Joint species distributions reveal crop type and field-specific assembly rules and idiosyncrasies in carabid beetle occurrences and abundances

Abel Masson [1], Yann Tricault [2], Sylvain Poggi [3], Lucile Muneret [4], Sandrine Petit [4], Stéphanie Aviron [5, 6], Aude Vialatte [7], Manuel Plantegenest [1]

[1] IGEPP, INRAE, Institut Agro, Univ Rennes, Rennes, France ;
[2] IGEPP, INRAE, Institut Agro, Univ Rennes, Angers, France ;
[3] IGEPP, INRAE, Institut Agro, Univ Rennes, Le Rheu, France ;
[4] Université de Bourgogne Europe, Institut Agro Dijon, INRAE, UMR Agroécologie, Dijon, France ;
[5] BAGAP, L'Institut-Agro Rennes-Angers, ESA, INRAE, Rennes, France ;
[6] Zone Atelier Armorique, Rennes, France ;
[7] Université de Toulouse, INRAE, DYNAFOR, Castanet-Tolosan, France

Corresponding Author

Abel Masson - Email: abel.masson@inrae.fr

Author Contributions

AM designed the study, performed the analyses and wrote the first draft, MP, YT and SyP designed and supervised the study, and acquired funding, MP, SA, LM, SaP and AV set up and coordinated the SEBIOPAG Network. All authors reviewed the manuscript and gave final approval.

Acknowledgments

The authors would like to thank Audrey Alignier for her contribution to the SEBIOPAG Network.

Conflict of Interest

None declared.

Abstract

1. Designing effective biodiversity conservation and pest regulation strategies in agroecosystems requires understanding how environmental gradients and assembly rules jointly structure ecological communities.

2. We fitted a hurdle joint species distribution model (jSDM) to presence–absence and conditional abundance data of 20 carabid species sampled across 57 arable fields in three French agricultural regions over five years, incorporating climate, landscape, and agricultural management covariates alongside four hierarchical random effects (sample, field, year, crop type).

3. Species distributions were primarily driven by region and temperature. Disturbance and resource gradients had little effect on occurrence, but pesticide pressure consistently reduced the conditional abundance of several species, particularly granivorous ones.

4. Beyond measured gradients, residual co-occurrence patterns revealed distinct species assemblages at both crop type and field levels. A latent crop-type gradient opposed a *Brassicaceae* assemblage — shaped by phenological match, trophic resources and interspecific interactions — to a maize–soybean assemblage. At field level, the idiosyncratic distribution of *Carabus auratus* was interpreted as a signal of extinction debt.

5. **Synthesis and applications.** Carabid community structure is constrained by crop type through biotic and abiotic mechanisms, with direct implications for biocontrol strategies. Short-term pesticide reduction may restore population densities without recovering species diversity, highlighting the need for abundance-based indicators of agri-environment schemes benefits. The detection of an idiosyncratic species distribution, consistent with an extinction debt, further illustrates the relevance of jSDMs for conservation in agricultural landscapes.

Keywords

Community ecology; Assembly rules; Agricultural Landscapes; Carabid beetles; Joint Species Distribution Model

1. Introduction

At the turn of the twentieth century, two theories still competed over the existence and nature of assembly rules governing ecological community structure (the set of species and their abundances): Clements's organismic hypothesis [Clements 1916] — positing that communities tend, through a deterministic ecological succession analogous to an ontogenesis, toward a recurrent discrete “organic entity” specific to each habitat [Kirchhoff 2020] — and Gleason's individualistic hypothesis [Gleason 1926], holding that communities emerge contingently from the response of each species to the environment. While history has largely favored Gleason's individualistic view, it also recognized the influence of species functional roles [Elton 1927] - later on of biotic interactions - that constrain the realized distribution of individuals on the environmental hyperspace [Hutchinson 1957], and which, as do sampling biases or unobserved ecological gradients, introduce interdependencies in species distributions [Ovaskainen & Abrego 2020]. By controlling for individual species-environment relationships to reveal residual co-variation patterns, joint species distribution models (jSDMs [Warton et al. 2015; Tikhonov et al. 2020]) reinvent a century-old debate over the existence and nature of assembly rules in residual community structure.

Integrating biotic interactions into species distribution models is notoriously difficult — because the number of pairwise interactions grows quadratically with species richness [Bystrova et al. 2021], because interaction networks are rarely known and hard to observe directly [Pascual & Dunne 2005], and more fundamentally because the dynamic processes by which species interactions shape co-variation patterns are not always recoverable from static data alone — as illustrated by [Connell 1980] on the *ghost of competition past* or more recently by [Zurell et al. 2018] with transient predator-prey dynamics. Within this inherent limitation, a productive approach is that of redundancy analysis (RDA [Legendre & Legendre 1998]), which has been widely used in community ecology for nearly four decades (e.g. see [Eyre et al. 2013] for an example on carabid communities) and consists in separating a constrained component of community structure — explained by environmental gradients — from an unconstrained residual structure that absorbs the joint effects of unobserved drivers and of biotic interactions strong and persistent enough to leave a detectable imprint on species distributions [Poggiato et al. 2021]. Joint species distribution models such as HMSC [Tikhonov et al. 2020] build on the same fundamental separation, with more flexible regression approaches for the constrained part, and a more parsimonious approach for the unconstrained residual part, based on a shared "interaction currency" [Kissling et al. 2012] in the form of a small set of latent factors.

Carabid beetles (*Coleoptera: Carabidae*) are among the most diverse, abundant and well-documented arthropod families in temperate agroecosystems [Thiele 1977; Lövei & Sunderland 1996; Kromp 1999; Holland 2002], with species distributed across a wide range of environmental gradients [Rainio & Niemelä 2003], offering a well-suited system for studying community structure. Their life cycle typically includes a peak of activity in spring or autumn, during which adults disperse across the dynamic landscape mosaic according to its properties (e.g. composition, fragmentation, connectivity) and their dispersal capacity, in search for food, oviposition, or overwintering sites. Habitat preferences vary among species and according to their traits (e.g. diet, or nychthemeral rhythm), between crop and non-crops habitats (e.g. woody, grassy margins or flower strips) [Knapp et al. 2022, Boetzl et al. 2024, Bannwart et al.

2025], and within crop fields according to soil structure [Holland 2002], microclimate (e.g. temperature, hygrometry), resource availability [Carbonne et al. 2022], and management practices [Holland and Luff 2000] such as pesticide applications [Geiger et al. 2010, Eyre et al. 2013] and tillage [Müller et al. 2022,]. Crop type plays a pivotal role in shaping carabid communities in agricultural landscapes [Eyre et al. 2013, Boetzl et al. 2024], for instance through phenology - governing the provision of resources during species peak activity [Duflot et al. 2016] - or through canopy cover, which buffer thermal stress for small dark-bodied species that are less tolerant of heat [Kegel 1990, Holland 2002]. Beyond abiotic drivers, crop type also recruits specific prey and natural enemies [De Heij et Willenborg 2020], with whom carabids are involved in trophic interactions, as generalist predators of crop pests (e.g. aphids, slugs and weed seeds [Sacco-Martret de Préville et al. 2022; Bohan et al. 2011]) and other natural enemies [Davey et al. 2013], or as parasites [Saska et al. 2008], but also non-trophic interactions, such as facilitation [Losey & Denno 1998] or aposematic association [Brandmayr et al. 2010].

This study expanded upon [Muneret et al. 2023], who used generalized linear models to explain the occurrence of individual carabid species based on region, climate, landscape, and agricultural resource and disturbance gradients. Beyond individual species responses to environmental covariates, our study probed the often promoted albeit rarely quantified contribution of additional assembly rules to community structure. Defined as any biotic or abiotic process altering species coexistence [Götzenberger et al. 2012], assembly rules were sought in residual co-occurrence and abundance covariation patterns allowed to emerge at local (sample, field), non-spatial (crop type) and temporal (year) levels. Such patterns may reflect, for instance, shared sensitivity to an unmeasured gradient at field level, consistent biotic interactions within crop types, or species succession across years. Specifically, we applied a joint species distribution model to predict carabid community structure as a function of eight environmental covariates and residual co-occurrence and abundance covariation at four random levels — sample, field, year, and crop type. The model was calibrated with data from 20 carabid species sampled twice a year over five years across 57 arable plots in three agricultural regions of France. Both assembly rules and their absence (idiosyncrasies) may be relevant to biodiversity conservation, or to improve predictions of community-level responses to global change, or to support pest regulation strategies by carabid beetle communities.

2. Material & Method

Using the same data and set of covariates as [Muneret et al. 2023], we extended their approach to a joint species distribution model (jSDM) implemented as a hurdle model [Martin et al. 2005] capturing species variation and covariation in both occurrence and conditional abundance (abundance given presence). The hierarchical structure of our model additionally allowed functional traits to mediate species responses to environmental covariates. All model definition and calibration were performed within the Hierarchical Modelling of Species Communities (HMSC) Bayesian framework [Ovaskainen et al. 2017].

2.1. Study design

The study was conducted within the SEBIOPAG network [Figure 1], comprising 57 arable fields across three agricultural regions of France (near Rennes, Dijon, and Toulouse) sampled twice a year from 2014 to 2018 (except Brittany in 2017) using one-week pitfall trapping sessions [Petit et al., 2021]. Sampling dates were determined using degree-day accumulation (1000 and 1500 DD) to ensure temporal comparability across years and regions. Crop type was included as a hierarchical grouping factor (21 categories). Samples (field × year × session) with unknown or rare crop type (fewer than 4 occurrences, approximately 10% of samples) were excluded from the analyses.

2.2. *Community surveys*

Collected carabid beetles were identified to the species level, except for a small number of individuals that could not be reliably assigned to species. Rare species (< 0.5% of total activity-density) and species for which trait information was unavailable (e.g. *Anisodactylus* spp.) were excluded, resulting in a final dataset of 20 species across 458 samples.

2.3. Environmental covariates

Three of the nine covariates, relating to broad environmental context — region, sampling session (at 1000 and 1500 cumulative degree-days), and arable land cover (proportion of cultivated area within a 1 km² circle buffer around focal fields) — were directly measured. The remaining six were derived from daily time series, relating to climate (mean daily precipitation and temperature), resource availability (fertilization intensity and simulated daily crop height), and disturbance (physical, based on expert-assigned scores for soil-disturbing operations, and chemical, based on the Treatment Frequency Index). For each, the cumulative sum of daily values over the year preceding each survey was the covariate entering the model. Following [Muneret et al. 2023], we also explored a seasonal alternative, summarizing seasonal cumulations onto the first axis of a PCA [Supp. Mat. 0]. As results were highly consistent between the two approaches, we present hereafter only the annually cumulated variables, which are more readily interpretable.

2.4. Species traits

Species functional traits were obtained from the databases carabid.org and BETSI [Joimel et al. 2021, https://portail.betsi.cnrs.fr/]. Traits included mean body size, as a species-level average, and breeding season (autumn or spring). Species without a clearly defined breeding season (e.g. *P. cupreus*) were coded as neither autumn nor spring breeders.

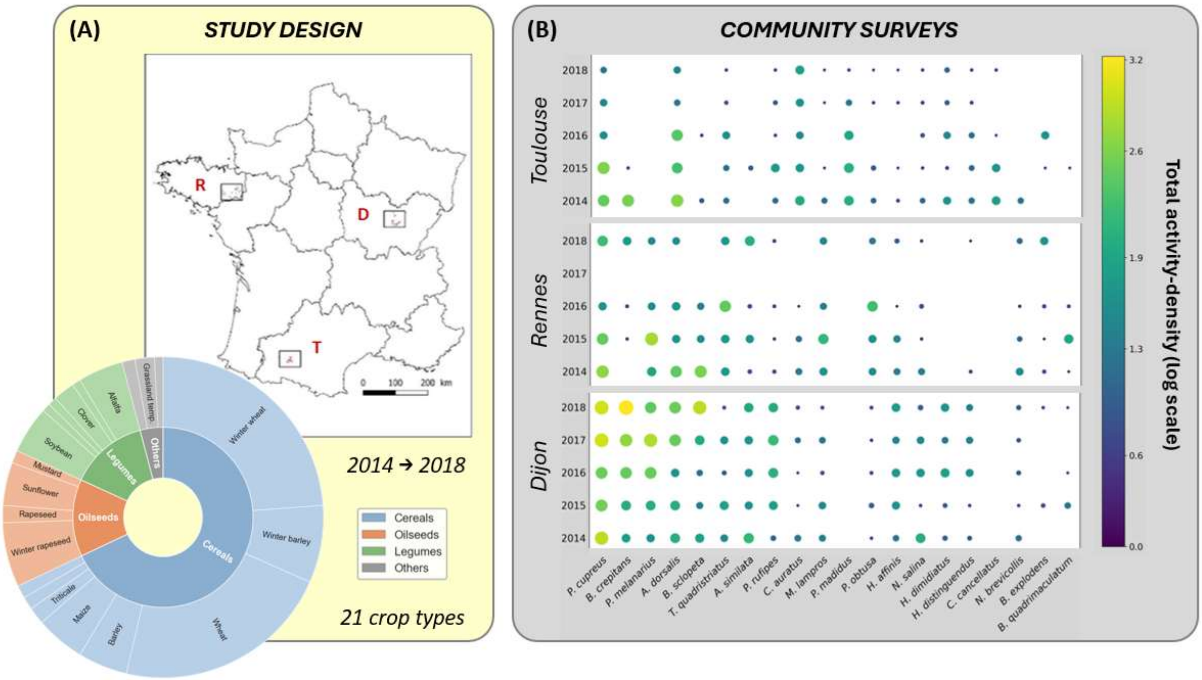


Figure 1. Overview of the experimental design and distribution of the 20 carabid species considered, stratified by region and year. (A) Study design showing the 57 monitored arable fields (red dots) across three agricultural regions of France (black rectangles): Rennes (R), Dijon (D), and Toulouse (T). Fields were sampled twice a year from 2014 to 2018 (except R in 2017), and crop type was recorded each year (21 crop types). (B) Total activity-density of each carabid species (x-axis) summed within each region × year combination displayed on a logarithmic scale (right-hand axis). See [Supp. Mat. 1] for more details.

2.5. *Environmental covariates selection*

To allow for nonlinearity in responses to covariates, we included linear and quadratic terms for all continuous predictors [Muneret et al. 2023; see Fig. 3]. We validated our covariate selection using a Poisson lognormal regression [PLN; Chiquet et al. 2021] as a computationally efficient surrogate for HMSC model. The full covariate PLN model achieved the lowest Bayesian Information Criterion (BIC) among the candidate models [see Table SM2.1 in Supp. Mat. 2].

2.6. *Specification of random effects*

Beyond accounting for hierarchical dependence among observations, random effects in HMSC also structure the multivariate covariance of species residuals, enabling the detection of scale-specific species associations. We defined four random effects — crop type, site, year, and sample —representing scales at which residual variation may arise from unmeasured drivers [see Table SM 3.1 in Supp. Mat. 3 for examples of biotic and abiotic drivers of species associations operating at each scale]. Climate covariates, measured at the year × region level, largely captured inter-annual variation once region is controlled for, limiting residual variance for the year random effect. It was nonetheless retained, as species co-occurrence or abundance covariation patterns may still arise at an annual level.

2.7. *Hierarchical joint species distribution model (jSDM)*

Several species were absent from one or more study regions due to their biogeographic range — for instance, *Carabus cancellatus* only recorded in Toulouse, or *Pterostichus madidus* and *P. melanarius* showing a geographic turnover between southern and northern regions [Avtaeva et al. 2021] — introducing structural zeros, i.e. absences that reflect biogeography rather than unfavorable local conditions. To account for this, we modelled presence–absence (PA) and conditional abundance (AB) – i.e. abundance given presence – as separate components within a hurdle jSDM framework [Ovaskainen et al. 2017, 2020]. The same set of environmental predictors, traits, and random effects were used to model occurrence and abundance.

The first component describes species occurrence (presence–absence, $\boldsymbol{PA}$) using a probit link:

$$\boldsymbol{PA} \sim Bernoulli(\boldsymbol{p}) \ ; \ probit(\boldsymbol{p}) \ = \ \boldsymbol{\eta_{PA}} \quad \textbf{[Eq. 1]}$$

The second component describes conditional abundance ($\boldsymbol{Y}$) using a normal distribution on log-transformed counts, with an offset accounting for sampling effort (number of pitfall traps, N_{PT}).

$$log(\boldsymbol{Y}) \mid \boldsymbol{Y} > 0 \sim Normal(\boldsymbol{\mu}, \boldsymbol{\sigma}^2) \ ; \ \boldsymbol{\mu} \ = \ \boldsymbol{\eta_{ab}} + log(\boldsymbol{N_{PT}}) \quad \textbf{[Eq. 1bis]}$$

In both components (Eq. 1 and 1 bis), the linear predictor $\boldsymbol{\eta}$ is the sum of the products of the covariates $\boldsymbol{X}$ by the species-specific responses $\boldsymbol{\beta}$, and a residual term capturing unexplained covariation in species PA or abundances. This residual term is structured by latent variables $\boldsymbol{Z}$, defined for each random-effect level – $g \in \{sample, site, crop\ and\ year\}$ – and combined with species-specific factor loadings $\boldsymbol{\Lambda}$ allowing residual variation to be shared among species.

$$\boldsymbol{\eta} \ = \ \boldsymbol{1}\, \boldsymbol{\beta}_0^T + \boldsymbol{X}\, \boldsymbol{\beta} \ + \ \boldsymbol{\Sigma}_{g \in G} \boldsymbol{Z}_g \boldsymbol{\Lambda}_g \quad \textbf{[Eq. 2]}$$

Species traits are incorporated both to inform the model (by sharing information across species) and to quantify their effects on species responses to environmental gradients. The species-specific coefficients $\boldsymbol{\beta}$ (and the intercept, $\boldsymbol{\beta_0}$) are modelled as a function of species traits $\boldsymbol{T}$ (mean body size and breeding season) through hierarchical parameters, $\boldsymbol{\gamma_0}$ and $\boldsymbol{\gamma}$ as follows:

$$\boldsymbol{\beta} \ = \ \boldsymbol{1}\ \boldsymbol{\gamma}_0^T \ + \ \boldsymbol{T}\ \boldsymbol{\gamma} \ + \ \varepsilon \quad \textbf{[Eq. 3]}$$

The covariate matrix $\boldsymbol{X}$ was scaled prior to model fitting, allowing the use of default priors as recommended with *HMSC* [Tikhonov et al., 2020]. Further details on model specification are given in [Supp. Mat. 4]. All model definition and fitting were performed using the HMSC R package v 3.3.6 [Tikhonov et al., 2020].

2.8. *Goodness of fit*

Model fitting relied on Markov chain simulation of posterior parameter distributions, with convergence assessed via the Gelman-Rubin index $\hat{R}$ [Gelman and Rubin 1992]. The explanatory fit of the hurdle model was assessed separately for each component using metrics appropriate to their respective response distributions: Tjur's coefficient of discrimination (Tjur R^2; [Tjur 2009]) for the binary presence-absence component, and the coefficient of determination (R^2) for the log-normal conditional abundance component.

3. Results

3.1. Model fitting

We retained 1,000 posterior samples for all models, following a thinning interval of 100 and a burn-in of 50,000. MCMC convergence was satisfactory: $\hat{R}$ [Gelman and Rubin 1992] exceeded 1.05 for only 1 out of 2032 parameters in the PA component, and 14 out of 2032 in the conditional abundance component, all confined to residual species covariance parameters ($\boldsymbol{\Omega}$ [see Supp. Mat. 3]), with no convergence failure among $\boldsymbol{\beta}$ or $\boldsymbol{\gamma}$ parameters. In the conditional abundance component, these marginal convergence issues were predominantly concentrated at the sample random effect level (Fig. SM 4.1 & 4.2).

The model showed good explanatory fit across both components for most species [Fig. SM 3.B.1]. For the presence-absence component, Tjur R^2 ranged from 0.14 (*Brachinus explodens*) to 0.60 (*C. cancellatus*), with a mean of 0.38, and for the conditional abundance component, R^2 ranged from 0.34 (*Carabus auratus*) to 0.99 (*B. explodens*), with a mean of 0.66. High R^2 values should nonetheless be interpreted with caution, as they predominantly concerned species with few records —notably *B. explodens* (n = 5 presences), for which overfitting is possible.

3.2. Variance partitioning

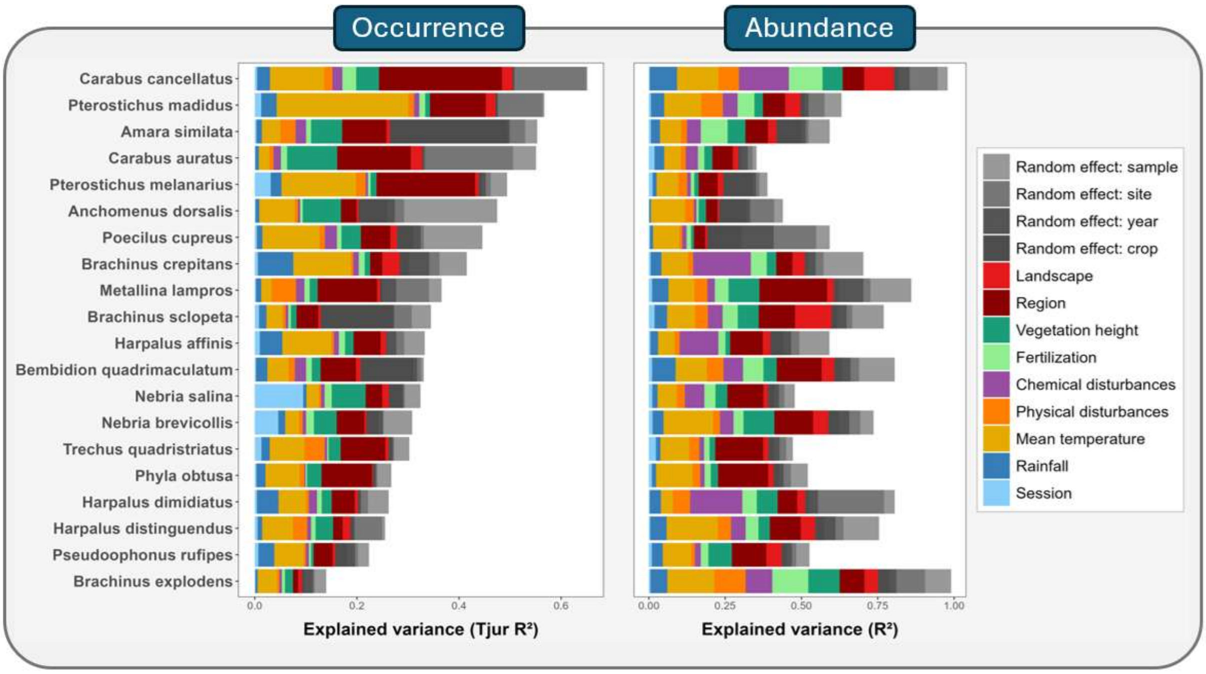


Figure 2. – Variance explained by environmental covariates and random effects, for the occurrence (left) and conditional abundance (right) components of the hurdle model. Bar length corresponds to the Tjur R^2 (occurrence component) and R^2 (conditional abundance component) for each species with colored segments representing the contributions of each fixed effect (climate, landscape, resources, disturbances) and random effect (crop type, site, year, sample). Species are ranked in decreasing order of Tjur R^2. Details are provided in [Supp. Mat. 3].

Variance partitioning revealed contrasting patterns of environmental control between the two model components (Figure 2). With the exception of **vegetation height**, which drove notable species variation in *A. similata*, *C. auratus*, *A. dorsalis*, and both *Nebria* species, and session in *Nebria* species and *P. melanarius,* the occurrence component was dominated by **region** and **mean temperature** (mean contributions of 0.08 and 0.07, respectively), all other covariates contributing modestly across species. Random effects accounted for a substantial but variable share of explained occurrence variance across species, ranging from nearly 50% for *A. similata*, *A. dorsalis* and *Brachinus* species to near zero for *P. melanarius*, *T. quadristriatus* and *P. obtusa*. Among random effects, crop type strongly influenced occurrence in *A. similata*, *B. sclopeta* and, to a lesser extent, *B. quadrimaculatum*, while field-level effects were prominent for both *Carabus* species and *P. madidus*. Year-level effects had negligible influence across all species, and the remaining random variance was carried by the sample level, notably for *A. dorsalis* and *P. cupreus*.

The contributions of fixed and random effects were more evenly distributed across gradients and species for the conditional abundance than for the occurrence component. **Region** and **mean temperature** remained the most significant fixed effects (mean contributions of 0.10 each), but **chemical disturbances** emerged as an important driver (mean contribution of 0.06), with marked effects on *C. cancellatus*, *B. crepitans*, *H. affinis* and *H. dimidiatus*. Random effects explained a comparable conditional abundance variance to that observed for occurrence, though with a more homogeneous distribution across hierarchical levels.

3.3. Species responses to environmental gradients

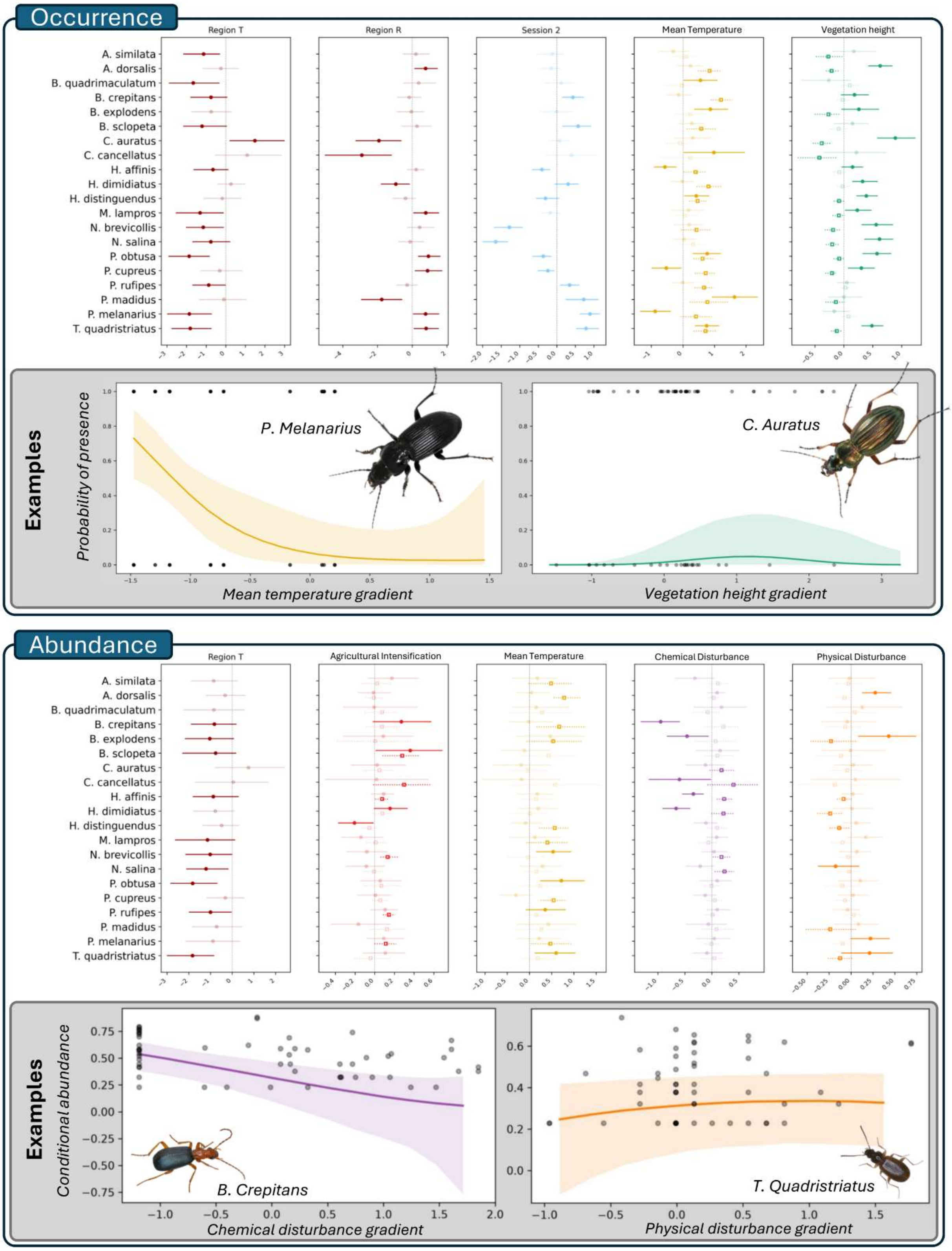


Figure 3. Posterior distributions of species responses to environmental covariates for **the two components of the Hurdle model** — occurrence (upper panel) and conditional abundance (lower panel). On each panel the upper part shows posterior credibility interval at

90% of the species-specific regression coefficients (β_{jk}, as in [Eq. SM3]) for a selection of covariates. Linear and quadratic effects are **displayed** on the same row: (filled circles = linear effects; open squares = quadratic effects). Points and squares indicate posterior medians. Faded responses denote effects with uncertain direction. Lower parts present examples of median marginal responses for two selected species along two continuous gradients. The solid line represents median marginal response, shaded bands the 90% credible interval, and grey points the observations. Complete results and details are provided in [Supp. Mat. 3].

Region was the main driver of species occurrence, with marked effects for *Carabus* species and *P. madidus*, and **Session** reflected clear phenological differentiation, with species falling into three groups: late-season species including *Trechus* and *Pterostichus* spp. (positive effect of session 2), early-season species including *Nebria* spp. (negative effect of session 2), and seasonally indifferent species. **Temperature** increased the probability of occurrence across species, except for *P. melanarius*, whose response to the temperature gradient was inverse of that of *P. madidus*. **Vegetation height** stood out as the only resource gradient influencing occurrence with strong and consistent unimodal responses across species as illustrated by *C. auratus* in [Figure 3]. No agricultural gradient — **landscape agricultural intensification, chemical disturbance**, or **physical disturbance** — had any detectable effect on species occurrence.

Although their influence on conditional abundance was more limited than on occurrence [Figure 3], **region** and **temperature** retain detectable effects: species abundances increased with mean temperature and were lower in Toulouse. Overall, landscape composition and agricultural disturbance gradients had a more pronounced influence on conditional abundance than on occurrence [see Supp. Mat. 3]. Responses to **physical soil disturbance** were heterogeneous, with unimodal responses peaking at intermediate disturbance levels in *B. explodens* and *T. quadristriatus*, and consistently negative responses in the three *Harpalus* species and *P. madidus*. The **proportion of annual cropland in the landscape**, had a strong positive effect on *B. sclopeta*, and a low positive effect on nine other species. Most species were indifferent to **pesticide pressure**, but when detected, its effect was negative, namely for *B. crepitans*, *B. explodens*, *C. cancellatus*, *H. affinis*, and *H. dimidiatus*, with *B. crepitans* showing the steepest declines along the **chemical disturbance** gradient [Figure 3], lower panels).

Finally, the overall scarcity of significant trait–environment interactions [Supp. Mat. 5] suggests that species responses to environmental gradients are only weakly structured by the considered functional traits (mean size and breeding season). Nevertheless, incorporating trait structure improved MCMC convergence and reduced computation time relative to the trait-free model.

3.4. Species Associations

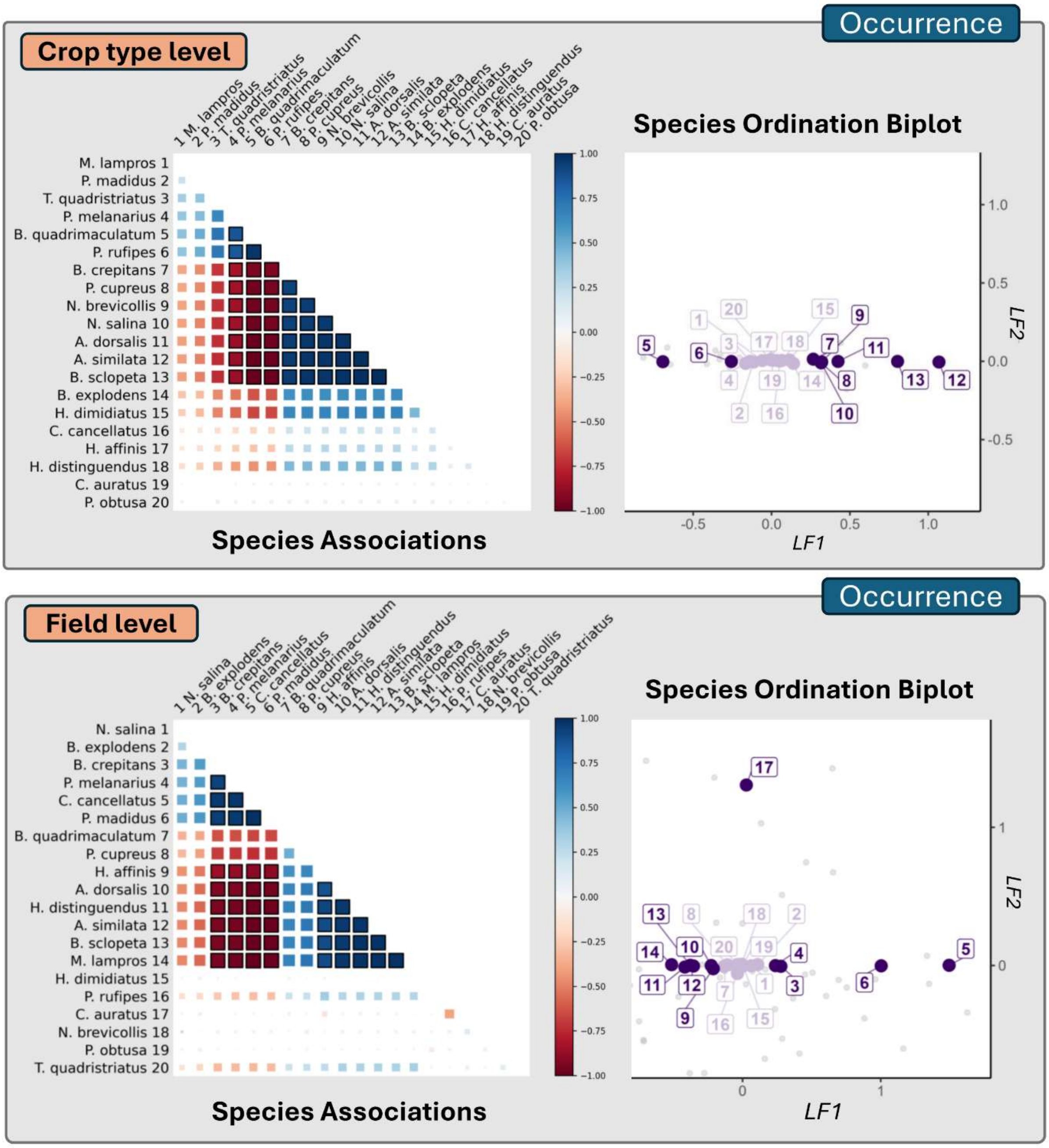


Figure 4a. Species associations for the occurrence component, estimated **at** the **crop type** (upper panels) **and field** (lower panels) random effect **levels**. Left panels: Pairwise residual correlation matrices between species. Color intensity indicates the strength of the residual correlation (blue: positive, red: negative); significant associations are outlined in black. Right panels: ordination biplots showing the projection of species onto the two latent factors, with grey points representing individual sampling units (either crop types or fields), violet points and labels representing species. Species for which at least one factor loading is significantly different from zero are in bold. Species associations at year and sample scale are provided in [Fig. SM 6.1 & 6.2].

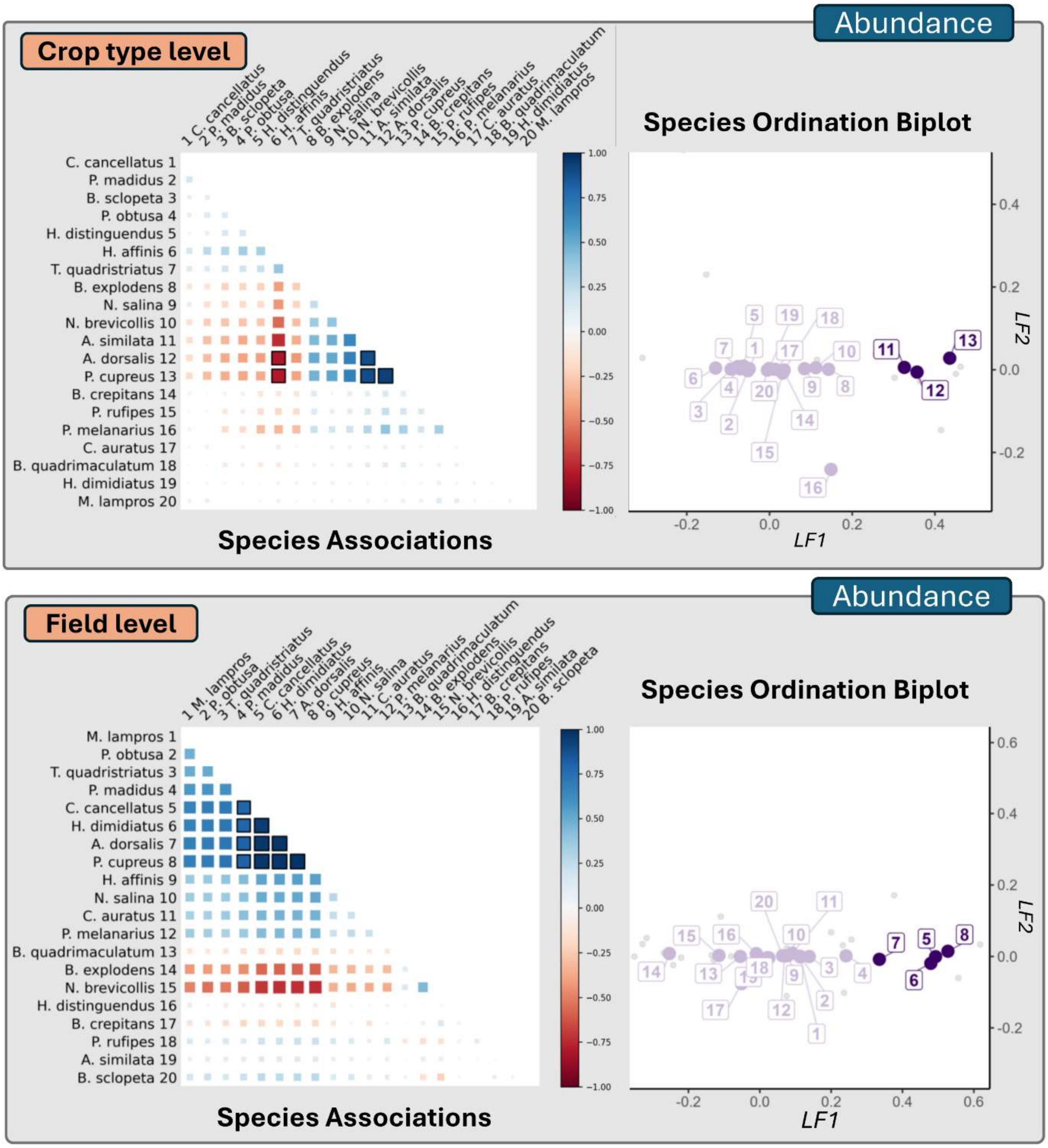


Figure 4.b: Residual species associations for the Conditional Abundance (AB) component, estimated at the crop type (upper panels) **and field** (lower panels) **random effect levels.** Layout and color conventions are as in Figure 4a.

At each level, the 190 possible pairwise species associations represent correlations (positive or negative) in residual occurrence or conditional abundance among species [Fig. SM 6.1 & 6.2]. At crop level, 45 associations in occurrence and 5 associations in conditional abundance were significant (nonrandom), and at field level respectively 45 and 10; these are detailed in [Figure 4 a-b] and below. At year level, 34 associations in occurrence were significant but remained negligible in effect size [Figure 2] and are not further discussed. At sample level, all associations in occurrence were significant (except those involving *C. cancellatus, P. madidus,* and *B. quadrimaculatum*) and positive, suggesting an effect on overall activity-density (e.g. trapping efficiency).

First, for the occurrence component at **crop type level** (Fig. 4a, upper panel), the species ordination biplot (right panel) was structured by a single latent factor (LF1), opposing two assemblages — (4–6) and (7–13) — while a third group — (1–3 & 14–20) — showed no significant associations; LF2 contributed little to the overall pattern. At one end of LF1, *A. similata* and *B. sclopeta* (12 and 13) showed the strongest association, and were positively associated with *P. cupreus*, *A. dorsalis*, the two *Nebria* species, and *B. crepitans* (species 7–11). At the other end, *P. rufipes*, *B. quadrimaculatum* and *P. melanarius* — the latter showing no significant response to LF1 — were negatively associated with all species from group (7–13). Second, **at field level** (Fig. 4a, lower panel), the same broad single-axis structure was recovered — assemblages (3–6) and (9–14) on opposing sides of LF1, and group (1–2, 15–20) bearing no significant associations — with a notable addition: *C. auratus* (17) responded idiosyncratically to LF2, reflecting a field-level response largely independent of the rest of the community. While the opposition between *P. madidus* and *C. cancellatus* (5–6) and assemblage (9–14) appeared to drive LF1, two relations absent at the crop type level emerged at the field level: *B. sclopeta* and *B. crepitans* positively associated at crop type level but negatively at field scale and a positive association between *Harpalus* spp. and *Amara* spp.

The structure observed for the occurrence component was not observed for the conditional abundance component [Figure 4.b, and Fig. SM 6.2]: species associations were scarce and of lesser influence. At the **crop type level**, only five associations were significant: a positive association among *A. dorsalis*, *A. similata* and *P. cupreus* — but not *B. sclopeta*, which was strongly associated with *A. similata* in the occurrence component — alongside *P. melanarius* as the sole driver of LF2. At the **field level**, an assemblage of four species — *H. dimidiatus*, *C. cancellatus*, *A. dorsalis* and *P. cupreus* — emerged.

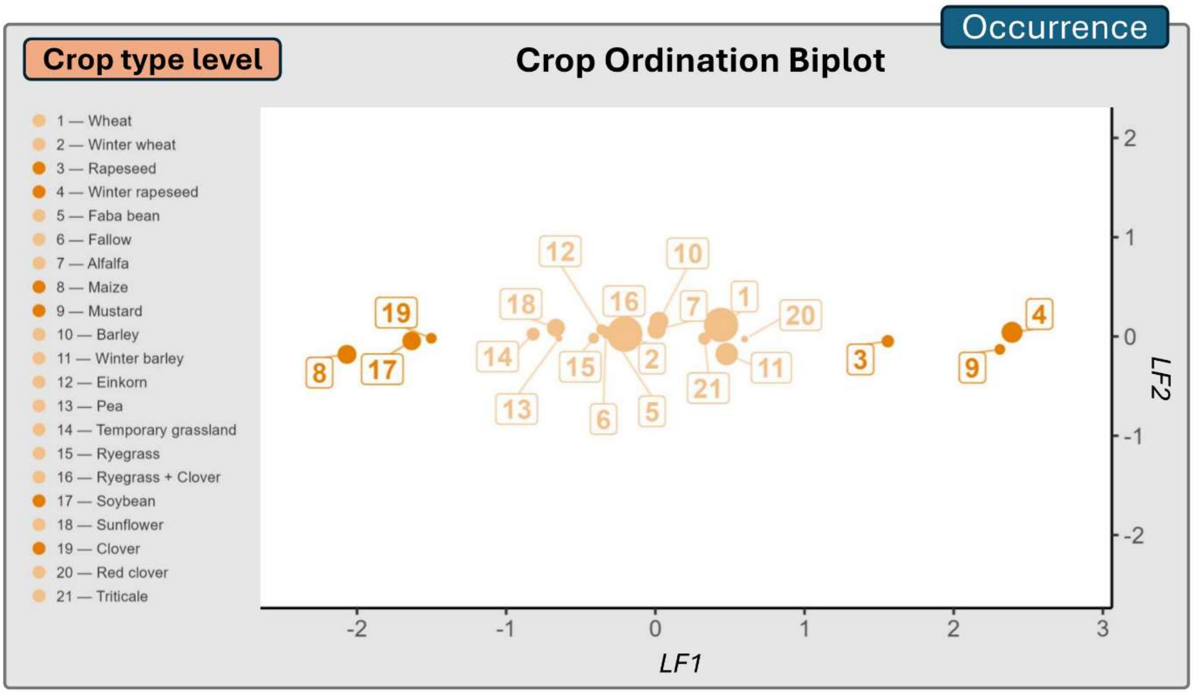


Figure 5. Crop type ordination biplot spanned by the two latent factors (LF1 and LF2) of the occurrence component defined at crop-type level. The 21 crop types are projected into this space, with coordinates corresponding to the latent factor values assigned during model calibration. Highlighted points (bright orange) indicate crop types whose latent factor values are significantly different from zero. Point size reflects the number of observations per crop type (ranging from 105 for winter wheat [2] to 4 for pea [13] [Supp. Mat. 1]).

Crop type ordination for the occurrence component spanned an amplitude approximately twice that of the species ordination, yet only six crop types were assigned latent factors significantly different from zero. Maize, soybean and clover anchored the negative pole, oilseed rape (including winter oilseed rape) and mustard the positive pole. Wheat, barley, winter wheat and winter barley, which together accounted for two thirds of observations in the SEBIOPAG dataset and harbor all carabid species represented therein, clustered near zero [see Figure 5].

4. Discussion

We designed a hurdle joint species distribution model within the HMSC framework, which we successfully fitted to the presence–absence and conditional abundance of 20 carabid beetle species over five years across 57 arable fields. Region and mean temperature emerged as the dominant drivers of both occurrence and conditional abundance. Resource (crop height and fertilization) and disturbance (pesticide use and soil perturbation) gradients had little effect on occurrence, but exerted a stronger influence on abundance, with pesticide pressure showing a clear negative effect on several species. Unobserved ecological gradients, of biotic or abiotic origin, captured by the latent factor structure of our model, accounted for a substantial share of the explained variance — up to 50% for occurrence and to a lesser extent for conditional abundance. These gradients were particularly pronounced at the crop type and field levels, occasionally revealing singular behaviors — such as the response of *C. auratus* at the field level — or crop type-specific species assemblages. Functional traits (mean size and breeding season) did little to explain species–environment relationships.

4.1. Response to covariates

By extending the results obtained with SDMs on the occurrences of six species to the joint occurrences of 20 species, we recovered the main findings of [Muneret et al. 2023] — which can be summarized as a strong effect of region, climate (primarily temperature), and to a lesser extent of vegetation height. We also recovered modest and mixed effects of disturbances (chemical and physical), and a very weak effect of landscape and fertilization.

4.1.1. Broad scale factors – region, session, landscape

Region was the first dominant factor for both occurrence and abundance, reflecting the biogeography of species, contrasting more northerly species (e.g. *P. melanarius* [Avtaeva et al. 2021], never found in Toulouse) with more southerly ones (e.g. *P. madidus*, *C. cancellatus*, found exclusively in Toulouse). This contrast is possibly linked to lower dispersal capacity among larger, flightless carabids (apterous or brachypterous) [Lindroth 1985, Homburg et al. 2013, Joimel et al. 2021], which may impede the dispersal flows between regions. **Session** effect -which was not included in [Muneret et al. 2023] - also influences occurrence (but has little effect on conditional abundance), distinguishing late-season species (positive response) such as *Pterostichus*, *Brachinus* and *Pseudoophonus* [Joimel et al. 2021, Saska 2008] from the two *Nebria*, which undergo summer diapause [Telfer & Butterfield 2004]. The weak effect of **landscape agricultural intensification** likely reflects the coarseness of our proxy —arable land cover within a 1 km^2 buffer — which may fail to capture the landscape properties most relevant to carabid beetles [Boetzl et al. 2023].

4.1.2. Climatic variables

Climate was the second major gradient shaping carabid communities in agricultural fields, mainly through a strong positive effect of temperature on both occurrence and conditional abundance. Climate variability is indeed largely responsible for interannual variability in carabid communities, which our model appears to capture well, as suggested by the negligible residual variance of the year random effect. Furthermore, the contrasting responses of *P. madidus* (positive) and *P. melanarius* (negative) to **temperature** may reflect their mutually exclusive regional distributions [present data; Avtaeva et al. 2021] in addition to any direct thermal response. Therefore, climate responses should not be over-interpreted [Beale et al. 2010; Dormann et al. 2007]. Finally, although microclimatic temperature was not measured, its effect on carabid locomotor activity and capture efficiency is expected to be captured by the sample-level random effect.

4.1.3. Disturbance and resource gradients

The effect of disturbances on carabid occurrence was relatively weak, but **chemical disturbance** showed a significant negative effect on the conditional abundance of several species. While a negative impact of disturbances on predatory insects [Attwood et al. 2008] and carabids in particular [Geiger et al. 2010] is expected, their limited effect on occurrence in this study may result from the selection of the 20 most frequent species in agricultural landscapes —supposedly tolerant to agricultural disturbances — and does not preclude a negative effect on the occurrence of rarer species. Furthermore, a marked negative effect of chemical disturbances on the conditional abundance of *B. crepitans*, *H. dimidiatus* and *H. affinis* confirmed the detrimental impact of pesticides on carabid communities. Because *Harpalus* species are predominantly granivorous [Joimel et al. 2021], they may be doubly impacted by this gradient — directly through insecticides, and indirectly through herbicides, which deplete the weed seedbank on which they depend. Consistent with [Muneret et al. 2023], **vegetation height** showed a humped-shaped effect on species occurrences, potentially delineating the contours of an ecological niche: sufficient crop height to ensure the availability of resources (e.g. prey, refuges, shade) [Holland 2002], yet low enough not to impede movement or exacerbate interactions with vegetation-dwelling predators such as spiders or coccinellids [Birkhofer et al. 2011]

4.2. <u>Latent gradients, species associations and their multiple drivers</u>

Species responses to the environmental covariates did not fully capture occurrence or abundance variability which can represent up to 40% of the model-explained variance in occurrence (and more so for conditional abundance). As pointed by [Booij et Noorlander 1992] and more recently by [Muneret et al. 2023], resource and disturbance gradients do not fully summarize the many mechanisms driving carabid community assembly. Our jSDM was precisely designed to address what is not explained by covariates by drawing species associations across years, crop types, fields and samples.

4.2.1. Species associations are not evidence of ecological interaction

Species associations in occurrence (residual co-occurrence patterns) or in abundance (residual abundance covariation patterns) are not unambiguous signals of interactions, and may reflect unobserved environmental gradients or sampling artefacts [Blanchet et al. 2020]. In our case, for instance, and in line with the only other known application of jSDMs to carabid communities [Elo et al. 2021], we recovered a positive association in occurrence among the largest carabids (*P. madidus*, *P. melanarius*, and *C. cancellatus*) at the field level (Figure 4a – Field Level). Yet because *P. madidus* and *P. melanarius* are never found in the same region, they cannot interact directly; and even among sympatric species, positive associations may simply reflect shared responses to unobserved environmental gradients. Indeed, because all three carabids are sometimes described as forest-associated species [Joimel et al. 2021], their association could result from a common response to the proximity of woodland or deadwood at certain fields [Thiele 1977], or to human activities as suggested by [Elo et al. 2021]. Alternatively, as scavengers, *P. madidus* and *C. cancellatus* may be co-captured in certain traps containing decomposing prey [Ferrante et al. 2017]. Finally, because they engage in intraguild predation [especially *P. melanarius* Currie et al. 1996], and are opposed along the LF1 gradient (Figure 4a – Field Level) to an assemblage including *Harpalus spp.* and *A. similata* — which respond to the chemical signature of *P. melanarius* [De Heij et al. 2023, Charalabidis et al. 2019] — a biotic dimension may also underlie LF1, apparently grouping very large predators in opposition to those that tend to avoid them. [Table SM 3.1 in Supp. Mat. 3 – Part C] extends the discussion on the possible biotic or abiotic origins of associations at every level in the context of our study (or close analogues).

4.2.2. Scale dependency of ecological assembly rules

Decomposing species associations across multiple random effects helps discriminate among their potential underlying drivers —unobserved abiotic factors operating at each scale, or biotic interactions operating locally that emerge at broader scales when mediated by shared habitat features or seasonal synchrony [Mod et al. 2020]. In our study, associations in occurrence showed a clear scale dependency. At the year level, few and weak associations were detected, possibly because unobserved gradients operating at that scale are already captured by the covariates — climatic ones in particular. At the sample level, by contrast, associations were numerous and uniformly positive, most likely reflecting microhabitat conditions — prey aggregation [Honek et al. 2001], vegetation density [Speight & Lawton, 1976], or microclimate [Honek 1988] — that generally promote trapping efficiency, or carabid activity-density across all species. The two scales at which associations were strongest and most discriminating, however, were field and crop type levels, for which examples of abiotic and biotic drivers are given in [Table SM 3.1].

To our knowledge, this study is the first to include a random effect at crop type level in a jSDM, and we argue that it constitutes an effective means to uncover structuring assembly rules at a relevant agroecological scale. Gradients emerging at the crop type level at least partly reflect unobserved factors, such as crop phenology, but this is true for all levels (field, year, sample) [Table SM 3.1]. Moreover, not only do gradients at crop type level avoid the need for a spatial or temporal resolution that could bias results [König et al. 2021], but more importantly, they reflect the biotic and abiotic assembly rules underlying the marked contrast in species assemblages observed across crop types [Booij & Noorlander 1992, Eyre et al. 2013] that is weakly explained by the resource and disturbance gradients included in this study. It would have been informative to examine associations emerging at other spatially implicit structuring levels, such as the previous crop or the crop rotation [Marrec et al. 2017].

4.2.3. Species assemblages and their biotic drivers at crop type level

At crop type level, the four taxa for which crop type had a predominant effect on explained variance in occurrence and abundance were positively associated (*A. similata, Brachinus spp., A. dorsalis* and *P. cupreus*). Biotic interactions likely structure this assemblage: *Brachinus spp.* are ectoparasites of *Amara* larvae [Saska et al., 2008], *A. dorsalis* is associated with *Brachinus* *spp*. through an aposematic relationship manifested by interspecific aggregation and rubbing behaviours [Zetto Brandmayr et al., 2006], and interspecific aggregation has been documented across all four taxa [Bonacci et al., 2025]. Widely documented, this assemblage constitutes the most compelling example in this study of the capacity of residual species associations to reflect genuine biotic interactions. The mechanism by which *P. cupreus* is associated with the rest of the quartet remains difficult to identify given its eudominant nature, yet this association is a long-standing observation in the carabidology literature. Later in the season, *Nebria* species (*N. brevicollis* and *N. salina*) complete the assemblage following their summer diapause, and are therefore associated without necessarily interacting with the other species. Altogether these species constitute one pole of the LF1 gradient at crop type level, while *B. quadrimaculatum* and *P. rufipes* structure a contrasting assemblage at the other. *P. melanarius* is also associated with both species, but its response to LF1 is non-significant, probably due to its ubiquist nature.

Positive associations among *P. cupreus*, *A. dorsalis* and *A. similata* observed in occurrence were confirmed in abundance, but not for *B. sclopeta*. While, the host-parasite relationship between *B. sclopeta* and *Amara spp*. is indeed expected to promote species co-occurrence, complex predator-prey dynamics may arise obscuring the relationships between species abundances.

4.2.4. Linking species assemblages with crop type suitability as habitat

The crop type ordination biplot - based on crop types' assigned latent factor values - reveals a composite ecological gradient, reflecting each crop suitability as habitat for different species assemblages, shaped primarily by the match between crop and species phenology, the resources provided by the crop, and species interactions within each crop.

The positive pole of crop type ordination corresponds to *Brassicaceae* crops — oilseed rape and mustard — with mustard recording the highest abundances of *A. similata* and of *B. sclopeta* [Supp. Mat. 1]. The strong association of *A. similata* and its affiliated species assemblage with oilseed rape is well documented [Labruyère et al. 2018], and has been primarily attributed to the coincidence of spring-breeding carabid activity peak with oilseed rape maturity before harvest [Labruyère et al. 2018], and the availability of trophic resources during this period — including larvae of *Meligethes aeneus* and *Dasineura brassicae* as prey [Haschek et al. 2012], *Brassicaceae* seeds released through pod shatter [De Heij & Willenborg 2020] which are preferred by carabids in laboratory trials [Kulkarni et al. 2017], and *Amara* larvae as hosts for ectoparasitic *Brachinus* [Saska et al. 2008].

At the negative pole, the assemblage of *B. quadrimaculatum*, *P. rufipes* and *P. melanarius* corresponds to the dominant species recorded in maize and soybean fields, which are the only crops hosting *B. quadrimaculatum* and exhibit markedly above-average abundances of *P. rufipes*. Maize and soybean communities are further characterised by the complete absence of *A. similata* — and, in maize, of *Brachinus spp.* and *A. dorsalis* — which positions them in opposition to the positive pole of LF1. Maize and soybean share a canopy structure combining dense vegetative cover with low plant density, leaving wide inter-rows of bare soil — a configuration to which *Bembidion* has been shown to respond favourably, in wheat fields with omitted rows [Honek, 1988], and in sugar beet fields where *Bembidion* are likewise consistently dominant [Hanson et al. 2016; Rosec et al. 2026]. Finally, the late harvest of maize and soybean relative to oilseed rape and cereals — and the continuity of habitat in clover, a perennial crop, further provides an uninterrupted summer refuge for autumn-breeding species such as *P. melanarius* and *P. rufipes*.

4.2.5. Field-level segregation of B. sclopeta and *B. crepitans*

In addition to the opposition between very large carnivores (*Pterostichus* and *Carabus*) and granivores (*Harpalus* and *Amara*), the first field-level gradient (LF1) reveals a clear dissociation between the two *Brachinus* species: *B. sclopeta* and *B. crepitans*. Competitive exclusion has never been demonstrated in carabid beetles [Lövei & Sunderland 1996]; On the contrary, multispecies *Brachinus* aggregations— typically on hot dry days — have been characterized by [Schaller et al. 2018] in North America, with up to eight species co-occurring. In another comprehensive study on *Brachinus* assemblages in North American wetlands, [Juliano 1985] argued that abiotic factors — particularly water availability — are more likely to drive species co-occurrences than biotic interactions. It is possible that an abiotic trait, such as the desiccation resistance described by [Juliano 1985] for North American species, or contrasting host preferences at parasitic larval stage may account for the negative association between *B. crepitans* and *B. sclopeta* along an as-yet-unidentified field-level environmental gradient.

4.2.6. Idiosyncratic field-level response of C. auratus as a putative ghost of its past distribution

At field level, *C. auratus* is the sole contributor of the second latent gradient (LF2), accounting for nearly one third of its explained variance. Unlike most *Carabus* species, which are predominantly found in forest environments, *C. auratus* occurs in agricultural landscapes, mainly in grasslands [Mayr et al. 2007] and field margins [Bennewicz & Barczak 2020]. It is positively associated with organic farming [Sander et al. 2006,], brachypterous with limited dispersal capacity and spring-breeder with a long larval period [Marggi 1992] which renders it vulnerable to landscape fragmentation and habitat disturbance [Sander et al. 2006]. To account for the idiosyncratic response of *C. auratus* to the second latent gradient, drawing on amateur observations and by analogy with other *Carabus* species in forest environments [Pett et al. 2024, Völler et al. 2017], we speculate that *C. auratus* has been affected by ongoing habitat degradation (e.g. the rarefaction of grassland habitats [Tscharntke et al. 2005]), and persists at some fields that are no longer favorable – that is, *C. auratus* is carrying an extinction debt. This situation has been documented on carabids by [Den Boer 1977], who estimated that local populations of *Abax parallelipedus* and *P. oblongopunctatus* could survive for 40-50 years in unfavorable habitats. Confirmation of our extinction debt hypothesis by longer-term data would demonstrate the utility of jSDMs to inform conservation measures, as those called for by [Homburg et al. 2014b] and more recently by [Pett et al. 2024] for other *Carabus* species.

5. Conclusion

This study demonstrated that joint species distribution models can effectively unravel assembly rules in carabid communities across agricultural landscapes, of biotic or abiotic origins and at different spatial or non-spatial levels. Beyond recovering the dominant role of region and climate in shaping species distributions — as previously shown by [Muneret et al. 2023] — we revealed a negative effect of chemical disturbance on species conditional abundances, and highlighted that a substantial share of community structure is driven by unobserved ecological processes at crop type and field levels.

From an applied perspective, three findings deserve particular attention. First, the negative effect of pesticide pressure on conditional abundance — but not on presence–absence — suggests that monitoring programmes relying solely on occurrence data may underestimate agricultural intensification impacts. Second, the idiosyncratic response of *C. auratus*, interpreted as a signal of extinction debt, illustrates the potential of jSDMs to flag species at risk of local extinction. Third, by quantifying crop type suitability for specific carabid assemblages, our model opens perspectives to assess biocontrol services provided by these key pest predators [Masson et al. 2025].

Data Availability Statement

Data are available via the French data portal Data INRAE https://doi.org/10.15454/BMLIQI (Muneret et al. 2023) and scripts are available on Zenodo repository at this address: https://doi.org/10.5281/zenodo.20734683

References

Attwood, S.J., Maron, M., House, A.P.N., Zammit, C., 2008. Do arthropod assemblages display globally consistent responses to intensified agricultural land use and management? Global Ecology and Biogeography 17, 585–599. https://doi.org/10.1111/j.1466-8238.2008.00399.x

Avtaeva, T.A., Sukhodolskaya, R.A., Brygadyrenko, V.V., 2021. Modeling the bioclimatic range of Pterostichus melanarius (Coleoptera, Carabidae) in conditions of global climate change. Biosystems Diversity 29, 140–150. https://doi.org/10.15421/012119

Bannwart, P., Gardarin, A., Petit, S., 2025. Do semi-natural habitats enhance overwintering of generalist predators in arable cropping systems? A meta-analysis. Biological Control 201, 105700. https://doi.org/10.1016/j.biocontrol.2025.105700

Beale, C.M., Lennon, J.J., Yearsley, J.M., Brewer, M.J., Elston, D.A., 2010. Regression analysis of spatial data. Ecology Letters 13, 246–264. https://doi.org/10.1111/j.1461-0248.2009.01422.x

Bennewicz, J., Barczak, T., 2020. Ground beetles (Carabidae) of field margin habitats. Biologia 75, 1631–1641. https://doi.org/10.2478/s11756-020-00424-y

Birkhofer, K., Wolters, V., Diekötter, T., 2011. Density-dependent and -independent effects on the joint use of space by predators and prey in terrestrial arthropod food-webs. Oikos 120, 1705–1711. https://doi.org/10.1111/j.1600-0706.2011.19546.x

Blanchet, F.G., Cazelles, K., Gravel, D., 2020. Co-occurrence is not evidence of ecological interactions. Ecology Letters 23, 1050–1063. https://doi.org/10.1111/ele.13525

Boetzl, F.A., Sponsler, D., Albrecht, M., Batáry, P., Birkhofer, K., Knapp, M., Krauss, J., Maas, B., Martin, E.A., Sirami, C., Sutter, L., Bertrand, C., Baillod, A.B., Bota, G., Bretagnolle, V., Brotons, L., Frank, T., Fusser, M., Giralt, D., González, E., Hof, A.R., Luka, H., Marrec, R., Nash, M.A., Ng, K., Plantegenest, M., Poulin, B., Siriwardena, G.M., Tscharntke, T., Tschumi, M., Vialatte, A., Van Vooren, L., Zubair-Anjum, M., Entling, M.H., Steffan-Dewenter, I., Schirmel, J., 2024. Distance functions of carabids in crop fields depend on functional traits, crop type and adjacent habitat: a synthesis. Proceedings of the Royal Society B: Biological Sciences 291, 20232383. https://doi.org/10.1098/rspb.2023.2383

Bohan, D.A., Caron-Lormier, G., Muggleton, S., Raybould, A., Tamaddoni-Nezhad, A., 2011. Automated Discovery of Food Webs from Ecological Data Using Logic-Based Machine Learning. PLOS ONE 6, e29028. https://doi.org/10.1371/journal.pone.0029028

Bonacci, T., Massolo, A., Brandmayr, P., Zetto Brandmayr, T., 2025. [Interspecific aggregation in Amara, Brachinus, Anchomenus and Poecilus — reference to be confirmed].

Booij, C.J.H., Noorlander, J., 1992. Farming systems and insect predators. Agriculture, Ecosystems & Environment 40, 125–135. https://doi.org/10.1016/0167-8809(92)90088-S

Brandmayr, P., Bonacci, T., Giglio, A., Talarico, F.F., Brandmayr, T.Z., 2010. The evolution of defence mechanisms in carabid beetles: a review. In: Casale, A., Vigna Taglianti, A. (Eds.), Atti del Museo Civico di Storia Naturale di Trieste.

Bystrova, D., Poggiato, G., Bektaş, B., Arbel, J., Clark, J.S., Guglielmi, A., Thuiller, W., 2021. Clustering species with residual covariance matrix in joint species distribution models. Frontiers in Ecology and Evolution 9, 601384. https://doi.org/10.3389/fevo.2021.601384

Carbonne, B., Bohan, D.A., Foffová, H., Daouti, E., Frei, B., Neidel, V., Saska, P., Skuhrovec, J., Petit, S., 2022. Direct and indirect effects of landscape and field management intensity on carabids through trophic resources and weeds. Journal of Applied Ecology 59, 176–187. https://doi.org/10.1111/1365-2664.14043

Charalabidis, A., Dechaume-Moncharmont, F.-X., Carbonne, B., Bohan, D.A., Petit, S., 2019. Diversity of foraging strategies and responses to predator interference in seed-eating carabid beetles. Basic and Applied Ecology 36, 13–24. https://doi.org/10.1016/j.baae.2019.02.003

Chiquet, J., Mariadassou, M., Robin, S., 2021. The Poisson-Lognormal Model as a Versatile Framework for the Joint Analysis of Species Abundances. Frontiers in Ecology and Evolution 9, 588292. https://doi.org/10.3389/fevo.2021.588292

Clements, F.E., 1916. Plant Succession: An Analysis of the Development of Vegetation. Carnegie Institution of Washington, Washington, D.C. 512 pp.

Connell, J.H., 1980. Diversity and the coevolution of competitors, or the ghost of competition past. Oikos 35, 131–138. https://doi.org/10.2307/3544421

Currie, C.R., Spence, J.R., Niemelä, J., 1996. Competition, Cannibalism and Intraguild Predation among Ground Beetles (Coleoptera: Carabidae): A Laboratory Study. The Coleopterists Bulletin 50, 135–148.

Davey, J.S., Vaughan, I.P., Andrew King, R., Bell, J.R., Bohan, D.A., Bruford, M.W., Holland, J.M., Symondson, W.O.C., 2013. Intraguild predation in winter wheat: prey choice by a common epigeal carabid consuming spiders. Journal of Applied Ecology 50, 271–279. https://doi.org/10.1111/1365-2664.12008

De Heij, S.E., Willenborg, C.J., 2020. Connected Carabids: Network Interactions and Their Impact on Biocontrol by Carabid Beetles. BioScience 70, 490–500. https://doi.org/10.1093/biosci/biaa039

De Heij, S.E., Charalabidis, A., Dechaume-Moncharmont, F.-X., Petit, S., 2023. Feeding behavior and mobility of carabid beetles in response to perceived risk of predation. European Journal of Entomology 120, 73–81. https://doi.org/10.14411/eje.2023.011

Den Boer, P.J., 1977. Dispersal power and survival. Carabids in a cultivated countryside. Miscellaneous Papers, Landbouwhogeschool Wageningen 14, 1–190.

Dormann, C.F., McPherson, J.M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Kissling, W.D., Kühn, I., Ohlemuüller, R., Peres-Neto, P.R., Reineking, B., Schröder, B., Schurr, F.M., Wilson, R., 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 30, 609–628. https://doi.org/10.1111/j.2007.0906-7590.05171.x

Duflot, R., Ernoult, A., Burel, F., Aviron, S., 2016. Landscape level processes driving carabid crop assemblage in dynamic farmlands. Population Ecology 58, 265–275. https://doi.org/10.1007/s10144-015-0534-x

Elo, M., Ketola, T., Komonen, A., 2021. Species co-occurrence networks of ground beetles in managed grasslands. Community Ecology 22, 29–40. https://doi.org/10.1007/s42974-020-00034-3

Elton, C.S., 1927. Animal Ecology. Sidgwick and Jackson, London. 207 pp.

Eyre, M.D., Luff, M.L., Leifert, C., 2013. Crop, field boundary, productivity and disturbance influences on ground beetles (Coleoptera, Carabidae) in the agroecosystem. Agriculture, Ecosystems & Environment 165, 60–67. https://doi.org/10.1016/j.agee.2012.12.009

Ferrante, M., Barone, G., Lövei, G.L., 2017. The carabid Pterostichus melanarius uses chemical cues for opportunistic predation and saprophagy but not for finding healthy prey. BioControl 62, 741–747. https://doi.org/10.1007/s10526-017-9829-5

Geiger, F., Bengtsson, J., Berendse, F., Weisser, W.W., Emmerson, M., Morales, M.B., Ceryngier, P., Liira, J., Tscharntke, T., Winqvist, C., Eggers, S., Bommarco, R., Pärt, T., Bretagnolle, V., Plantegenest, M., Clement, L.W., Dennis, C., Palmer, C., Oñate, J.J., Guerrero, I., Hawro, V., Aavik, T., Thies, C., Flohre, A., Hänke, S., Fischer, C., Goedhart, P.W., Inchausti, P., 2010. Persistent negative effects of pesticides on biodiversity and biological control potential on European farmland. Basic and Applied Ecology 11, 97–105. https://doi.org/10.1016/j.baae.2009.12.001

Gelman, A., Rubin, D.B., 1992. Inference from Iterative Simulation Using Multiple Sequences. Statistical Science 7, 457–472. https://doi.org/10.1214/ss/1177011136

Gleason, H.A., 1926. The individualistic concept of the plant association. Bulletin of the Torrey Botanical Club 53, 7–26. https://doi.org/10.2307/2479933

Götzenberger, L., de Bello, F., Bråthen, K.A., Davison, J., Dubuis, A., Guisan, A., Lepš, J., Lindborg, R., Moora, M., Pärtel, M., Pellissier, L., Pottier, J., Vittoz, P., Zobel, K., Zobel, M., 2012. Ecological assembly rules in plant communities — approaches, patterns and prospects. Biological Reviews 87, 111–127. https://doi.org/10.1111/j.1469-185X.2011.00187.x

Hanson, H.I., Palmu, E., Birkhofer, K., Smith, H.G., Hedlund, K., 2016. Agricultural Land Use Determines the Trait Composition of Ground Beetle Communities. PLOS ONE 11, e0146329. https://doi.org/10.1371/journal.pone.0146329

Haschek, C., Drapela, T., Gaderbauer, P., Kromp, B., Zaller, J.G., 2012. Carabid beetle condition, reproduction and density in winter oilseed rape affected by field and landscape parameters. Journal of Applied Entomology 136, 665–674. https://doi.org/10.1111/j.1439-0418.2011.01694.x

Holland, J.M., 2002. Carabid beetles: their ecology, survival and use in agroecosystems. In: Holland, J.M. (Ed.), The Agroecology of Carabid Beetles. Intercept, Andover, pp. 1–40.

Holland, J.M., Luff, M.L., 2000. The Effects of Agricultural Practices on Carabidae in Temperate Agroecosystems. Integrated Pest Management Reviews 5, 109–129. https://doi.org/10.1023/A:1009619309424

Homburg, K., Brandt, P., Drees, C., Assmann, T., 2014. Evolutionarily significant units in a flightless ground beetle show different climate niches and high extinction risk due to climate change. Journal of Insect Conservation 18, 781–790. https://doi.org/10.1007/s10841-014-9685-x

Homburg, K., Schuldt, A., Drees, C., Assmann, T., 2013. Broad-scale geographic patterns in body size and hind wing development of western Palaearctic carabid beetles (Coleoptera: Carabidae). Ecography 36, 166–177. https://doi.org/10.1111/j.1600-0587.2012.07488.x

Honěk, A., 1988. The effect of crop density and microclimate on pitfall trap catches of Carabidae, Staphylinidae (Coleoptera), and Lycosidae (Araneae) in cereal fields. Pedobiologia 32, 233–242. https://doi.org/10.1016/S0031-4056(23)00236-6

Honěk, A., Martinková, Z., 2001. Aggregation of ground beetles (Carabidae, Coleoptera) on winter rape seeds dispersed on the ground. Plant Protection Science 37, 97–102.

Hutchinson, G.E., 1957. Concluding remarks. Cold Spring Harbor Symposia on Quantitative Biology 22, 415–427. https://doi.org/10.1101/SQB.1957.022.01.039

Joimel, S., Nahmani, J., Hedde, M., Auclerc, A., Beaumelle, L., Bonfanti, J., Cortet, J., Ganault, P., Maunoury-Danger, F., Pey, B., 2021. A large database on functional traits for soil ecologists: BETSI. Applied Soil Ecology 158, 523–528.

Juliano, S.A., 1985. Habitat associations, resources, and predators of an assemblage of Brachinus (Coleoptera: Carabidae) from southeastern Arizona. Canadian Journal of Zoology 63, 562–568. https://doi.org/10.1139/z85-250

Kegel, B., 1990. Diurnal activity of carabid beetles living on arable land. In: Stork, N.E. (Ed.), The Role of Ground Beetles in Ecological and Environmental Studies. Intercept, Andover, pp. 65–76.

Kirchhoff, T., 2020. The myth of Frederic Clements's mutualistic organicism, or: on the necessity to distinguish different concepts of organicism. History and Philosophy of the Life Sciences 42, 24. https://doi.org/10.1007/s40656-020-00317-y

Kissling, W.D., Dormann, C.F., Groeneveld, J., Hickler, T., Kühn, I., McInerny, G.J., Montoya, J.M., Römermann, C., Schiffers, K., Schurr, F.M., Singer, A., Svenning, J.-C., Zimmermann, N.E., O'Hara, R.B., 2012. Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. Journal of Biogeography 39, 2163–2178. https://doi.org/10.1111/j.1365-2699.2011.02663.x

Knapp, M., Štrobl, M., Venturo, A., Seidl, M., Jakubíková, L., Tajovský, K., Kadlec, T., González, E., 2022. Importance of grassy and forest non-crop habitat islands for overwintering of ground-dwelling arthropods in agricultural landscapes: A multi-taxa approach. Agriculture, Ecosystems & Environment 339, 108155. https://doi.org/10.1016/j.agee.2022.108155

König, C., Wüest, R.O., Graham, C.H., Karger, D.N., Sattler, T., Zimmermann, N.E., Zurell, D., 2021. Scale dependency of joint species distribution models challenges interpretation of biotic interactions. Journal of Biogeography 48, 1541–1551. https://doi.org/10.1111/jbi.14106

Kromp, B., 1999. Carabid beetles in sustainable agriculture: a review on pest control efficacy, cultivation impacts and enhancement. Agriculture, Ecosystems & Environment 74, 187–228.

Kulkarni, S.S., Dosdall, L.M., Spence, J.R., 2017. Seed Detection and Discrimination by Ground Beetles (Coleoptera: Carabidae) Are Associated with Olfactory Cues. PLOS ONE 12, e0170593. https://doi.org/10.1371/journal.pone.0170593

Labruyere, S., Petit, S., Ricci, B., 2018. Annual variation of oilseed rape habitat quality and role of grassy field margins for seed eating carabids in arable mosaics. Agricultural and Forest Entomology 20, 234–245. https://doi.org/10.1111/afe.12250

Legendre, P., Legendre, L., 1998. Numerical Ecology, 2nd ed. Elsevier, Amsterdam. 853 pp.

Lindroth, C.H., 1985. The Carabidae (Coleoptera) of Fennoscandia and Denmark. Vol. 15, Part 1. Scandinavian Press, Leiden. 225 pp.

Losey, J.E., Denno, R.F., 1998. Positive predator–predator interactions: enhanced predation rates and synergistic suppression of aphid populations. Ecology 79, 2143–2152.

Lövei, G.L., Sunderland, K.D., 1996. Ecology and behavior of ground beetles (Coleoptera: Carabidae). Annual Review of Entomology 41, 231–256. https://doi.org/10.1146/annurev.en.41.010196.001311

Marggi, W.A., 1992. Faunistik der Sandlaufkäfer und Laufkäfer der Schweiz (Cicindelidae & Carabidae), Coleoptera. Documenta Faunistica Helvetiae 13. CSCF, Neuchâtel.

Marrec, R., Caro, G., Miguet, P., Badenhausser, I., Plantegenest, M., Vialatte, A., Bretagnolle, V., 2017. Spatiotemporal dynamics of the agricultural landscape mosaic drives distribution and abundance of dominant carabid beetles. Landscape Ecology 32, 1513–1528. https://doi.org/10.1007/s10980-017-0576-x

Martin, T.G., Wintle, B.A., Rhodes, J.R., Kuhnert, P.M., Field, S.A., Low-Choy, S.J., Tyre, A.J., Possingham, H.P., 2005. Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. Ecology Letters 8, 1235–1246. https://doi.org/10.1111/j.1461-0248.2005.00826.x

Masson, A., Rastello, K., Sacco-Martret de Préville, A., Tricault, Y., Poggi, S., Canard, E., Etienne, M.-P. and Plantegenest, M. (2025), Unveiling the Hidden Feast: A Model to Translate Molecular Detection Into Predation Rate—Application Example on Biological Control by Generalist Predators in Agricultural Fields. Mol Ecol Resour, 25: e70033. https://doi.org/10.1111/1755-0998.70033

Mayr, S., Wolters, V., Dauber, J., 2007. Ground beetles (Coleoptera: Carabidae) in anthropogenic grasslands in Germany: effects of management, habitat and landscape on diversity and community composition. Applied Soil Ecology 37, 178–189.

Mod, H.K., Chevalier, M., Luoto, M., Guisan, A., 2020. Scale dependence of ecological assembly rules: Insights from empirical datasets and joint species distribution modelling. Journal of Ecology 108, 1967–1977. https://doi.org/10.1111/1365-2745.13434

Müller, P., Neuhoff, D., Nabel, M., Schiffers, K., Döring, T.F., 2022. Tillage effects on ground beetles in temperate climates: a review. Agronomy for Sustainable Development 42, 65. https://doi.org/10.1007/s13593-022-00803-6

Muneret, L., Ricci, B., Vialatte, A., Aviron, S., Ducourtieux, C., Biju-Duval, L., Petit, S., 2023. Carabid beetles have hump-shaped responses to disturbance and resource gradients within agricultural landscapes. Journal of Applied Ecology 60, 581–591. https://doi.org/10.1111/1365-2664.14357

Ovaskainen, O., Abrego, N., Halme, P., Dunson, D., 2016. Using latent variable models to identify large networks of species-to-species associations at different spatial scales. Methods in Ecology and Evolution 7, 549–555. https://doi.org/10.1111/2041-210X.12501

Ovaskainen, O., Tikhonov, G., Norberg, A., Guillaume Blanchet, F., Duan, L., Dunson, D., Roslin, T., Abrego, N., 2017. How to make more out of community data? A conceptual framework and its implementation as models and software. Ecology Letters 20, 561–576. https://doi.org/10.1111/ele.12757

Ovaskainen, O., Abrego, N., 2020. Joint Species Distribution Modelling: With Applications in R. Cambridge University Press, Cambridge, UK. 372 pp.

Pascual, M., Dunne, J.A., 2005. Ecological Networks: Linking Structure to Dynamics in Food Webs. Oxford University Press, Oxford.

Petit, S., Deytieux, V., Cordeau, S., 2021. Landscape-scale approaches for enhancing biological pest control in agricultural systems. Environmental Monitoring and Assessment 193, 75. https://doi.org/10.1007/s10661-020-08812-2

Pett, B.L., Raymond, B., Hackman, J.R., Hotchkiss, A., Knott, R., Royle, N.J., 2024. Contrasting patterns of habitat use in a threatened carabid (Carabus intricatus) and a sympatric congener in ancient temperate rainforest. Insect Conservation and Diversity 17, 213–224. https://doi.org/10.1111/icad.12717

Poggiato, G., Münkemüller, T., Bystrova, D., Arbel, J., Clark, J.S., Thuiller, W., 2021. On the Interpretations of Joint Modeling in Community Ecology. Trends in Ecology & Evolution 36, 391–401. https://doi.org/10.1016/j.tree.2021.01.002

Rainio, J., Niemelä, J., 2003. Ground beetles (Coleoptera: Carabidae) as bioindicators. Biodiversity and Conservation 12, 487–506. https://doi.org/10.1023/A:1022412617568

Rosec, M., Tricault, Y., Poggi, S., Canard, E., de Préville, A.S.M., Gendron, P., Barascou, L., Doehler, M., Buzy, S., Arnoult, A., Le Ralec, A., Plantegenest, M., Masson, A., 2026. Quantifying Species- and Community-Level Predation by Carabids on Pests and Beneficials in Wheat and Beet Fields. Entomologia Experimentalis et Applicata n/a. https://doi.org/10.1111/eea.70107

Sacco–Martret de Préville, A., Ortiz-Martinez, S., Plantegenest, M., Canard, E., 2022. Effect of Conservation Agriculture on Aphid Biocontrol by Generalist (Carabid Beetle) and Specialist (Parasitoids Wasp) Natural Enemy Communities in Winter Wheat. Frontiers in Ecology and Evolution 10, 893787. https://doi.org/10.3389/fevo.2022.893787

Sander, A.-C., Purtauf, T., Wolters, V., Dauber, J., 2006. Landscape genetics of the widespread ground-beetle Carabus auratus in an agricultural region. Basic and Applied Ecology 7, 555–564. https://doi.org/10.1016/j.baae.2006.01.001

Saska, P., Honěk, A., 2008. Synchronization of a coleopteran parasitoid, Brachinus spp. (Coleoptera: Carabidae), and its host. Annals of the Entomological Society of America 101, 533–538. https://doi.org/10.1603/0013-8746(2008)101[533:SOACPB]2.0.CO;2

Schaller, J.C., Davidowitz, G., Papaj, D.R., Smith, R.L., Carrière, Y., Moore, W., 2018. Molecular phylogeny, ecology and multispecies aggregation behaviour of bombardier beetles in Arizona. PLOS ONE 13, e0205192. https://doi.org/10.1371/journal.pone.0205192

Speight, M.R., Lawton, J.H., 1976. The influence of weed-cover on the mortality imposed on artificial prey by predatory ground beetles in cereal fields. Oecologia 23, 211–223. https://doi.org/10.1007/BF00345003

Telfer, G., Butterfield, J.E.L., 2004. The control of reproductive diapause in Nebria salina (Coleoptera: Carabidae). Ecological Entomology 29, 482–487. https://doi.org/10.1111/j.0307-6946.2004.00620.x

Thiele, H.U., 1977. Carabid Beetles in Their Environments: A Study on Habitat Selection by Adaptations in Physiology and Behaviour. Springer, Berlin.

Tikhonov, G., Opedal, Ø.H., Abrego, N., Lehikoinen, A., de Jonge, M.M.J., Oksanen, J., Ovaskainen, O., 2020. Joint species distribution modelling with the r-package Hmsc. Methods in Ecology and Evolution 11, 442–447. https://doi.org/10.1111/2041-210X.13345

Tjur, T., 2009. Coefficients of Determination in Logistic Regression Models—A New Proposal: The Coefficient of Discrimination. The American Statistician 63, 366–372. https://doi.org/10.1198/tast.2009.08210

Tscharntke, T., Klein, A.M., Kruess, A., Steffan-Dewenter, I., Thies, C., 2005. Landscape perspectives on agricultural intensification and biodiversity – ecosystem service management. Ecology Letters 8, 857–874. https://doi.org/10.1111/j.1461-0248.2005.00782.x

Völler, E., Griebeler, E.M., Baur, B., 2017. The pace of range expansion: a long-term study on the flightless ground beetle Carabus hortensis (Coleoptera: Carabidae). Journal of Insect Conservation 21, 647–658. https://doi.org/10.1007/s10841-017-0043-7

Warton, D.I., Blanchet, F.G., O’Hara, R.B., Ovaskainen, O., Taskinen, S., Walker, S.C., Hui, F.K.C., 2015. So Many Variables: Joint Modeling in Community Ecology. Trends in Ecology & Evolution 30, 766–779. https://doi.org/10.1016/j.tree.2015.09.007

Zetto Brandmayr, T., Bonacci, T., Massolo, A., Brandmayr, P., 2006. What is going on between aposematic carabid beetles? The case of Anchomenus dorsalis (Pontoppidan 1763) and

Brachinus sclopeta (Fabricius 1792) (Coleoptera Carabidae). Ethology Ecology & Evolution 18, 335–348. https://doi.org/10.1080/08927014.2006.9522700

Zurell, D., Pollock, L.J., Thuiller, W., 2018. Do joint species distribution models reliably detect interspecific interactions from co-occurrence data in homogenous environments? Ecography 41, 1812–1819. https://doi.org/10.1111/ecog.03315

Supp. Mat. 0 – PCA on seasonally cumulated covariates vs. annual covariates model

In a JSDM, all species share the same explanatory variables, yet carabid beetles differ in their phenology and may respond to environmental gradients at different times of year. We therefore needed a way to synthesize the seasonal variation of each gradient into a single predictor applicable across all species. Two approaches were considered: (i) the annual cumulative value of each gradient, as also used by [Muneret et al. 2023], which is straightforward but discards seasonal structure; and (ii) the first principal component (PC1) of a PCA performed on the four seasonal cumulative values (spring, summer, autumn, winter), which retains the dominant axis of seasonal co-variation and may better capture the overall gradient signal across the year. For each of the six gradients (vegetation height, fertilization, chemical disturbance, physical disturbance, mean temperature, and rainfall), we fitted both formulations in the hurdle model and compared their explanatory performance (Tjur R² and R²) and variance partitioning. The two approaches yielded nearly identical results across all species and both model components, and the annual cumulative formulation was therefore retained as the primary parameterization for its greater interpretability.

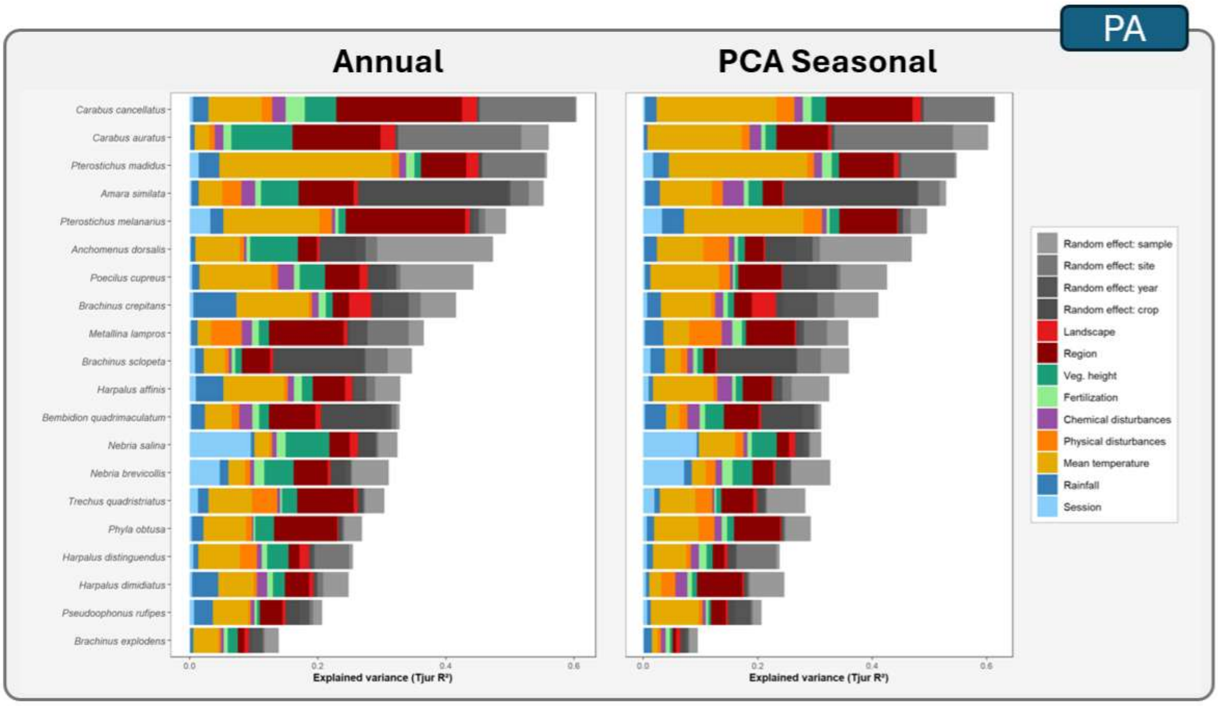


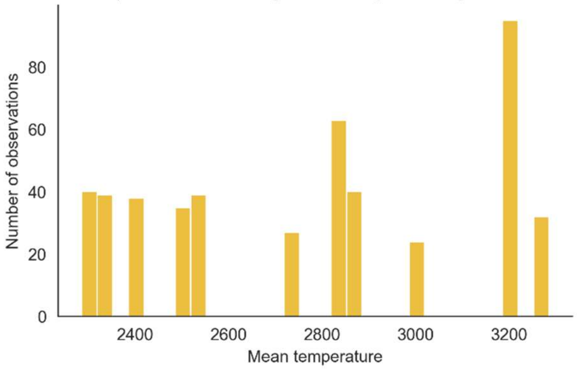
Figure SM 0.1 – Variance explained by environmental covariates and random effects, with **annual covariates** (cumulation over the previous year) on the left, or **the first axis of the PCA on seasonal covariates** (cumulation over each season prior to sampling) on the right. Bar length corresponds to the Tjur R² (PA component) and R² (AB component) for each species with coloured segments representing the contributions of each fixed effect (climate, landscape, resources, disturbances) and random effect (crop type, site, year, sample). Species are ranked in decreasing order of Tjur R².

Supp. Mat. 1 – Useful Infographics on SEBIOPAG Dataset

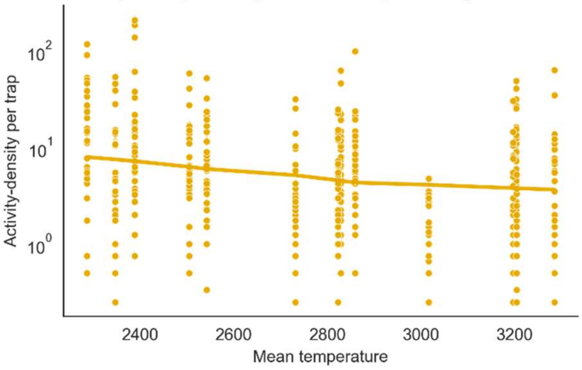
The following figures show the distribution of values for the various gradients mobilized in our study (on the right side – A – of each graph), and how they correlate with the total activity-density measured for each sample (on the left side – B). See also the associated Notebook in [Code].

Climate Gradient (Temperature and Rainfall)

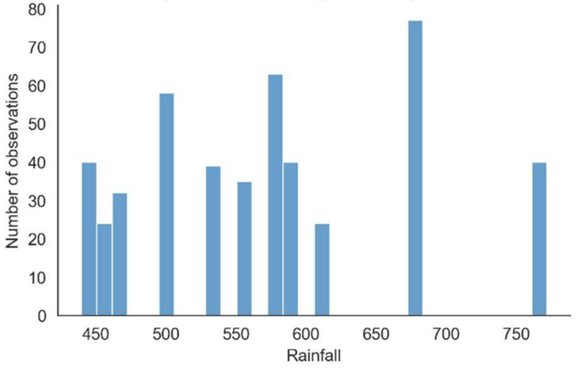
A) Distribution along Mean temperature gradient



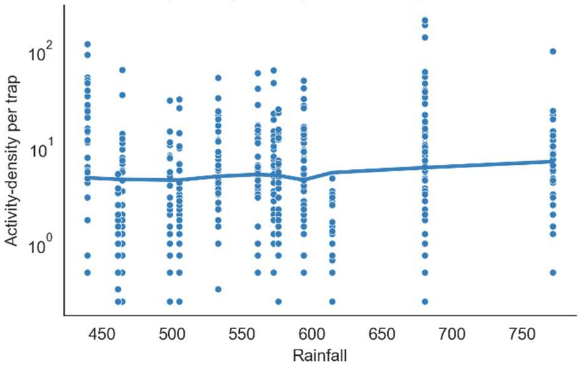
B) Activity-density vs Mean temperature gradient



A) Distribution along Rainfall gradient

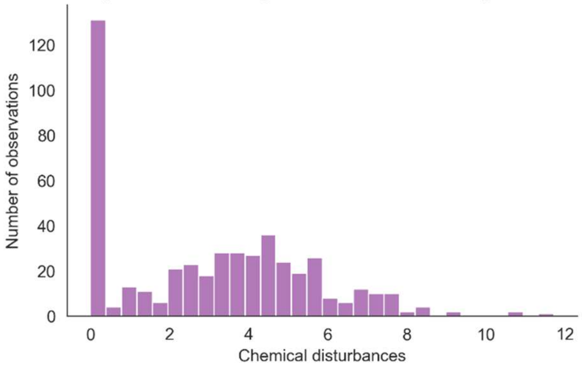


B) Activity-density vs Rainfall gradient

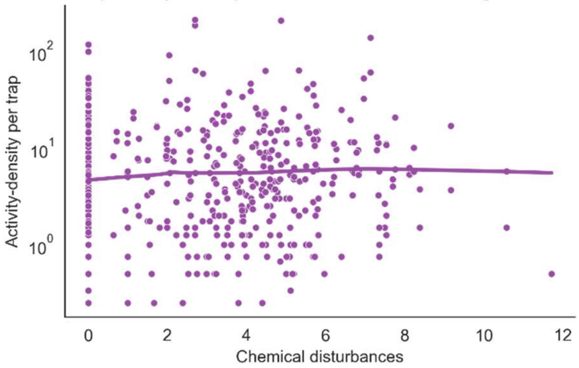


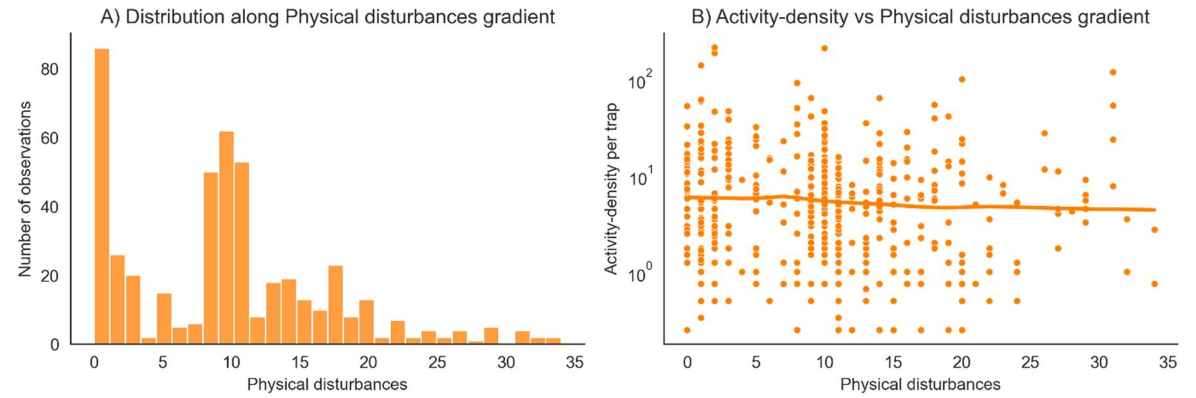
Disturbance Gradient (Chemical and Physical)

A) Distribution along Chemical disturbances gradient

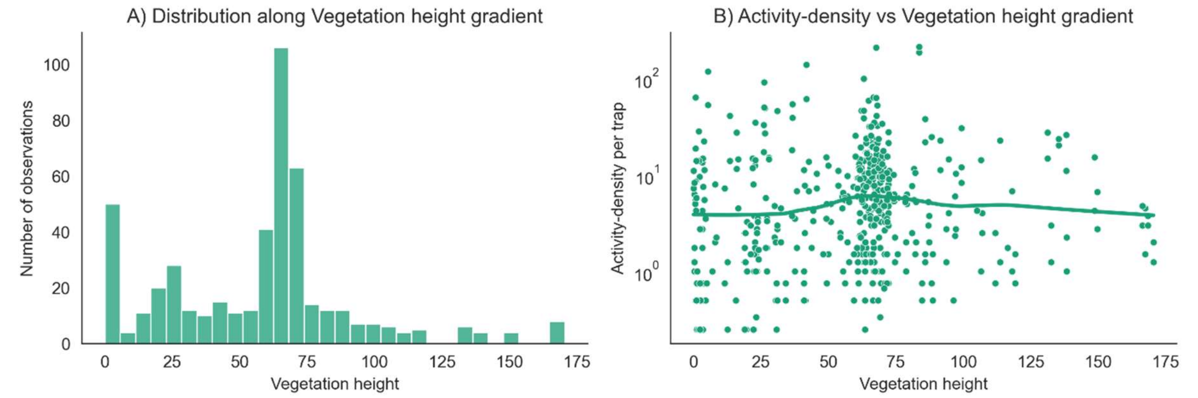


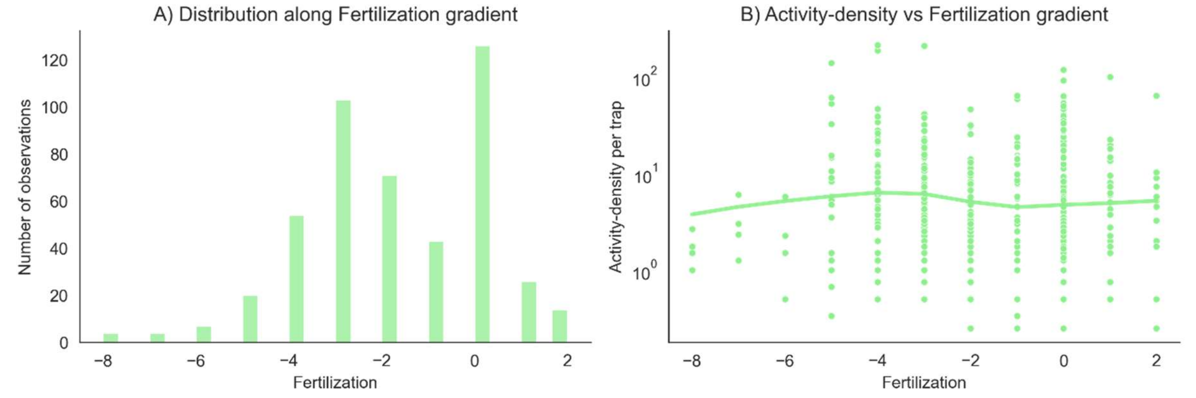
B) Activity-density vs Chemical disturbances gradient

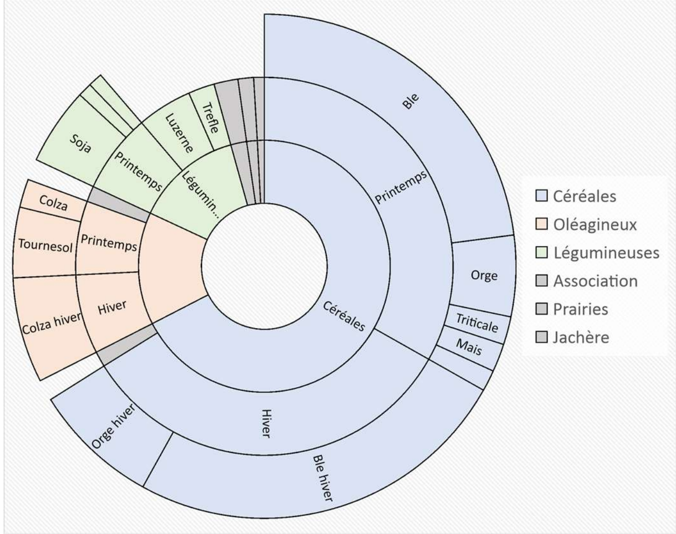




Ressource Gradient (Vegetation Height and Fertilization)



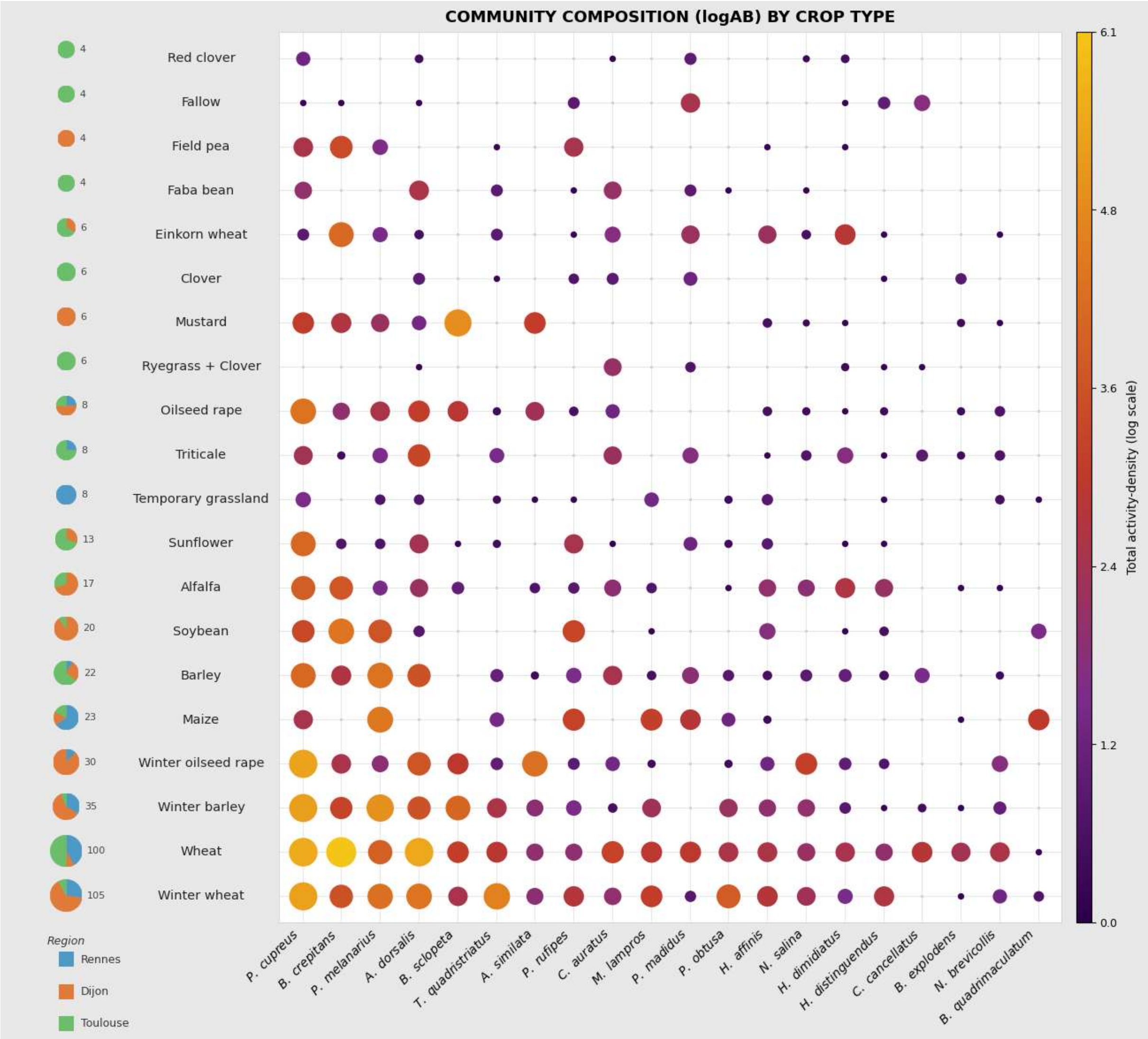




Crop Types

On the SEBIOPAG experimental design, 56 different crop types are recorded, of which 35 (shown in grey) are represented only once. Among the remaining 90%, cereals dominate, with wheat (spring or winter) being particularly prevalent. Outside cereals, few crop types exceed 20 observations, except for winter oilseed rape, soybean, alfalfa, and sunflower.

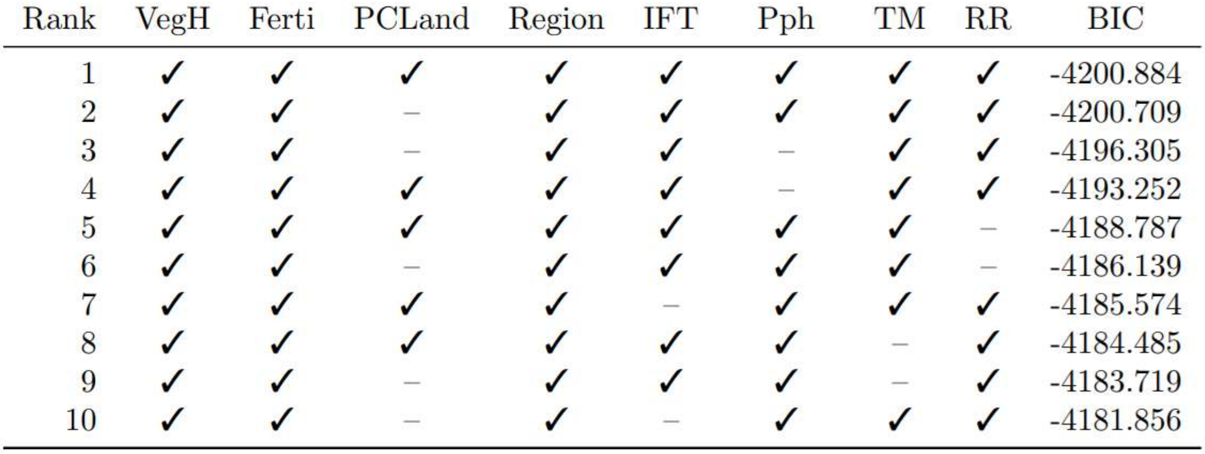
Infographics: Community Composition by Crop Types



Infographics SM 1.1 - Total activity-density of each carabid species across crop types in the SEBIOPAG dataset, based on conditional abundance (logAB). Each cell represents the total log-transformed abundance of a given species (columns) recorded across all samples of a given crop type (rows), displayed on a logarithmic scale (color and bubble size, see colorbar). Crop types are ranked by increasing number of sampled fields (indicated on the left, ranging from 4 for red clover to 105 for winter wheat). Pie charts on the left indicate the regional composition of samples for each crop type (blue: Rennes; orange: Dijon; green: Toulouse).

Supp. Mat. 2. – PLN Model Selection

Given the computational cost of fitting HMSC models, we validated our covariate selection using a multivariate Poisson lognormal regression [PLN; Chiquet et al., 2019], whose focus on community-wide species abundances makes it a conceptually appropriate surrogate for screening the 256 candidate covariate combinations. This combination achieved the lowest Bayesian Information Criterion (BIC) among all candidate models [see Table SM 2.1 below]



Rank	VegH	Ferti	PCLand	Region	IFT	Pph	TM	RR	BIC
1	✓	✓	✓	✓	✓	✓	✓	✓	-4200.884
2	✓	✓	–	✓	✓	✓	✓	✓	-4200.709
3	✓	✓	–	✓	✓	–	✓	✓	-4196.305
4	✓	✓	✓	✓	✓	–	✓	✓	-4193.252
5	✓	✓	✓	✓	✓	✓	✓	–	-4188.787
6	✓	✓	–	✓	✓	✓	✓	–	-4186.139
7	✓	✓	✓	✓	–	✓	✓	✓	-4185.574
8	✓	✓	✓	✓	✓	✓	–	✓	-4184.485
9	✓	✓	–	✓	✓	✓	–	✓	-4183.719
10	✓	✓	–	✓	–	✓	✓	✓	-4181.856

Note : The BIC is defined as $BIC = -2log(\hat{L}) + k\ log(n)$*, where* $\hat{L}$ *is the maximized likelihood of the model,* k *is the number of estimated parameters, and* n *is the sample size*

Table SM 2.1. Ranking of candidate models describing species abundances as a function of resource (VegH, Fert), global environment (PCLand, Region), perturbation (IFT, Pph), and climate (TM, RR) covariates. Check marks indicate variables (linear + quadratic effects) included in each model. Models are ranked according to the Bayesian Information Criterion (BIC), with lower values indicating a better trade-off between goodness of fit and model parsimony.

Supp. Mat. 3. – Full model specification & further details

Part A - Hierarchical joint species distribution model

Hurdle Model

Joint variations in carabid species occurrences and abundances are modelled using a hurdle joint species distribution model implemented within the HMSC framework [Ovaskainen et al. 2017a, 2020]. The hurdle structure decomposes the response into two components modelled separately: a presence–absence (PA) component and a conditional abundance (AB) component describing abundance given presence.

PA component. The probability of occurrence of species *j* in sample *i* is modelled through a probit link:

$$PA_{\,i,j} \sim Bernoulli(p_{i,j}) \;\; ; \; probit(p_{i,j}) \; = \; \eta_{PA,i,j} \;\; [\mathbf{Eq.SM1}]$$

AB component. The conditional abundance of species *j* in sample *i* — i.e. abundance given presence — is modelled using a normal distribution on log-transformed counts, with an offset accounting for sampling effort (number of pitfall traps, $N_{PT,i}$):

$$log(Y_{\boldsymbol{i,j}}) \mid Y_{\boldsymbol{i,j}} > 0 \sim Normal(\mu_{i,j}, {\sigma^2}_j) \;\; ; \; \mu_{i,j} \; = \; \eta_{PA,i,j} \; + \; log(N_{PT,i}) \quad [\mathbf{Eq.SM1bis}]$$

Although count data are inherently discrete, this log-normal approximation is commonly used within the HMSC framework, which does not natively support zero-truncated count distributions. Its adequacy is expected to increase with species abundance and number of detections.

Linear predictor

In both components, the linear predictor η_{ij} is specified in $[\mathbf{Eq.SM2}]$ as the sum **(i)** an intercept $\beta_{j,0}$ modulating the baseline abundance of species j, **(ii)** the product of the 17 predictors $\{X_k\}_{k\in[\![1,17]\!]}$ by the species-specific responses $\{\beta_{j,k}\}_{k\in[\![1,17]\!]}$, and **(iii)** a residual term that is partially shared across species to represent residual variation and covariation.

More precisely, the residual term is defined as the product of latent factors specified for each observation i and factor loadings (responses to the latent factors) estimated for each species j. In our model, latent factors are defined at four levels g corresponding to four random effects $\{sample,\; site,\; crop,\; and\; year\}$. At each level, the number of latent factors included N_g is effectively determined through shrinkage during model fitting - in our case, the shrinkage procedure retained one dominant latent factor per level, except at the field level where two latent factors were retained. The sum of the products of latent factors and factor loadings $\sum_{\ell=1}^{N_g} z_{g,c(i),\ell} \cdot \lambda_{g,j,\ell}$ - where $c(i)$ denotes the modality of level g associated with observation i (e.g. $year$ 2015) - expresses the contribution of random effect g to η_{ij}.

$$\eta_{i,j} = \log\left(N_{PT,i}\right) + \beta_{j,0} + \sum_{k=1}^{17} X_{i,k}\beta_{j,k} + \sum_{\substack{g \in \{sample,\ site,\\ crop,\ year\}}} \sum_{\ell=1}^{N_g} z_{g,c(i),\ell}\lambda_{g,j,\ell}\ [\mathbf{Eq.SM2}]$$

Traits Hierarchy

Species traits are incorporated both to inform the model (by sharing information across species) and to quantify their effects on species responses to each predictor (X_k). The species-specific coefficients $\beta_{j,k}$ (and the intercepts, $\beta_{j,0}$) were modelled as a function of species traits $(T_{j,m})_{m\in\{1,2,3\}}$ —mean body size (continuous) and two dummy variables indicating spring and autumn breeders (with no breeding season as the reference level)—via hierarchical parameters, $\gamma_{0,k}$ and $(\gamma_{m,k})_{m\in\{1,2,3\}}$, as follows:

$$\beta_{j,k} = \gamma_{0,k} + \sum_{m=1}^{3} T_{j,m}\,\gamma_{m,k} + \varepsilon_{j,k}\quad [\mathbf{Eq.SM3}]$$

Part B – Generating and Displaying Results

B.1. Variance Partitioning [Figure 2]

Model explanatory power for each species was assessed using the evaluateModelFit() function of the HMSC R package, returning Tjur R^2 for presence–absence (PA) component and R^2 for conditional abundance (AB) component [see Fig SM 3.B.1 below]. For the AB component, R^2 is computed as the squared Pearson correlation between posterior mean predictions and observed values on the log-count scale, and therefore reflects the proportion of variance in log-transformed counts explained by the model. The Tjur R^2 is defined as R^2_Tjur = $\bar{p}_1 - \bar{p}_0$, where $\bar{p}_1$ and $\bar{p}_0$ are the mean predicted occurrence probabilities at fields where the species was respectively observed and absent (Tjur, 2009).

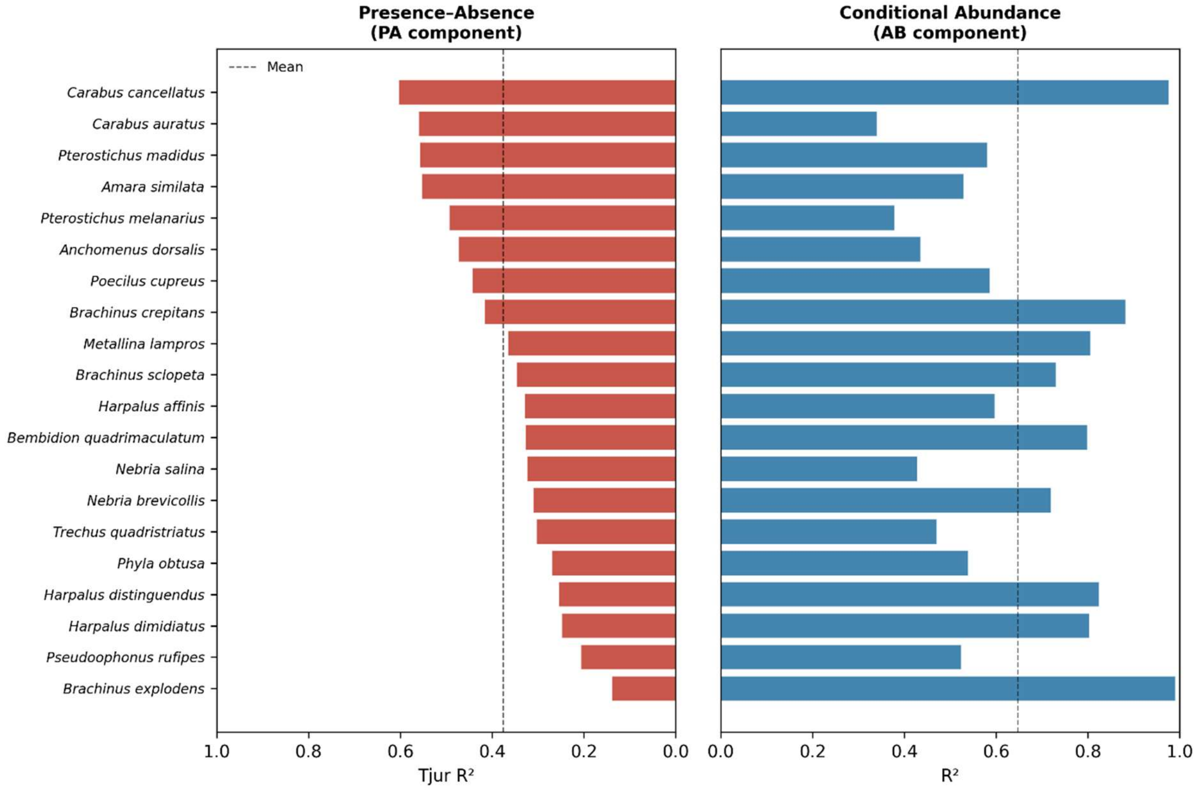


Fig SM 3.B.1. Explanatory power for Presence-Absence (PA) (Tjur R^2, left panel**) and conditional abundance (AB) (**R^2, right panel**) components**. Species are ranked in decreasing order of Tjur R^2 (top to bottom). Dashed lines indicate the cross-species mean for each component. Note that the x-axis of the PA panel is reversed to facilitate visual comparison between the two components.

Note: R^2 values for the AB component should be interpreted in light of the number of presences available per species. Typically, the R^2 of *Brachinus explodens* ($R^2 = 0.99$) is based on only 5 presences and likely reflects overfitting rather than genuine explanatory power.

Variance partitioning was obtained using the computeVariancePartitioning() function, which decomposes the variance of the linear predictor η into components explained by each fixed and random effect included in the model. The proportion of variance explained by each fixed effect X_k is expressed as $Var\left(\beta_{j,k}X_k\right)/\sum_{i=1}^{17} Var\left(\beta_{j,i}X_i\right)$. On [Fig. 2] we pooled the variance explained by the linear and the quadratic terms for each variable X_k. The proportion of variance explained by each random effect g is expressed as the matrix product of the corresponding factor loadings, $\Omega_{j,j}^{(g)} = \sum_{l=1}^{N_g} \lambda_{g,j,l}^2$.

B.2. Predicting species-specific responses to environmental gradients [Figure 3]

Posterior distributions of the species-specific regression coefficients ($\beta_{j,k}$, as in Eq. SM 3]) for each explanatory variable are shown in [Fig. 3]. The combined (linear and quadratic) response of species abundance to a focal variable depends on the reference values of the other variables in the model. To visualize this response, we used the $\mathrm{constructGradient()}$ function, which varies the focal variable across its observed range while holding other variables at their reference values (in our case Dijon region, first sampling session, and median values for all continuous variables). We then estimated the variation in the marginal prediction of the linear predictor η along this gradient using the $\mathrm{predict()}$ function. [Fig. 3] shows for examples of the variation in the linear predictor for *P. melanarius* and *C. auratus* PA along mean temperature and vegetation height gradients, and for *B. crepitans* and *T. quadristriatus* AB (conditional abundance) along Chemical and Physical disturbance gradients.

B.3. Computing pairwise species associations [Figure 4a & 4b]

For each random effect level g, the species association matrix $\boldsymbol{\Omega}^{(g)}$ is defined from the factor loadings as $\boldsymbol{\Omega}^{(g)} = \boldsymbol{\Lambda}_g {\boldsymbol{\Lambda}_g}^{\boldsymbol{T}}$, where matrix coefficient (i,j) is given by: $\Omega_{i,j}^{(g)} = \sum_{l=1}^{N_g} \lambda_{g,i,l}\lambda_{g,j,l}$. Species association matrices $\boldsymbol{\Omega}^{(g)}$ are symmetric; Therefore, only the lower triangle is shown [Figure 4a & 4b]. Diagonal coefficients $\Omega_{j,j}^{(g)} = \sum_{l=1}^{N_g} {\lambda^2}_{g,j,l}$ quantify the contribution of random effect g to the variance of the linear predictor η for species j. Off-diagonal coefficients describe residual correlations of η (= associations) between species induced by random-effect g. Association matrices were estimated at the crop, field, year, and sample levels using the $\mathrm{computeAssociations()}$ function of the HMSC R-package. For each level (g) $\mathrm{computeAssociations()}$ standardizes $\Omega^{(g)}$ to correlations: $OmegaCor_{i,j}^{(g)} = \Omega_{i,j}^{(g)} / sqrt(\Omega_{i,i}^{(g)} \cdot \Omega_{j,j}^{(g)})$.

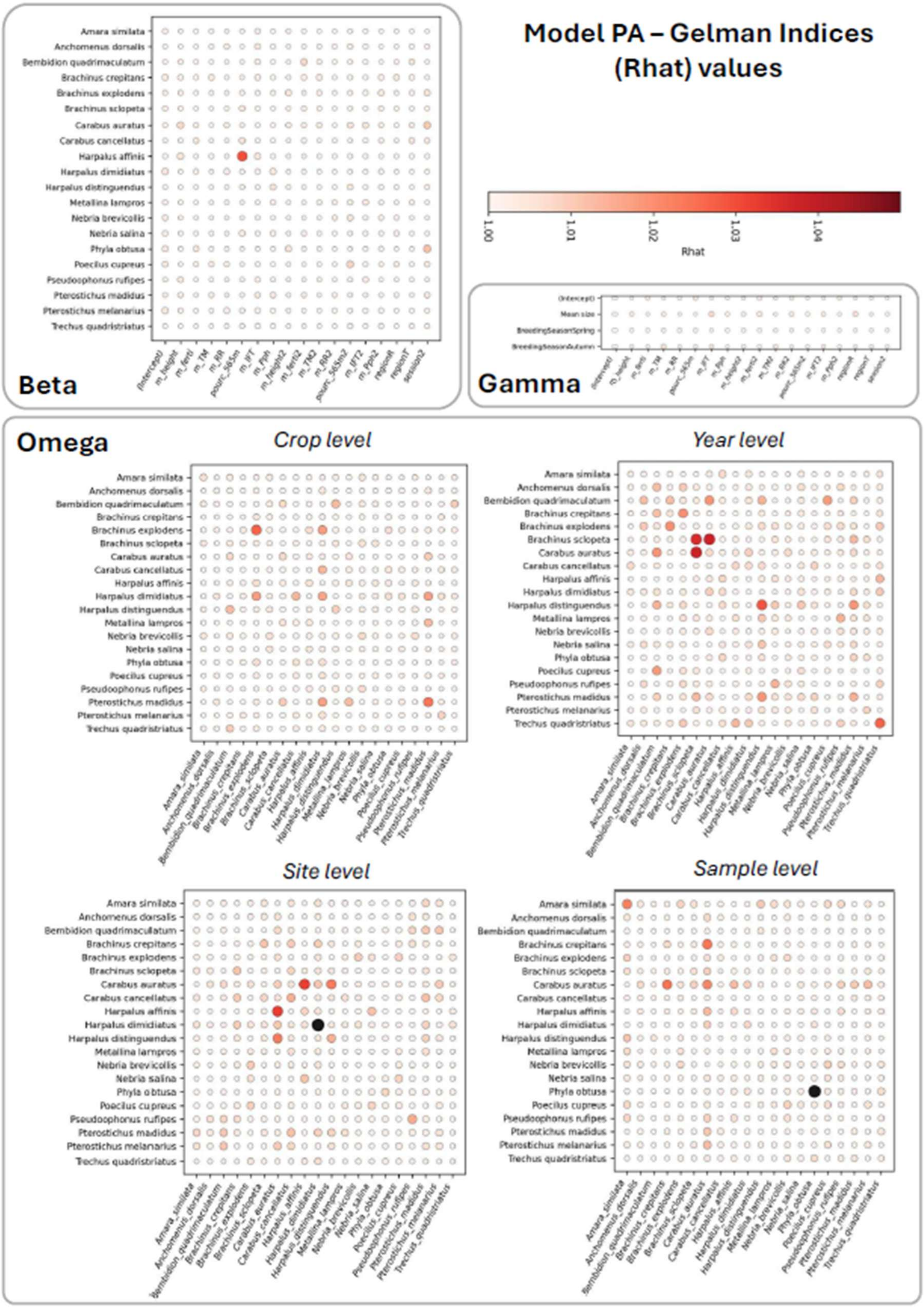
Ordination biplots were constructed separately for each random effect level by projecting the posterior means of the latent factor values $z_{g,c(i),\ell}$ (one point per unit of the random level) and the posterior means of the factor loadings $\lambda_{g,i,l}$ (one point per species) onto the plane defined by the two dominant latent factors. Units close together share similar residual community composition after accounting for fixed effects; species with similar loading vectors respond similarly to the same latent gradient.

Part C – Interpreting Results

Table SM 3.1 - Patterns of species associations and examples of biotic and abiotic mechanisms that may underlie them. Where possible, examples are drawn from or closely related to our study system. The table assumes that two species display high factor loadings at the crop, year, field, or sample scale (*columns*) — meaning that a structuring mechanism operating at that scale shapes the distribution of both species — and that they are positively associated, negatively associated, or unassociated (*lines*). Examples were selected to illustrate mechanisms that are likely to be sufficiently strong to be captured by the model. For the sake of parsimony, only one example is given per cell, although in practice associations may reflect combinations of multiple biotic or abiotic structuring mechanisms.

			Crop	Year	Field	Sample
Species association	Positive	Biotic	**Ectoparasitism** of Amara larvae by Brachinus spp. in oilseed rape fields [Saska et al. 2008]; **aposematic association** between Brachinus spp. and A. dorsalis [Brandmayr et al. 2006].	**Shared predation on a prey whose availability varies among years —** e.g. *P. cupreus* and *P. melanarius* on slugs [Oberholzer & Frank 2003], whose populations fluctuate interannually [Symondson et al. 2002]	**Shared avoidance of a common predator —** e.g. *Amara spp*. and *Harpalus spp*. both responding to chemical cues of P. melanarius [De Heij et al. 2023].	**Shared predation on aggregated prey —** e.g. seed patches favouring the simultaneous capture of P. rufipes and C. fuscipes [Honek et al. 2001].
		Abiotic	**Shared sensitivity to vegetation structure** — e.g. *B. lampros* and *B. quadrimaculatum* both responding favourably to bare soil inter-rows [Honek 1988; present study].	**Co-emergence from diapause** — e.g. *Nebria salina* and *N. brevicollis* [Lindroth 1985], whose triggering signals vary among years.	**Shared sensitivity to soil** [Luff 1996, Thomas et al. 2002] or field margins [Thomas et al. 2001, Boetzl et al. 2024] **properties**.	**Shared microhabitat** favouring the activity of both species — e.g. weed cover [Speight & Lawton 1976] or local temperature patches [Honek 1997].
	Negative	Biotic	**Competitive exclusion at crop scale** — a mechanism unlikely to occur in carabids at this scale [Niemelä et al. 1993].	**Specialization on different preys**, e.g. on seeds in *Amara* species [Saska et al. 2019], whose availability may be inversely correlated across years.	**Competitive exclusion at field scale –** e.g. between *P. cupreus* and *P. melanarius* [Thomas et al. 2001], although likely weak [Niemelä et al. 1993]	**Predation on differently distributed prey patches** — e.g. different carabid species spatially associated with different oilseed rape pests [Williams et al. 2010]
		Abiotic	**Opposite sensitivity to sowing date —** linked to the contrast between spring and autumn breeders [Bannwart et al. 2025; present study].	**Successional shift in a newly established habitat** — e.g. grassland, generating opposing dynamics between granivores and carnivores in the first years of succession [Purtauf et al. 2004]	**Differential sensitivity to a water availability gradient,** linked to interspecific variation in desiccation resistance [Juliano 1985].	**Succession between early- and late-season species,** found at the same fields each year but not within the same sampling units.
	None	**Biotic -** The symmetry of associations makes it unlikely that a strong interaction would go uncaptured by the factor loadings of both species involved. Conversely, a real interaction may not be captured if it is too weak or highly context-dependent.				
		Abiotic	**Unique sensitivity to a crop-specific gradient,** such as the specialization of *A. similata* for *Brassicaceae* fields [Marrec et al. 2017, present study]**.**	**Differential sensitivity to long-term trends** — e.g. contrasting gradual vs. sudden declines of *Demetrias* and *B. lampros* over 30 years [Holland 2002].	**Ghost of past distribution,** to which the species responds idiosyncratically. Hypothesis proposed for C. auratus [Essl et al. 2024, present study]	**Strong local aggregation** – e.g. in *Brachinus*, involving up to 8 species [Schaller et al. 2018], e.g. under a rock.

Supp. Mat. 4. – Sample quality



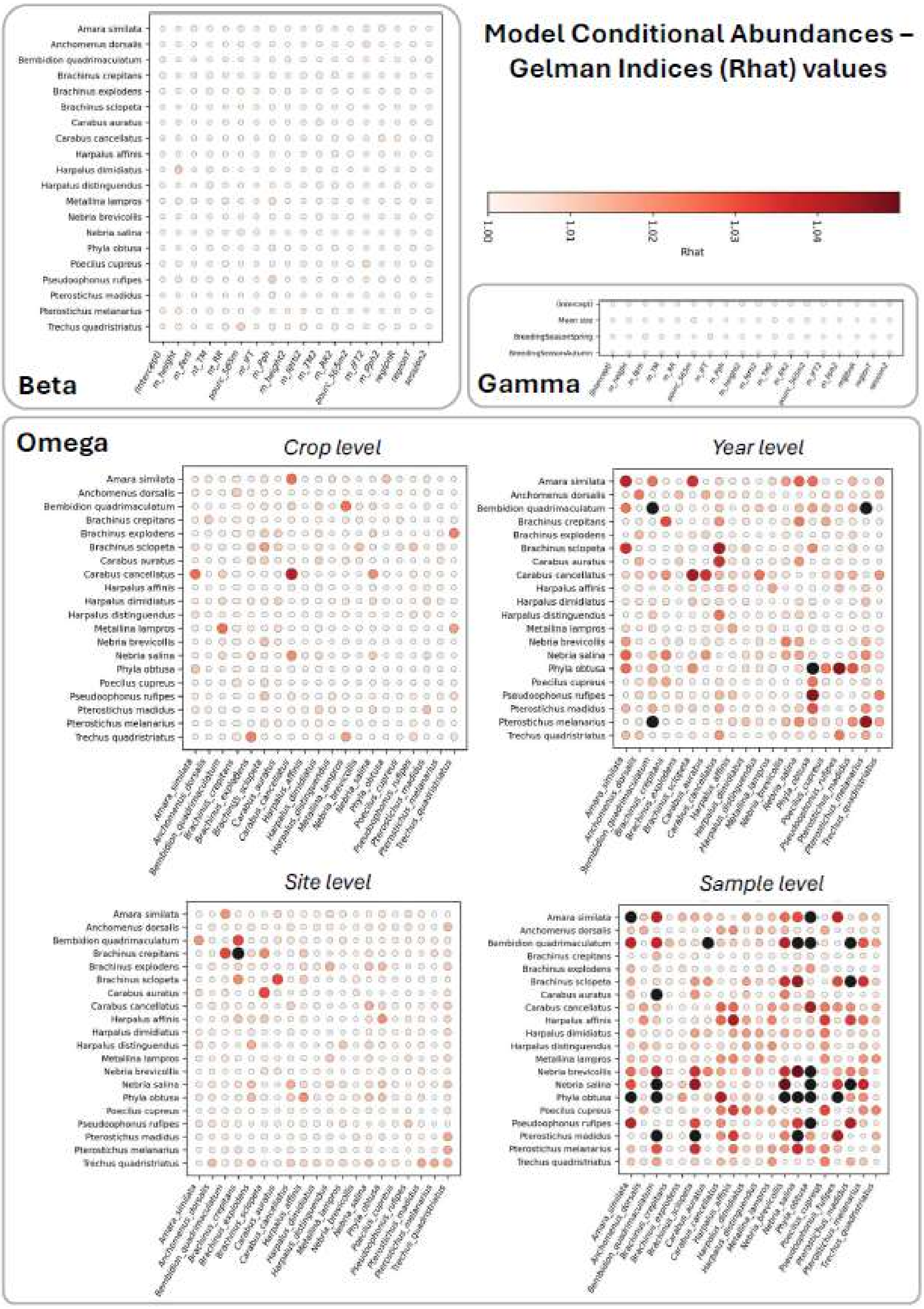


Figure SM. 4.1 - 4.2– Gelman-Rubin index values (Rhat) for (A) species responses to the various environmental gradients (gradient or factor on the x-axis, carabid species on the y-axis), and for (B) the hierarchical parameters structuring these responses according to species traits (gradient or factor on the x-axis, traits on the y-axis). Figure 4.1 for the PA and Figure 4.2 or the AB components of the model.

Supp. Mat. 5 – Effects of traits on species responses to environmental gradients

Figure SM 5.1. Fourth-corner analysis of trait–environment relationships, linking species traits (y-axis) to their responses to environmental covariates (x-axis). Green upward triangles indicate positive effects and red downward triangles negative effects; grey dashes indicate effects with uncertain direction (support < 0.90). The intercept on the covariate axis (x-axis) corresponds to the species baseline occurrence or abundance [$\beta_{j,0}$ in Eq. SM2], whereas the intercept on the trait axis (y-axis) corresponds to the baseline effect of each covariate across all species, regardless of trait values [$\beta_{0,k}$ in Eq. SM3] (see Supp. Mat. 3 for further details). Results are shown separately for the Presence-Absence (PA, left panel) and Conditional Abundance (AB, right panel) components of the Hurdle model.

Most detected effects involved the intercept terms rather than traits *per se*, reflecting baseline differences in occurrence and abundance across species and covariates rather than trait-mediated responses. The clearest trait signal was a positive association between breeding seasonality (spring or autumn breeders vs. non-seasonal species) and species responses to **pesticide pressure** (IFT) in the AB component. The negative effect of body size on regional responses in the AB component likely constitutes a statistical artefact, driven primarily by the consistently strong regional responses of large-bodied *Carabus* species and *P. madidus* rather than by a genuine size-mediated mechanism.

Supp. Mat. 6. – On Species Associations

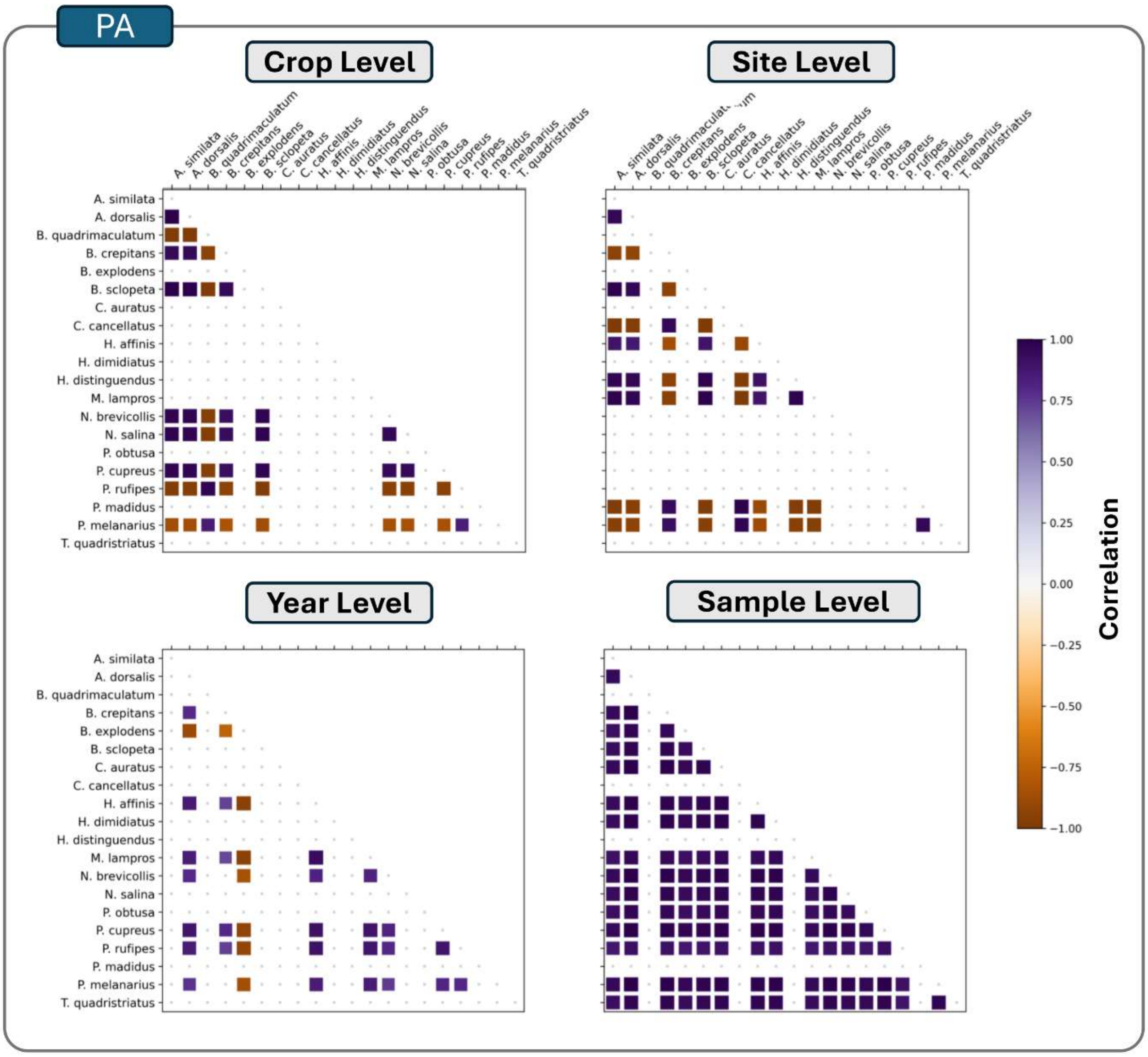


Figure SM 6.1 - Pairwise species associations for the PA model at the crop (A), field (B), year (C), and sample (D) levels, as computed using the computeAssociations() function of the HMSC R-package. Positive associations are shown in blue, negative ones in maroon. Only associations with support values (probability of having the assigned sign) greater than 0.9 are shown (see Figure 4 for the full set of associations, and ordinations biplots based on species factor loadings).

Note : Details on how species associations are computed are available in [Supp. Mat. 3]

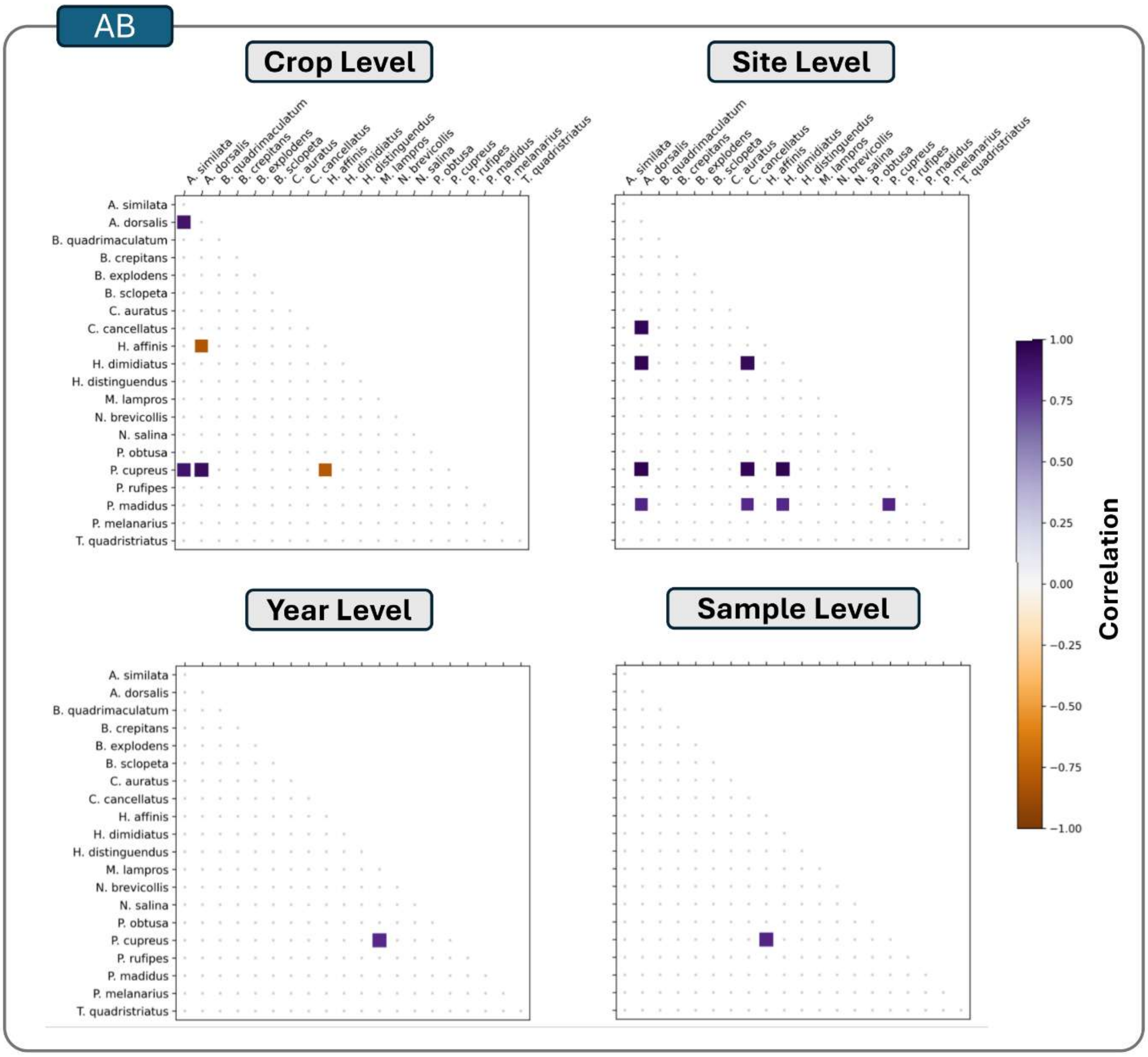


Figure SM 6.2 - Pairwise species associations for the AB model at the crop (A), field (B), year (C), and sample (D) levels, as computed using the computeAssociations() function of the HMSC R-package. Positive associations are shown in blue, negative ones in maroon. Only associations with support values (probability of having the assigned sign) greater than 0.9 are shown (see Figure 4 for the full set of associations, and ordinations biplots based on species factor loadings).

Note: Details on how species associations are computed are available in [Supp. Mat. 3]

References

Brandmayr, P., Bonacci, T., Massolo, A., Zetto Brandmayr, T., 2006. What is going on between aposematic carabid beetles? The case of Anchomenus dorsalis (Pontoppidan 1763) and Brachinus sclopeta (Fabricius 1792) (Coleoptera Carabidae). Ethology Ecology & Evolution 18, 335–348. https://doi.org/10.1080/08927014.2006.9522700

Chiquet, J., Robin, S., Mariadassou, M., 2019. Variational Inference for sparse network reconstruction from count data. Proceedings of the 36th International Conference on Machine Learning, 1162–1171.

De Heij, S.E., Ali, K., Prager, S., Willenborg, C.J., 2023. Feeding behavior and mobility of carabid beetles in response to perceived risk of predation. European Journal of Entomology 120, 81–92. https://doi.org/10.14411/eje.2023.011

Essl, F., García-Rodríguez, A., Lenzner, B., Alexander, J.M., Capinha, C., Gaüzère, P., Guisan, A., Kühn, I., Lenoir, J., Richardson, D.M., Rumpf, S.B., Svenning, J.-C., Thuiller, W., Zurell, D., Dullinger, S., 2024. Potential sources of time lags in calibrating species distribution models. Journal of Biogeography 51, 89–102. https://doi.org/10.1111/jbi.14726

Holland, J.M., 2002. Carabid beetles: their ecology, survival and use in agroecosystems, pp. 1–40.

Honek, A. (1997). The Effect of Plant Cover and Weather on the Activity Density of Ground Surface Arthropods in a Fallow Field. *Biological Agriculture & Horticulture*, *15*(1–4), 203–210. https://doi.org/10.1080/01448765.1997.9755195

Honěk, A., Martinková, Z., 2001. Aggregation of ground beetles (Carabidae, Coleoptera) on winter rape seeds dispersed on the ground. Plant Protection Science 37, 97–102. https://doi.org/10.17221/8371-PPS

Honěk, A. (1988). The effect of crop density and microclimate on pitfall trap catches of Carabidae, Staphylinidae (Coleoptera), and Lycosidae (Araneae) in cereal fields. *Pedobiologia*, *32*(3-4), 233-242.

Juliano, S.A., 1985. Habitat associations, resources, and predators of an assemblage of Brachinus (Coleoptera: Carabidae) from southeastern Arizona. Canadian Journal of Zoology 63, 1683–1691. https://doi.org/10.1139/z85-250

Lindroth, C.H., 1985. The Carabidae (Coleoptera) of Fennoscandia and Denmark. Vol. 15, Part 1. Scandinavian Press, Leiden. 225 pp.

Luff, M.L., 1996. Use of Carabids as environmental indicators in grasslands and cereals. Annales Zoologici Fennici 33, 185–195.

Marrec, R., Caro, G., Miguet, P., Badenhausser, I., Plantegenest, M., Vialatte, A., Bretagnolle, V., Gauffre, B., 2017. Spatiotemporal dynamics of the agricultural landscape mosaic drives distribution and abundance of dominant carabid beetles. Landscape Ecology 32, 2383–2398. https://doi.org/10.1007/s10980-017-0576-x

Muneret, L., Ricci, B., Vialatte, A., Aviron, S., Ducourtieux, C., Biju-Duval, L., Petit, S., 2023. Carabid beetles have hump-shaped responses to disturbance and resource gradients within agricultural landscapes. Journal of Applied Ecology 60, 581–591. https://doi.org/10.1111/1365-2664.14357

Niemelä, J., 1993. Interspecific Competition in Ground-Beetle Assemblages (Carabidae): What Have We Learned? Oikos 66, 325–335. https://doi.org/10.2307/3544821

Oberholzer, F., & Frank, T. (2003). Predation by the carabid beetles Pterostichus melanarius and Poecilus cupreus on slugs and slug eggs. *Biocontrol Science and Technology*, *13*(1), 99-110.

Ovaskainen, O., Tikhonov, G., Dunson, D., Grøtan, V., Engen, S., Sæther, B.-E., Abrego, N., 2017. How are species interactions structured in species-rich communities? A new method for analysing time-series data. Proceedings of the Royal Society B: Biological Sciences 284, 20170768. https://doi.org/10.1098/rspb.2017.0768

Purtauf, T., Dauber, J., Wolters, V., 2004. Carabid communities in the spatio-temporal mosaic of a rural landscape. Landscape and Urban Planning 67, 185–193. https://doi.org/10.1016/S0169-2046(03)00038-0

Saska, P., 2008. Composition of weed community determines carabid assemblage.

Saska, Pavel, Alois Honěk, and Zdenka Martinková. "Preferences of carabid beetles (Coleoptera: Carabidae) for herbaceous seeds." *ACTA ZOOLOGICA ACADEMIAE SCIENTIARUM HUNGARICAE: AN INTERNATIONAL JOURNAL OF ANIMAL TAXONOMY AND ECOLOGY* 65.Suppl. (2019): 57-76.

Schaller, J.C., Davidowitz, G., Papaj, D.R., Smith, R.L., Carrière, Y., Moore, W., 2018. Molecular phylogeny, ecology and multispecies aggregation behaviour of bombardier beetles in Arizona. PLOS ONE 13, e0205192. https://doi.org/10.1371/journal.pone.0205192

Speight, M.R., Lawton, J.H., 1976. The influence of weed-cover on the mortality imposed on artificial prey by predatory ground beetles in cereal fields. Oecologia 23, 211–223. https://doi.org/10.1007/BF00361237

Symondson, W.O.C., Sunderland, K.D., Greenstone, M.H., 2002. Can Generalist Predators be Effective Biocontrol Agents? Annual Review of Entomology 47, 561–594. https://doi.org/10.1146/annurev.ento.47.091201.145240

Thomas, C.F.G., Parkinson, L., Griffiths, G.J.K., Garcia, A.F., Marshall, E.J.P., 2001. Aggregation and temporal stability of carabid beetle distributions in field and hedgerow habitats. Journal of Applied Ecology 38, 100–116. https://doi.org/10.1046/j.1365-2664.2001.00574.x

Thomas, C.F., Holland, J.M., Brown, N., 2002. The spatial distribution of carabid beetles in agricultural landscapes, pp. 305–344.

Williams, I.H., Ferguson, A.W., Kruus, M., Veromann, E., Warner, D.J., 2010. Ground Beetles as Predators of Oilseed Rape Pests: Incidence, Spatio-Temporal Distributions and Feeding, in: Biocontrol-Based Integrated Management of Oilseed Rape Pests, pp. 115–149.