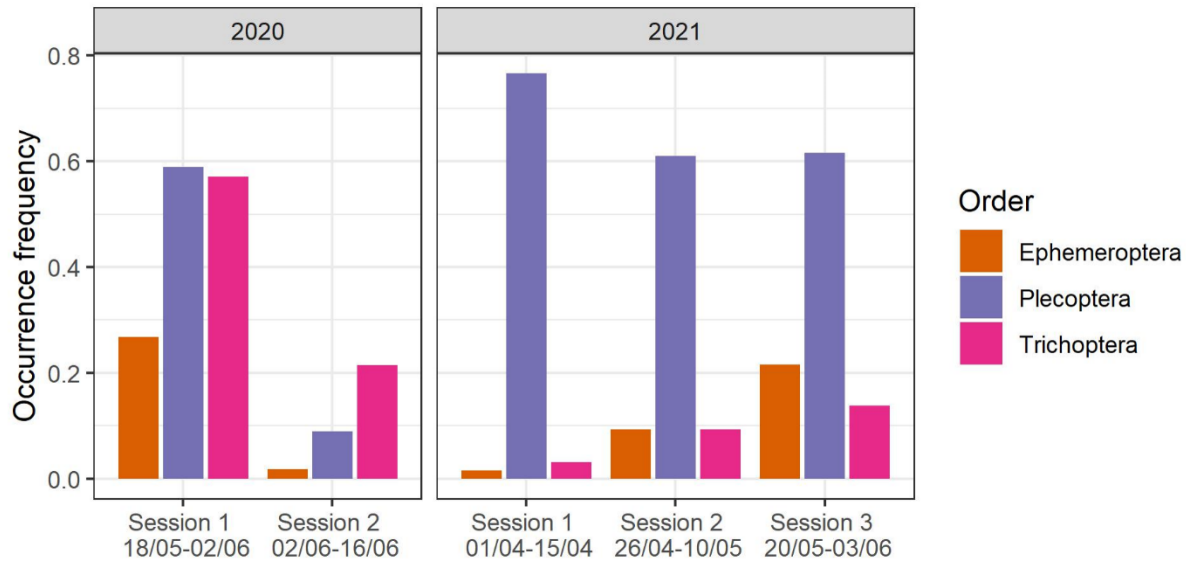
607 B: Sticky trap locations. The colors represent the sampling years (yellow: 2020, 56 traps; red: 2021,

608 64 traps). The elements indicated in blue are the water bodies (dark blue: main stream, Le Guyoult and

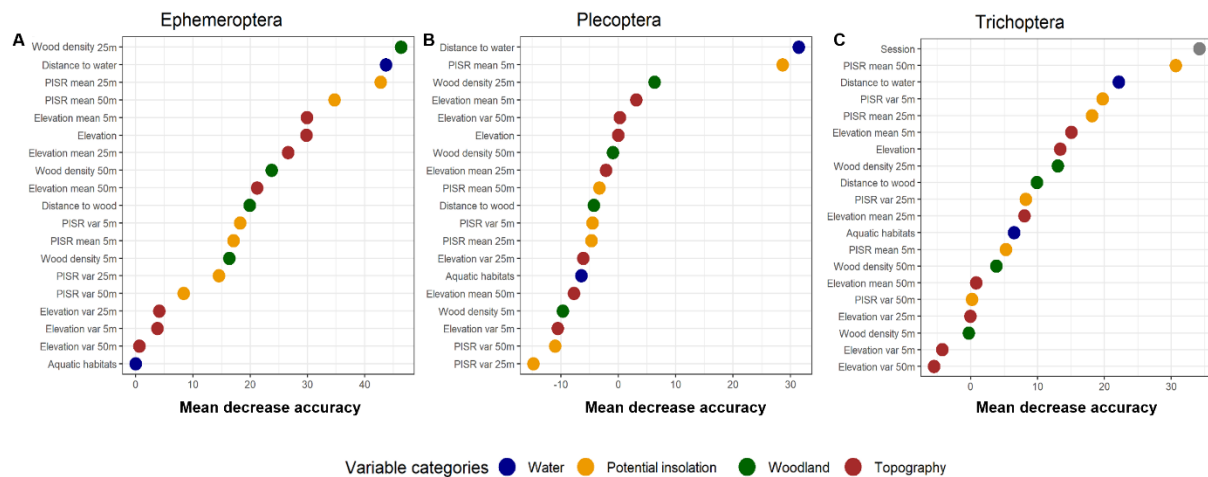
609 its tributaries, light blue: ponds). C: Picture of a sticky trap.

610





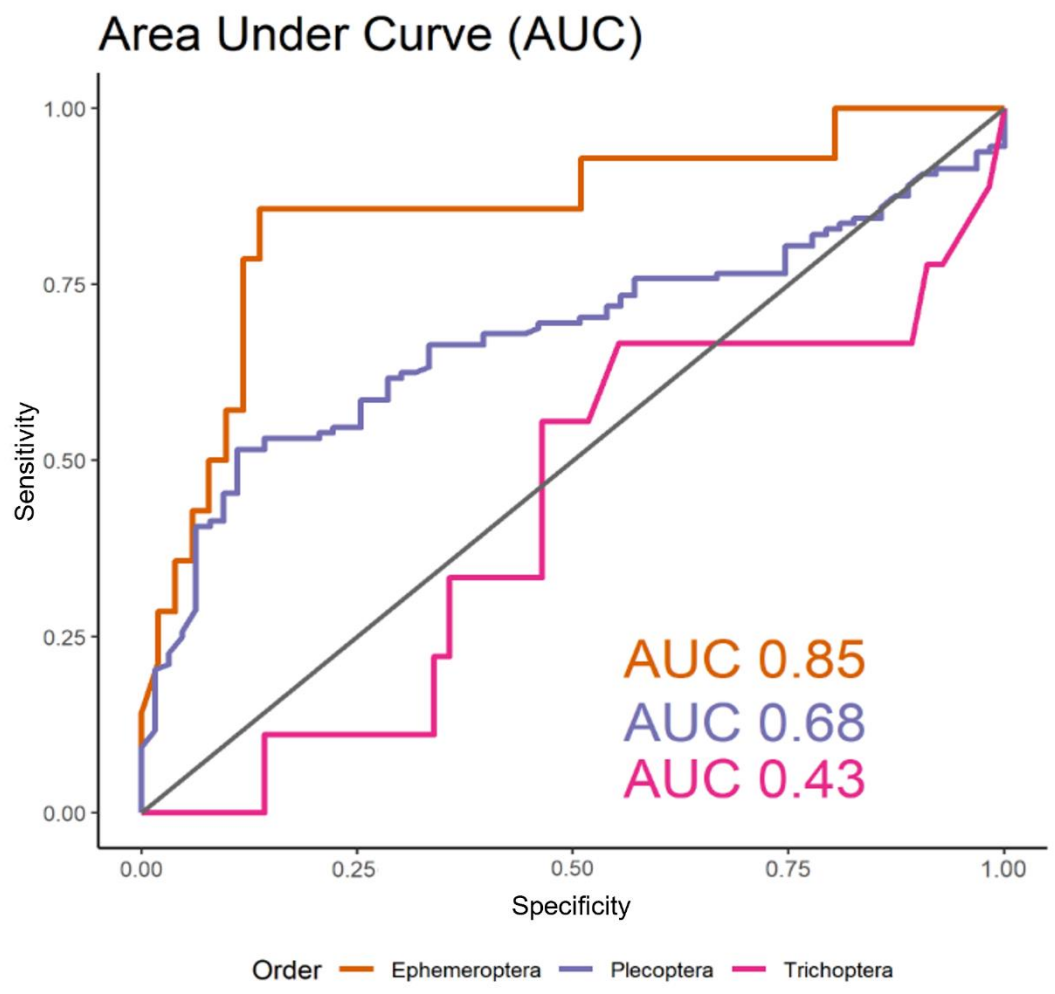
611  
 612 Fig. 2. Occurrence frequency per session and order. The occurrence frequency is the number of sticky  
 613 traps with individuals of the considered order divided by the total number of sticky traps for a session.  
 614



615  
 616 Fig. 3 Plot of variable importance selected from random forest models for Ephemeroptera (A),  
 617 Plecoptera (B) and Trichoptera (C). The mean decrease accuracy is a measure of variable importance  
 618 and expresses how much accuracy the model losses by excluding each variable. It represents the  
 619 average increase in the error of a tree in the forest when random permutations are made in the out-of-

620 bag (OOB) samples. The greater the increase in the error generated by the random permutations of an  
621 explanatory variable, the more important the variable is.

622

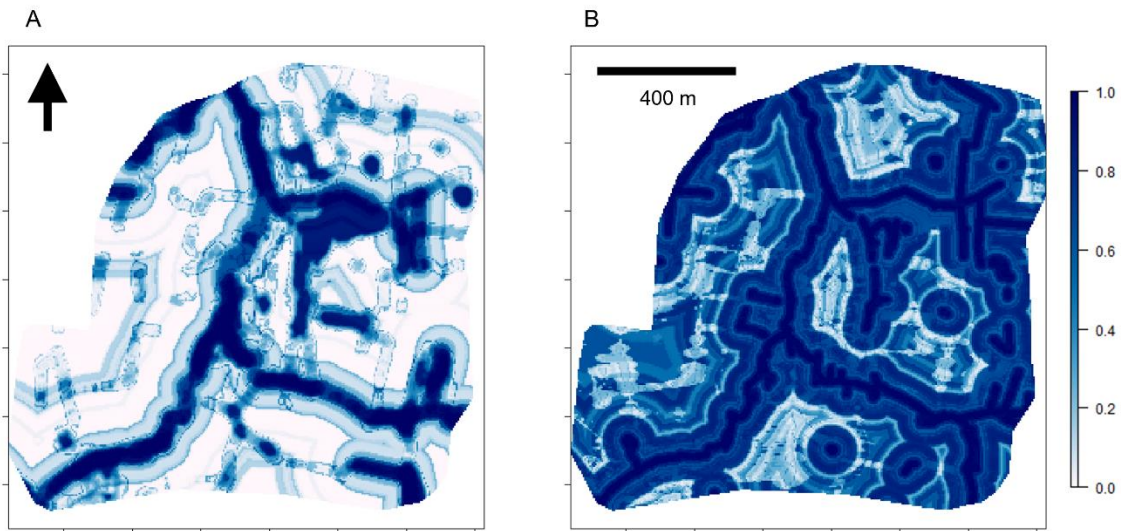


623

Fig. 4 Area

624 under the curve (AUC). For each order, models were calibrated using a random forest algorithm on the  
625 dataset 2020. The AUC was calculated from the predicted model on environmental variables and order  
626 occurrence of 2021. Sensitivity is the proportion of true occurrence. Specificity is the proportion of  
627 true absence.

628



629

630 Fig. 5 Prediction of the probability of occurrence for Ephemeroptera (A) and Plecoptera (B). Dark blue  
631 areas represent high probabilities of occurrence and white areas represents low probabilities of  
632 occurrence.

633

## Supplementary materials

### **S1. Hydrographic networks**

For each year and each order, we selected the hydrographic network that best explained the distribution of the aquatic insects following Gerber et al. (2022a). For the year 2020, the hydrographic networks used were directly extracted from Gerber et al. (2022a). For 2021, fourteen hydrographic networks were considered in the analyses (Table S1). As the hydrographic network usually decreases from spring to summer, we mapped manually all the water bodies at 4 periods (April 12<sup>th</sup>, May 5<sup>th</sup>, May 27<sup>th</sup> and June 29<sup>th</sup>). Wetland and puddle were not considered. For networks mapped manually in the field for each date, we distinguished 4 aquatic habitats (the main stream, the tributaries, the ditches with running/standing water and the ponds, Table S1). We also considered and tested all the hydrographic networks, with and without ponds.

Table S1. Description of the fourteen hydrographic networks based on aquatic habitats.

Network	Source	Description		
1	OpenStreetMap (OSM) hydrographic map	It includes the drainage of the Guyoult stream and two permanent ponds		
2	The complete drainage of the Guyoult stream, obtained by manual mapping (compilation of networks 3, 6 and 9) during the field period (April-June)	It includes the drainage of the Guyoult stream with its intermittent and permanent tributaries (unmapped by OSM), and all the full-water ponds during the sampling period		
		Main stream	Tributaries	Ditches with running/standing water
3		Yes	Yes	Yes
4	April 12 <sup>th</sup>	Yes	Yes	No
5		Yes	No	No
6	May 5 <sup>th</sup>	Yes	Yes	Yes

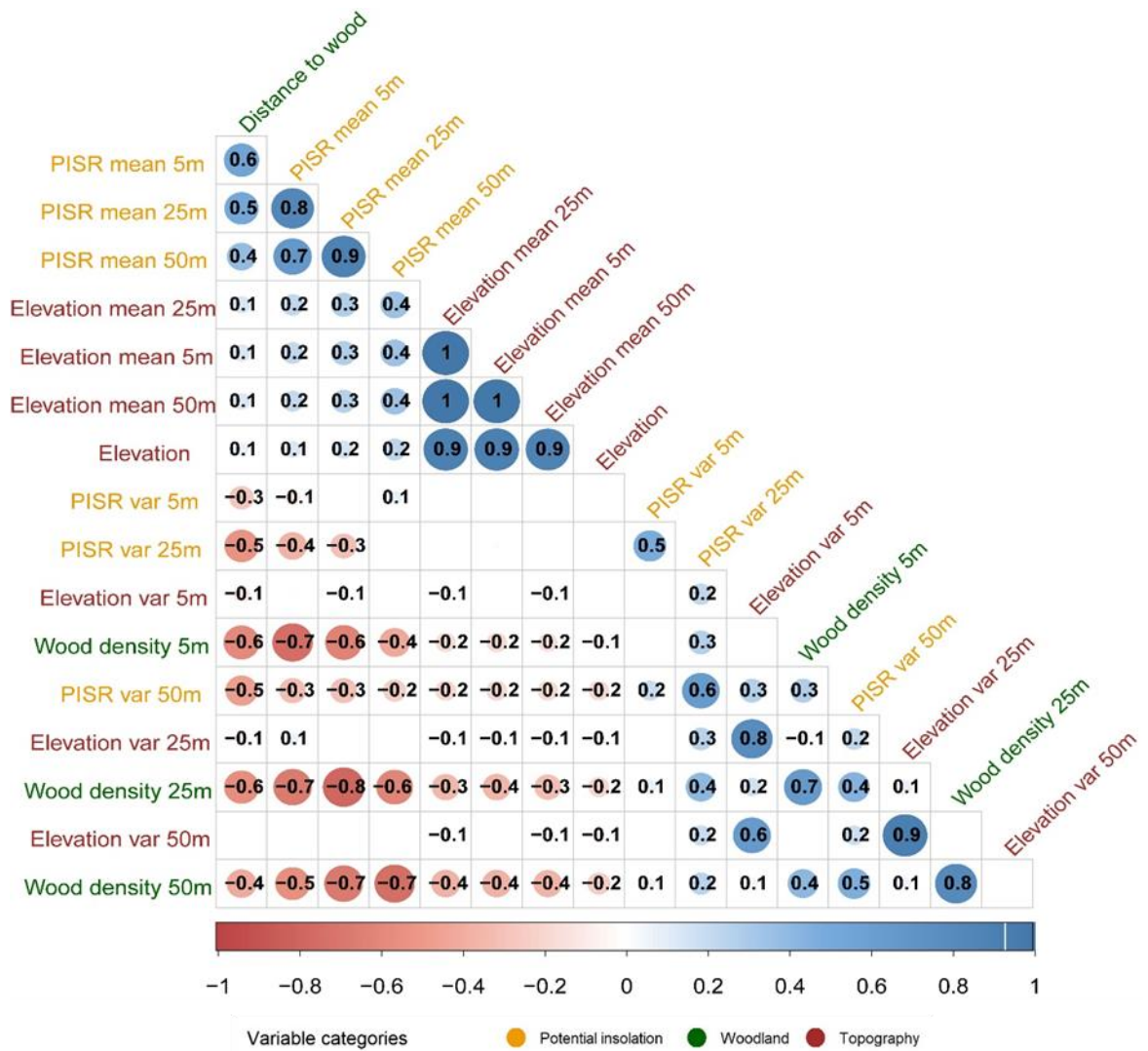
7		Yes	Yes	No
8		Yes	No	No
9		Yes	Yes	Yes
10	May 27 <sup>th</sup>	Yes	Yes	No
11		Yes	No	No
12		Yes	Yes	Yes
13	June 29 <sup>th</sup>	Yes	Yes	No
14		Yes	No	No

Abundance of Plecoptera in 2020 is best explained by the network 1 without ponds, (pseudo  $R^2 = 0.43$ ), and in 2021, by the full network surveyed the 27<sup>th</sup> of May, ponds included (pseudo  $R^2 = 0.22$ ). Abundance of Trichoptera in 2020 is best explained by the network surveyed the 5<sup>th</sup> of May, ponds excluded (pseudo  $R^2 = 0.27$ ) and in 2021 by the network 2, ponds included (pseudo  $R^2 = 0.22$ ). Abundance of Ephemeroptera in 2020 is best explained by the network 3 including the main stream only (pseudo  $R^2 = 0.71$ ), and in 2021 by the network 2, ponds excluded (pseudo  $R^2 = 0.55$ ). For Megaloptera, the best network (3) includes the main stream and the ponds (pseudo  $R^2 = 0.63$ ).

**Table S2. List of the 19 factors derived from the environmental maps.**

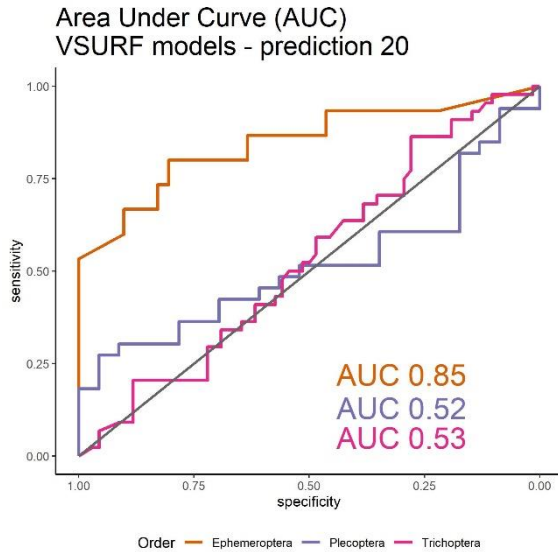
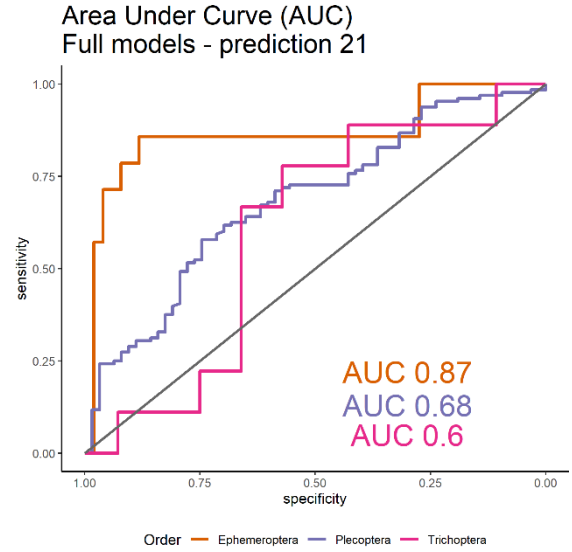
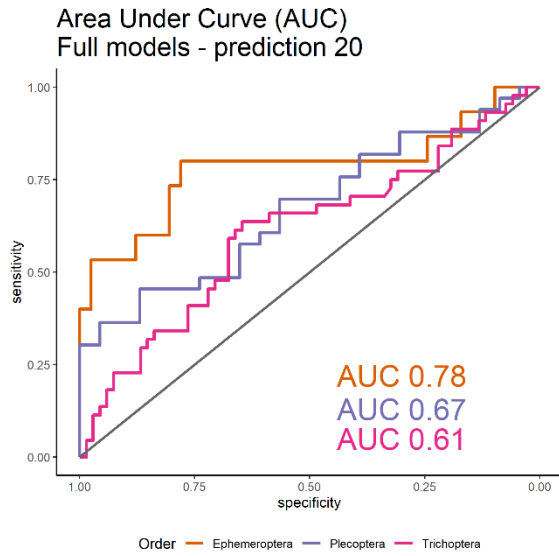
Variables names	Description
Aquatic habitat density 5 m	The number of aquatic habitat pixels divided by the area in a 5 m buffer radius
Aquatic habitat density 25 m	The number of aquatic habitat pixels divided by the area in a 5 m to 25 m radius ring buffer
Aquatic habitat density 50 m	The number of aquatic habitat pixels divided by the area in a 25 m to 50 m radius ring buffer
Distance from the nearest wooded area	Distance from the nearest wooded area (m)
Distance from the nearest aquatic habitat	Distance from the nearest aquatic habitat (m)
Elevation	Elevation (m) of the sticky trap
Elevation mean 5 m	Mean elevation (m) in a 5 m buffer radius
Elevation mean 25 m	Mean elevation (m) in a 5 m to 25 m radius ring buffer
Elevation mean 50 m	Mean elevation (m) in a 25 m to 50 m radius ring buffer
PISR mean 5 m	Potential Incoming Solar Radiation mean (kWh.m <sup>-2</sup> ) in a 5 m buffer radius
PISR mean 25 m	Potential Incoming Solar Radiation mean (kWh.m <sup>-2</sup> ) in a 5 m to 25 m radius ring buffer
PISR mean 50 m	Potential Incoming Solar Radiation mean (kWh.m <sup>-2</sup> ) in a 25 m to 50 m radius ring buffer
PISR var 5 m	Potential Incoming Solar Radiation variance (kWh.m <sup>-2</sup> ) in a 5 m buffer radius
PISR var 25 m	Potential Incoming Solar Radiation variance (kWh.m <sup>-2</sup> ) in a 5 m to 25 m radius ring buffer
PISR var 50 m	Potential Incoming Solar Radiation variance (kWh.m <sup>-2</sup> ) in a 25 m to 50 m radius ring buffer
Session	Sampling session
Wood density 5 m	Number of wooded pixels divided by the area in a 5 m buffer radius
Wood density 25 m	Number of wooded pixels divided by the area in a 5 m to 25 m radius ring buffer
Wood density 50 m	Number of wooded pixels divided by the area in a 25 m to 50 m radius ring buffer





**Fig. S3 Correlation matrix of variables.** Numbers are the Pearson correlation coefficients. The higher the coefficient, the larger the circle. Squares are blank (*i.e.* no circle) when the coefficient is zero. The colors of the legend represent the 3 categories of variables considered, red for topography, green for woodland and yellow for the potential insolation.





**Fig. S4 Area under curve (AUC)** For each order, models were calibrated using random forest algorithm. Then the AUC were calculated from the predicted model on environmental variables and order occurrence. A: AUC obtained with all the variables, calibration on dataset of 2021 and prediction on dataset of 2020. B: AUC obtained with all the variables, calibration on dataset of 2020 and prediction on dataset of 2021 dataset. C: AUC obtained from variables selected with the VSURF procedure, calibration on dataset of 2021 and prediction on dataset of 2020. Sensitivity is the proportion of true presence; specificity is the proportion of true absence.