1   **Appendix S1**. Supplementary methods.

2   *Supporting Information* to Martin, E. A. et al. The interplay of landscape composition and

3   configuration: new pathways to manage functional biodiversity and agro-ecosystem services

4   across Europe.

5

6   **Data preparation**

7   *Sites and landscapes*

8   Raw data collected from researchers across Europe represented sampling and experiments for

9   144 arthropod families from 1960 sites in 10 countries and 29 regions in Europe. We defined

10   sites as belonging to the same study if they were measured by the same group of researchers,

11   using at least one identical method in all sites and years. According to country regulations,

12   land use maps were either directly collected from data holders, or a script was provided to

13   contributors to run calculations on locally held maps. Vector maps of land use were required

14   to be sufficiently detailed to detect boundaries between crop field polygons in order to enable

15   calculation of configuration measures (Fig. S1). As this is not the case of publicly available

16   land cover data, high quality maps compiled and partially ground-validated by study

17   participants were used. These maps were based on digitization of 1-2 m resolution

18   orthophotos and/or on official land use maps obtained from national or regional county

19   offices for each study. Minimum mapping units of the included maps, which allowed

20   boundary detection between crop fields, were 4*4 m (i.e., the smallest elements that were

21   mapped were ca. 4*4 m; see an example land use map in Fig. S1).

22   For all studies, land use maps were classified into five categories: arable fields (including

23   managed grasslands in rotation, annual and perennial crops), forest, semi-natural habitat

24   (such as hedges, grassy ditches, unmanaged grasslands, shrubs, fallows), urban areas and

25   water (see also Fig. S1). In the following, the classes 'arable land' and 'semi-natural habitat'

26   are thus not the inverse of each other, but are complemented by additional classes (e.g.

27   forest). We calculated proportions of arable land and semi-natural habitat in circular sectors

28   of 0.1, 0.25, 0.5, 1, 2 and 3 km radius around sites, or up to the maximum available radius for

29   maps with limited extent (Table S1). Sectors were centered on sampling locations within

30   fields, which averaged 2.3±1.9 ha in size (mean±standard deviation). To minimize overlap

31   between landscape sectors, radii above 500 m were excluded that led to sector overlaps of

32   >20% for a given study.

33   In a second step, non-crop classes (semi-natural habitat, forest, urban and water) were

34   merged. Using the same sectors, we calculated edge density as the total length of borders

35   between crop fields and between crop and non-crop areas, in km per ha of the landscape

36   sectors. Maps were classified and landscape metrics were calculated using user-defined work-

37   flows in ArcGIS 9.3 (ESRI 2011) and R. Following preliminary assessment of landscape

38   gradients, studies from Hungary (Bald01, Kova01, Kova02; Table S1), the only Eastern

39   European country in our dataset, were found to have greatly outlying ranges of % semi-

40   natural habitat compared to all other studies, due to the presence of large amounts of semi-

41   natural grasslands in this region. This is combined with only low values of edge density. To

42   avoid highly unequal ranges of % SNH between high and low edge density sites, we excluded

43   Hungarian studies from further analysis. Furthermore, organic sites were excluded because

44   only few studies sampled organic and conventional fields in the same landscapes. As a result,

45   the total number of site replicates included in analyses was 1,637, corresponding to 1,515

46   different landscapes of up to 3 km radius around fields (some sites having been sampled

47   multiple times per study, in several studies or in several years).

48

49   *Arthropod data*

50   Arthropods were collected according to the target functional group using methods that varied

51   by study and crop. These include pan traps, pitfall or funnel traps, malaise traps, sweep-

52   netting, visual surveys / counting, trap nests, suction sampling, and a distance method

53   (counting webs / m2). We refer to the methods of published and unpublished studies listed in

54   Appendix S2 for detailed information on each sampling method. As a rule the same sampling

55   methods and effort were applied across all sites of one study. In cases where a method was

56   applied only in a subset of sites or sampling effort varied between sites, we accounted for this

57   using model random structures (see Model formulation and analysis) and by weighting

58   abundances by the sampling effort as described below.

59   Collection of arthropods using these methods took place inside 'focal' fields (see Fig S1) at

60   varying distances from field edges, along transects with increasing distance from the edge,

61   and/or in field margins. When margins were sampled, this is noted in the column 'Site

62   description' of Table S1. For samples taken inside fields, distances from field edges and the

number of distances sampled per field varied between studies. We refer to Appendix S2 for published and unpublished information on sampling location within fields. However, the same distances from field edges were sampled across all sites of each study. We account for differences in sampling locations between studies and in some studies for multiple sampling locations per site using model random structures (see Model formulation and analysis).

We standardized data formats provided by researchers ensuring that all species observed in a study were listed in all sampled sites of that study, including when their abundance was zero. We thus ensured that sites with zero abundance were retained in subsequent subsets of functional species groups. Total abundance of arthropods in each sampling location of fields or margins was the sum of all individuals sampled in that location by a given method, in a given year and for a given study. If data of several sampling rounds were provided in one year, we calculated mean values across rounds for that year. For studies in which sampling efforts differed between sites, we first rescaled the sampling effort between 0 and 1 by dividing it by the maximum sampling effort in all sites (Newbold *et al.* 2015), then standardized abundances by the rescaled sampling effort in each site.

Ecosystem service data

Data on pollination, pest control and yields were compiled into an ecosystem service index (ESI). Measurements included visitation rates of pollinators (per flower and time), fruit set and weight, seed set, predation and parasitism of pests, pest damage, pest density in vs. outside enemy exclusion cages, and crop yields as biomass/area or biomass/plant. Detailed units and measurement methods applied per study for each ecosystem service are specified in Table S3. Several measurements were frequently performed per study. ESI values were defined according to the type of measurement considered:

1) For cage experiments comparing measures with and without pollination or pest control, only open and closed treatments were considered (excluding intermediate levels such as vertebrate-only exclosures). If several replicates existed per exclusion treatment, these were aggregated to mean values per treatment and site. Exclosures were defined as pollinator exclosures if the main focus of the study was on pollination, and as enemy exclosures if the main focus was on pest control. Similarly, fruit, pod and seed set were defined as pollination variables (not yield variables) when the study included a focus on pollination. The ecosystem service index (ESI) for cage experiments was calculated as:

95   (a) $\begin{cases} ESI = 1 - \dfrac{R_{closed}}{R_{open}} & R_{open} > R_{closed} \\ \\ ESI = 0 & R_{open} \leq R_{closed} \end{cases}$ (b) $\begin{cases} ESI = 1 - \dfrac{R_{open}}{R_{closed}} & R_{open} < R_{closed} \\ \\ ESI = 0 & R_{open} \geq R_{closed} \end{cases}$

96   where R_{open} is the service measure in open treatments and R_{closed} is the service measure in

97   controls excluding either pollinators or natural enemies (e.g. Gardiner *et al.* 2009, Rusch *et*

98   *al.* 2013). For measures representing positive services (seed set, fruit weight, fruit set, yield,

99   seed predation), higher service provision corresponds to higher values in open than in closed

100   cages (a). Inversely, for measures representing negative services (exclosure pest density, crop

101   damage), higher service provision corresponds to lower values in open compared to closed

102   cages (b). We set the ESI to zero when no difference was found between treatments or when

103   differences between treatments led to negative ESI (i.e. no service provided; this occurred in

104   24 (6%) of observations of pollination service provision). In one study measuring exclosure

105   pest density, initial densities varied between sites. Here we calculated R_{open} and R_{closed} as

106   the growth rate of pests in open and closed treatments, following: $R = \dfrac{N_t}{N_{t0}}$ where N_t is the

107   density of pests at the end, and N_{t0} is the density of pests at the start of the experiment.

108   2) Other measures included direct assessments of ecosystem service without a comparative

109   exclosure experiment (e.g. proportion of parasitized pests), and one assessment based on

110   exclosures for which treatment outcomes (i.e. the exclosure ESI described above) were

111   processed directly by data contributors as the average difference in seed set between open and

112   bagged plants (Dain01-seed set; Table S3). In these cases, the ESI was the actual value of the

113   measurement without further processing. Differences in ranges of the resulting ESI between

114   studies are provided in Table S3.

115   In order to consider effects of landscape variables in models, we did not standardize the ESI

116   within studies, years and measurement types. Instead, we account for differences in

117   measurement type and resulting ESI range within and between studies by including random

118   structures in mixed effects models (see below, § Model formulation and analysis), and

119   ln(x+1)-transformed ESI variables to conform to assumptions of normality and

120   homoscedasticity (see also Appendix S3 for results using standardized responses).

121

122

Functional trait classification and cluster regression

Functional group and trait classification

In a first step, we classified organisms into functional groups of natural enemies, pollinators and pests according to literature, online sources (e.g. www.discoverlife.org) and author's knowledge. We defined natural enemies as predatory or parasitic on other arthropods at any life stage. We defined pests as causing damage to crops that has been found to reach economic injury levels. Pollinators were organisms able to transport pollen between flowers or flower parts. If organisms belonged to several functional groups throughout their life cycle, we classified them according to the life stage sampled (e.g. syrphids sampled as adults were defined as pollinators, but aphidophagous syrphid larva were defined as predators). Non-pest herbivores, ants, decomposer beetles and parasitoids of bees could not be placed into these groups and were not considered in functional group and trait analyses. Table S2 reports the detailed classification of taxa into functional groups. Carabids of the genera *Amara* (Bonelli, 1810)*, Ophonus* (Dejean, 1821) and *Harpalus* (Latreille, 1802), and genera of Staphylinidae based on Clough *et al.* (2007) were considered non-predatory and were excluded from functional trait analyses that focussed on enemies, pollinators and pests. References used for further taxa are listed in trait descriptions below. Honeybees (*Apis mellifera* [Linnaeus, 1758]) were excluded from all analyses (and from functional trait classification) because their abundance in the agricultural landscape is strongly influenced by the placement of managed hives.

With the aim of creating usable trait classes relevant for a broad range of taxa (Table S2), we then classified organisms into broadly defined categorical traits based on existing databases (syrphids: SyrphTheNet, Speight *et al.* 2010; carabids: carabids.org, Homburg *et al.* 2014; staphylinids: Gossner et al. 2015), published literature and expert knowledge. The traits we included are defined in Table 1 with additional detail as follows:

Diet breadth: for bees excluding honeybees, we defined diet breadth as specialist for oligolectic, and generalist for polylectic species. Lecticity was extracted from existing sources when available (e.g. Gagic *et al.* 2015; Kremen & M'Gonigle 2015; Normandin *et al.* 2017; Ascher & Pickering 2018; Bees, Wasps & Ants Recording Society 2018; 247 out of 358 species). The remaining species were assumed to be polylectic if belonging to parasitic families (24 species) or all other species of the genus were polylectic (11 species). Some species were assumed to be oligolectic because all other species of the genus are oligolectic

155 (11 species). Species only identified to genus or family (29 species) were assumed to share

156 the lecty of the majority of species in the same genus (family). The remaining 36 species

157 were assumed polylectic. We believe this is a safe assumption, as they fall in genera with a

158 predominance of polylectic species (e.g. *Lasioglossum* [Curtis, 1833], *Eucera* [Scopoli,

159 1770]). For spiders, stenophagy was based on Pekár & Toft (2015). If known for a given

160 species, it was extrapolated to the whole genus as stenophagy tends to be consistent within

161 genus. However, in genera in which stenophagy is unknown for common European species

162 but known for some exotic species, we did not classify species as stenophagous (e.g. *Dipoena*

163 [Thorell, 1869]). Predators including carabid, staphylinid and lady beetles, aculeate wasps,

164 lacewings, earwigs, other Coleoptera and Hemiptera (soldier and checkered beetles; damsel,

165 shore, flower and pirate bugs; other bugs of the families Lygaeidae, Pyrrhocoridae) feed on

166 more than one family of prey and were classified as generalists. Similarly, adult tachinids

167 (Stireman *et al.* 2006), other flies (non-hoverflies), polyphagous butterflies (Tolman &

168 Lewington 2008) and pollen-feeding beetles of the genus *Oedemera* (Olivier, 1789)

169 (Oedemeridae) were classified as generalist pollinators. Tachinids sampled as larvae were

170 considered generalist enemies (Stireman *et al.* 2006). Planthoppers followed Nickel &

171 Remane (2002) (mono- and oligophages as specialists, polyphages as generalists). Generalist

172 pest herbivores included two shield bugs (*Pentatoma rufipes* [Linnaeus, 1758] and

173 *Carpocoris purpureipennis* [De Geer, 1773], Pentatomidae), species of plant bugs

174 (*Adelphocoris lineolatus* [Goeze, 1778]*, Lygocoris pabulinus* [Linnaeus, 1761]*, Lygus

175 gemellatus* [Herrich-Schaeffer, 1835]*, L. pratensis* [Linnaeus, 1758]*, L. rugulipennis*

176 [Poppius, 1911], Miridae; other plant bugs also include generalist non-pest herbivores and

177 predators; Wheeler 2001; Cassis & Schuh 2011), the garden chafer beetle (*Phyllopertha

178 horticola* [Linnaeus, 1758], Scarabeidae; Milne & Laughlin 1956) and the cowpea aphid

179 (*Aphis craccivora* [C.L.Koch, 1854], Aphididae). In contrast, oligo- and monophagous

180 butterflies, parasitoid wasps (considered enemies), the aphid midge *Aphidoletes aphidimyza*

181 (Rondani, 1847) (Cecidomyiidae), and other pest herbivores (e.g. several aphids, the cereal

182 leaf beetle, the rape pollen beetle) were classified as diet specialists feeding mainly on host

183 plants or prey of no more than two families. The larvae of non-aphidophagous hoverflies

184 most often rely on food sources in only one specific 'microhabitat' (relative to the scale of

185 landscapes), but the actual species consumed are unknown (e.g. dead wood, stagnant water;

186 Speight *et al.* 2010). While these microhabitats do not represent a taxonomic family, we

187 chose to classify hoverflies as diet specialists due to their reliance on only one type of

188 microhabitat or one main family of prey (aphids) for predacious species.

6

189   Agricultural specialism: all pest herbivores that are diet specialists (including the cabbage

190   white *Pieris brassicae* [Linnaeus, 1758] [Lepidoptera], sampled as adult and thus considered

191   here as a pollinator) were defined as agricultural specialists (i.e. they are diet specialists on

192   agricultural host plants). Furthermore, aphidophagous syrphids (Speight *et al.* 2010;

193   considered pollinators when sampled as adults), the aphid midge *A. aphidimyza* (Watanabe *et*

194   *al.* 2014) and parasitoid wasps sampled in our study systems were defined as agricultural

195   specialists whose prey or hosts include mainly agricultural pests. All other taxa that are diet

196   generalists, able to feed on hosts or prey both within and outside agricultural fields, were

197   defined as non-agricultural specialists. Taxa that are diet specialists but whose hosts or prey

198   are neither agricultural host plants, nor pests specialized on agricultural hosts, were also

199   defined as non-agricultural specialists. These include diet specialist bees, spiders,

200   planthoppers, butterflies, and non-aphidophagous hoverflies (see taxon-specific references

201   above).

202   Diet life history: we defined taxa who experience a shift in the type of above-ground diet

203   resources during their life cycle (plant matter *vs*. flower products *vs*. animal matter, e.g.

204   carnivorous larvae *vs*. nectar- and honeydew-feeding adults) as including Lepidoptera,

205   Diptera, parasitoid wasps (O'Neill 2001), the aphid midge (Watanabe et al. 2014), sawflies

206   (Boevé 2008), and oedemerids (Vázquez 2002). Other taxa were defined as maintaining a

207   similar or overlapping above-ground diet type throughout their life cycle (e.g. carabid beetles

208   carnivorous at larval and adult stages, aculeate wasps carnivorous as larvae and omnivorous

209   as adults; O'Neill 2001).

210   Overwintering habitat: data on whether or not species overwinter (and survive to emerge in

211   spring) in agricultural fields are difficult to obtain, and likely strongly depend on winter field

212   management (till depth, presence of stubble) and/or the presence of a winter crop (Sarthou *et*

213   *al.* 2014). Based on expert knowledge and published literature, we defined organisms

214   considered to overwinter mainly outside crops, and that are not likely to remain or survive the

215   winter within crops, as 'overwintering in non-crop areas'. The majority of taxa in our datasets

216   (see Table S2 for detailed listing), including bees, wasps, tachinids, butterflies, non-predatory

217   hoverflies, bugs, several beetle taxa, 12 species of pests were defined as overwintering in

218   non-crop habitats based on expert knowledge and literature (e.g. Brodeur & McNeil 1990;

219   Leather 1993; Nickel 2003; Öberg *et al.* 2008). If no information was available for a taxon

220   (e.g. identified only to family or order), we assumed its overwintering habitat to be the same

221   as the majority of taxa in the same family (order). We defined organisms with generalist

overwintering requirements and/or that may also overwinter within crops as taxa for which empirical data or expert knowledge suggested this. These included 10 species of pests (following Milne & Laughlin 1956; Leather 1993; Alford 2014; Sutter *et al.* 2018), one spider (*Oedothorax apicatus* Blackwall 1850, Mestre *et al.* 2018), carabid and staphylinid beetles (Pfiffner & Luka 2000; Frank & Reichhart 2004) and aphidophagous syrphids (Raymond *et al.* 2014). However, some carabid species known to be mostly associated with non-crop habitats (forest, swamp) were defined as overwintering in non-crop areas. When published literature was insufficient to distinguish overwintering requirements of individual species, we generalized these to the whole taxon and resorted to expert knowledge for refinements wherever possible (coauthors M.H.E. for spiders, J.K. for butterflies and aphids; F. Bötzl and L. Pfiffner for carabids).

Dispersal mode: we defined dispersal mode as "flight" for carabid species with developed wings and "ground" otherwise (species with undeveloped, dimorphic and unknown wing shape). Further, we classified Orbicularia spiders (Araneidae, Tetragnathidae, Theridiidae and Linyphiidae) as aerial wind dispersers which frequently balloon, as these have been found to balloon more often than non-Orbicularia which we classified as ground-dispersers (Entling *et al.* 2011). Other beetles (checkered, soldier beetles) and earwigs with no aerial dispersal were defined as ground-dwellers. Taxa able to actively fly were classified as active fliers ('flight'; bees, aculeate wasps, butterflies, flies, bugs, beetles following Gossner *et al.* 2015). Taxa capable of active flight but are also frequently transported on wind currents were defined as 'flight/wind' dispersers (parasitoid wasps, aphids, the aphid midge *A. aphidimyza*, cereal leaf and pollen beetles, weevils; Sawyer & Haynes 1986; Westbrook *et al.* 2000; Skellern *et al.* 2017; Ludwig *et al.* 2018).

Stratum: taxa that forage mainly by walking, jumping or web-building on the ground or within vegetation were defined as occupying the ground and vegetation strata. These included aphids, spiders, earwigs, beetles (except lady, cereal leaf and pollen beetles), shore, shield, and damsel bugs. Taxa that frequently forage by flying between target hosts or to find prey were defined as occupying the aerial stratum, including bees, wasps, butterflies, flies, lacewings, midges, rape stem weevils, lady, cereal leaf and pollen beetles, planthoppers, saw- and stoneflies.

Full trait data for all species are accessible at doi:10.5061/dryad.6tj407n. Species and family names were resolved using online databases (NCBI 2018; World Spider Catalog 2018) and

254 classified using R package taxize v.0.9.0 (Chamberlain & Szöcs 2013; Chamberlain *et al.*
255 2016).

Cluster regression

257 Functional groups of organisms with similar impacts on, or responses to, their environment
258 can be defined *a priori*, using existing knowledge or assumptions on which species should
259 have similar impacts or responses to the environment. Alternatively, functional groups can be
260 defined using classification methods that identify emergent groups with similar
261 environmental responses. We used one such method, cluster regression, to aggregate the
262 diversity of responses of individual species into emergent groupings (trait syndromes) formed
263 with less subjectivity than *a priori* groupings (Kleyer *et al.* 2012).

264 Cluster regression is a multivariate approach that first builds species groups from all possible
265 combinations of their functional traits, then separates trait combinations (syndromes) that are
266 responsive to environmental gradients from those that are neutral. The method assumes that
267 the species pool in a landscape can be classified based on correlations between traits due to
268 underlying tradeoffs or allometries, and that the emergent groups can have similar responses
269 to the environment (Kleyer *et al.* 2012; Lavorel *et al.* 1997; Lavorel and Garnier 2002). The
270 first step of this approach involves forming groups based on hierarchical clustering of the
271 functional trait space. For this, the trait space is defined successively using one to all six
272 available functional traits. In a second step, the resulting groups based on all possible
273 combinations of traits are regressed against environmental variables (see below). Goodness of
274 fit measures of these regressions then allow to compare and identify the most parsimonious
275 trait combinations that are responsive to environmental variables.

276 Following Kleyer *et al.* (2012), we performed a cluster regression on the six categorical traits
277 defined in Table 1 using Ward's hierarchical agglomerative clustering algorithm based on
278 Gower distance between categorical traits (function hclust in R package stats v.3.4.3, using
279 Ward's minimum variance method with clustering criterion; Murtagh & Legendre 2014).
280 Clusters were bootstrapped 500 times to assess cluster stability and the Calinski-Harabasz
281 stopping criterion was used to determine the optimal number of clusters for each combination
282 of traits. Using R code adapted from Kleyer *et al.* (2012), we then performed a cluster
283 regression of all stable cluster combinations. For this, we used linear models to regress the
284 ln(x+1)-transformed abundance of all stable cluster combinations against linear and 2nd
285 degree polynomial explanatory landscape variables, measured at 500 m radius around sites.

286   As 500 m is the minimum radius for all studies, all species of all studies were included.

287   Tested explanatory landscape variables were % seminatural habitat (SNH), % arable land and

288   edge density (ED). These variables were hypothesized to act as environmental filters for

289   different species groups (see Introduction and Methods). Performing a cluster regression with

290   these variables allowed us to identify shared responses to environmental filters of different

291   possible species groups (clusters) obtained by hierarchical clustering. These regressions used

292   simplified models compared to the main analyses (see Methods and 'Model formulation'

293   below). They were used to identify the most parsimonious combinations of traits and

294   numbers of clusters for the set of landscape variables that we explore in subsequent analyses

295   (see Statistical analyses; Kleyer *et al.* 2012). Prior to regression, abundances of each study

296   were standardized between 0 and 1. Following Kleyer *et al.* (2012), linear and 2nd degree

297   polynomials of landscape variables were applied in univariate and multivariate additive

298   models with all combinations of uncorrelated variables. After model-averaging significant

299   univariate and multivariate models (with $p \leq 0.05$) (Burnham & Anderson 2002), we

300   calculated the R^2 of averaged models as the squared correlation between observed and

301   predicted (model-averaged) values for each cluster. We then selected optimal parsimonious

302   trait combinations as those for which the number of clusters and the correlation between

303   observed and predicted values did not increase with the inclusion of additional traits

304   (Bernhardt-Römermann *et al.* 2008). Clusters obtained from parsimonious trait combinations

305   were defined as trait syndromes. As trait syndromes may differ between functional groups

306   (due to interactions between effect and response traits; Lavorel & Garnier 2002), we

307   performed these analyses separately for pollinators, enemies and pests.

308

309   **Model formulation and analysis**

310   Here we present the structure of full models analysed in R. Due to several studies reporting

311   average values of abundance over multiple censuses in each site, generalized models for

312   count data were not used. Full linear mixed models included edge density, one composition

313   variable (% semi-natural habitat or % arable land) and their interaction as fixed effects. In

314   addition, we expected non-linearity including hump shapes for several parameters (Fig. 1).

315   We thus included quadratic terms (R function poly() scaled using package polypoly v.0.0.2;

316   Mahr 2017) for predictors and their interactions as recommended by Hainmueller *et al.*

317   (2017) (see all terms included in full models excepting intercepts Fig. S5-7). Below, we

318 provide an explanation of how to interpret quadratic interactions in these models (§ Model fit

319 and interpretation). As we expected changes at low values of predictors to have stronger

320 impacts than at high values, we ln(x+1)-transformed all landscape predictors. This

321 transformation improved model fits (R^2 calculated as the variance explained by marginal and

322 conditional effects, respectively; Nakagawa & Schielzeth 2013) compared to untransformed

323 variables and was maintained in all analyses.

324 To account for multiple nesting of sites according to study, year and measurement method,

325 we built random structures as follows. Random intercepts were SY (study-year, the

326 combination of one study in one year); SYM (sampling method nested in study-year, for

327 studies with multiple sampling methods e.g. targeting different taxonomic groups) and

328 SYMB (block within method within study-year). Blocks grouped together observations

329 sampled from multiple distances in the same field (3 studies), from multiple locations in the

330 same landscape (e.g. fields and their margins; 11 studies), or from multiple locations in the

331 same region for studies including several regions (2 studies). Since these random terms are

332 strictly nested, specifying crossed effects is equivalent to specifying nested ones (Newbold *et*

333 *al.* 2015). As similar studies were frequently performed in the same area, occasionally in the

334 same year, and studies with multiple years usually used different sites each year, we did not

335 nest year within study. Instead, we considered each year of multi-year studies to be an

336 independent dataset and used study-year combinations as the highest hierarchical unit.

337 Further, landscape effects may vary according to study and year. Initial models included

338 random slopes with respect to study-year for edge density, landscape composition and their

339 interaction (Barr *et al.* 2013). To avoid overparameterization, we then obtained parsimonious

340 random structures for each model accounting for the variance explained by each random term

341 using the 'RePsychLing' procedure (Baayen *et al.* 2015; Bates *et al.* 2015). In functional

342 group analyses, we excluded study-year-method combinations (SYM) where fewer than 4

343 sites and <20% of all sites had non-zero abundance. For some responses and landscape radii,

344 variances were abnormally inflated despite the use of orthogonal polynomials. This is likely

345 due to high variable correlation for particular groups and radii (O'Brien 2007). To account

346 for this, we excluded these responses and scales from interpretation. As all CIs of these cases

347 included zero, the overall interpretation of results was not affected. The three excluded cases

348 were: all enemies, 'gd, crop' enemies in models with % arable at 100 m radius, and 'wind,

349 non crop' enemies in models with % SNH at 3 km radius. In these cases, variance inflation

350   factors (VIF) were above 10 and exceeded the mean VIF across radii for these responses by

351   >300% (they were respectively 51, 101, and 75).

352   Models including polynomial terms may potentially produce artefacts at the boundaries of

353   observed values of explanatory variables. To confirm the robustness of results obtained with

354   polynomial models, we ran additional analyses using generalized additive mixed modelling

355   (GAMM, R package mgcv v.1.8-24; Wood 2017). We present the detailed methods and

356   results of these analyses in Appendix S4. GAMM analyses led to similar results to those

357   presented in the main text, thus confirming overall conclusions (see Appendix S4).

358   Performing within-model contrasts by testing interactions of functional groups or response

359   syndromes with landscape terms is the only way to formally test differences in effects

360   between groups (Brown *et al.* 2014; Forstmeier *et al.* 2016). Here, due to the imbalanced

361   nature of the contrasts (different groups sampled in different studies) and to avoid model

362   overcomplexity (Brown *et al.* 2014), we do not include such interactions in models and

363   instead consider functional groups as separate response variables. In results, no effects found

364   for a certain group thus indicate only that, and they do not indicate a significant difference

365   from a group where effects were found.

366   To examine whether landscape predictors had significantly higher explanatory power when

367   applied to trait syndromes within functional groups compared to broad groups of arthropods,

368   we compared the marginal R^2 of models relating to trait syndromes (n=132) to those of

369   models relating to pollinators, enemies and pests (n=36). We tested significance of the

370   difference by applying a Wilcoxon rank sum test for non-parametric data on the marginal R^2

371   values obtained at each scale for each response variable.

372

Model fit and interpretation

Interpretation of model terms

375   Quadratic terms were included in models to account for expected hump shapes of the effect

376   of landscape predictors across full gradients (see hypotheses in Introduction and Fig 1d). In

377   addition, we expect the non-linearity of effects (e.g. decreases of abundance and/or services

378   at high values of the predictors) to depend on interactions between the landscape variables

379   (shaded grey area in Fig.1d). For example, decreases in abundance at high amounts of semi-

380   natural or arable habitat, caused by a lack of complementary crop and non-crop resources,

381   may be lifted when edge density between patches is high, increasing spillover. In models, this

382   is formalized by including interactions between quadratic and linear effects of composition

383   and configuration variables (ED:SNH^2 and SNH:ED^2, Fig. S5-7). A significant

384   ED:SNH^2 term indicates that a quadratic effect of % SNH is present at some values of ED,

385   but changes or disappears at other values of ED. Conversely, the effect of ED depends on the

386   value of SNH^2 (e.g. it can be positive at low SNH^2 = low or high % SNH, but not at high

387   SNH^2 = intermediate SNH). Following the principle of marginality, lower order terms (e.g.

388   ED and SNH) are only interpreted as main effects if they are not involved in significant

389   higher-order interactions (Nelder 1994; Hao & Zhang 2017). The direction of effects in the

390   presence of significant interactions is shown in Figs. 2-4 and can be interpreted from model

391   coefficients (Figs. S5-7) as follows:

392   Full models (excluding error and random terms) are coded as:

393    $$Y = b_0 + b_1 SNH + b_2 SNH^2 + b_3 ED + b_4 ED^2 + b_5 ED \times SNH + b_6 SNH \times ED^2 + b_7 ED \times SNH^2 \quad (1)$$

394   Where Y is the response variable, $b_0 \ldots b_7$ are the estimates of each term, and SNH and ED

395   are mean-centered %SNH and edge density, respectively. For any given value of SNH, Y is a

396   simple function of ED and ED^2 (and vice versa for any given value of ED), following:

397    $$Y = a_0 + a_1 ED + a_2 ED^2 \quad \text{where} \quad \begin{cases} a_0 = b_0 + b_1 SNH + b_2 SNH^2 \\ a_1 = b_3 + b_5 SNH + b_7 SNH^2 \\ a_2 = b_4 + b_6 SNH \end{cases} \quad (2)$$

398    $$Y = a'_0 + a'_1 SNH + a'_2 SNH^2 \quad \text{where} \quad \begin{cases} a'_0 = b_0 + b_3 ED + b_4 ED^2 \\ a'_1 = b_1 + b_5 ED + b_6 ED^2 \\ a'_2 = b_2 + b_7 ED \end{cases} \quad (3)$$

399   In particular, if either covariate is at its mean (mean-centered SNH = 0 or mean-centered ED

400   = 0, respectively), then:

401    $$Y = b_0 + b_3 ED + b_4 ED^2 \qquad SNH = 0 \qquad (2a)$$

402    $$Y = b_0 + b_1 SNH + b_2 SNH^2 \qquad ED = 0 \qquad (3a)$$

403   In equations (2) and (3), the parameters a_1 and a'_1 determine the slope of quadratic effects at

404   mean predictor values. The parameters a_2 and a'_2 determine the direction and degree of

405   curvature of quadratic effects: positive values indicate a convex curve ('bowl-shaped'), and

406   negative values a concave one ('hump-shaped'). For example, for all arthropods at scale 3:

407   both b_3 and b_4 are non-significant (i.e., not significantly different from zero; Fig. S5), thus,

408   when % SNH is at its mean, Y is not significantly different from a constant and the

409   abundance of all arthropods does not vary significantly according to edge density (eq. 2a).

410   However, when %SNH is higher than the mean (mean-centered SNH >0), then:

411    $$Y = a_0 + a_1 ED + a_2 ED^2 \quad \text{where} \quad \begin{cases} a_1 \approx 2SNH + 1.8SNH^2 > 0 \\ a_2 \approx 0 \end{cases}$$

412   Thus $Y = a_0 + a_1 ED + a_2 ED^2$ where $a_1 > 0$ and $a_2 \approx 0$, indicating a significant positive

413   (linear) effect of edge density when %SNH is higher than the mean. Similarly, all arthropods

414   do not vary significantly with %SNH when edge density is at its mean (b_1 and b_2 non-

415   significant; eq. 3a). However, when edge density is higher than the mean (mean-centered ED

416   >0), then:

417    $$Y = a'_0 + a'_1 SNH + a'_2 SNH^2 \quad \text{where} \quad \begin{cases} a'_1 \approx 2ED > 0 \\ a'_2 \approx 1.6ED > 0 \end{cases}$$

418   Thus $Y = a'_0 + a'_1 SNH + a'_2 SNH^2$ where a'_1 and a'_2 are positive, indicating a convex effect of

419   %SNH on all arthropods, with positive slope when % SNH is at the mean.

420

421   *Model cross-validation*

422   Full models were fit for all responses and scales. To verify the absence of overfitting, we

423   used a cross-validation approach testing whether prediction error of the models increases

424   when applied to new data (Hastie *et al.* 2009). We partitioned the data for each response into

425   'training' and 'test' datasets (80 vs. 20% of the data, respectively) with 200 permutations,

426   respecting the grouping structure of the data (all observations from one study were kept in

427   one partition) using R package groupdata2 v.1.0.0 (Olsen 2017). We used 4-fold cross-

428   validation (respecting data structure by study and method) to obtain Root Mean Square Errors

429   (RMSE) of predictions on the training datasets, and compared these to the RMSE of

430   predictions on the 'test' datasets using one-sample t-tests of the difference in RMSE between

431   test and training data (ΔRMSE). Hereby, a significant positive difference in RMSE would

432   indicate that the prediction error is higher when applying predictions to new data, and would

433   thus suggest that overfitting is an issue.

434   For all responses and scales, the prediction error on test datasets was not significantly higher

435   than on training datasets (all p>0.1). Overall, differences in error between test and training

436   data were non-significant for all response variables, indicating that the full models including

437   quadratic terms did not overfit the data. In two cases ('All pests' with %SNH at scale 6:

438   ΔRMSE = 9, and 'Pest control' with % arable at scale 6: ΔRMSE = 2.4), differences in error

439   were higher than the maximum difference of 1.5 for all other response variables. However,

440   excluding these models does not affect the interpretation of results.

441

Supplementary references

443   Alford, D.V. (2014). *Pests of fruit crops: a colour handbook*. Crc Press.

444   Ascher, J.S. & Pickering, J. (2018). *Discover Life bee species guide and world checklist (Hymenoptera:*
445   *Apoidea: Anthophila)*. Available at:
446   http://www.discoverlife.org/mp/20q?guide=Apoidea_species. Last accessed 13 November
447   2018.

448   Baayen, H., Bates, D., Kliegl, R. & Vasishth, S. (2015). RePsychLing: Data sets from Psychology and
449   Linguistics experiments. R package version 0.0.4.

450   Barr, D.J., Levy, R., Scheepers, C. & Tily, H.J. (2013). Random effects structure for confirmatory
451   hypothesis testing: Keep it maximal. *J. Mem. Lang.*, 68, 255–278.

452   Bates, D., Kliegl, R., Vasishth, S. & Baayen, H. (2015). Parsimonious Mixed Models. *ArXiv150604967*
453   *Stat*.

454   Bees, Wasps & Ants Recording Society. (2018). *Bees, Wasps & Ants Recording Society*. Available at:
455   http://www.bwars.com. Last accessed 6 July 2018.

456   Bernhardt-Römermann, M., Römermann, C., Nuske, R., Parth, A., Klotz, S., Schmidt, W., *et al.* (2008).
457   On the identification of the most suitable traits for plant functional trait analyses. *Oikos*, 117,
458   1533–1541.

459   Boevé, J.L. (2008). Sawflies (Hymenoptera: Tenthredinidae). In: *Encyclopedia of Entomology*. {ed.
460   Capinera, J.L.} Springer, Dordrecht. pp. 3252–3257.

461   Brodeur, J. & McNeil, J.N. (1990). Overwintering microhabitat selection by an endoparasitoid
462   (Hymenoptera: Aphidiidae): induced phototactic and thigmokinetic responses in dying hosts.
463   *J. Insect Behav.,* 3, 751–763.

464   Brown, A.M., Warton, D.I., Andrew, N.R., Binns, M., Cassis, G. & Gibb, H. (2014). The fourth-corner
465   solution–using predictive models to understand how species traits interact with the
466   environment. *Methods Ecol. Evol.*, 5, 344–352.

467   Burnham, K.P. & Anderson, D.R. (2002). *Model selection and multimodel inference: a practical*
468   *information-theoretic approach*. Springer, New York.

469   Cassis, G. & Schuh, R. t. (2011). Systematics, Biodiversity, Biogeography, and Host Associations of
470   the Miridae (Insecta: Hemiptera: Heteroptera: Cimicomorpha*). Annu. Rev. Entomol.*, 57, 377–
471   404.

472   Chamberlain, S., Szocs, E., Boettiger, C., Ram, K., Bartomeus, I., Baumgartner, J., *et al.* (2016). taxize:
473   Taxonomic information from around the web. R package version 0.7.8.

474   Chamberlain, S.A. & Szöcs, E. (2013). taxize: taxonomic search and retrieval in R. *F1000Research*, 2.

475   Clough, Y., Ekroos, J., Báldi, A., Batáry, P., Bommarco, R., Gross, N., *et al.* (2014). Density of insect-
476   pollinated grassland plants decreases with increasing surrounding land-use intensity. *Ecol.*
477   *Lett.*, 17, 1168–1177.

478   Clough, Y., Kruess, A. & Tscharntke, T. (2007). Organic versus conventional arable farming systems:
479   functional grouping helps understand staphylinid response. *Agric. Ecosyst. Environ.*, 118, 285–
480   290.

481   Entling, M.H., Stämpfli, K. & Ovaskainen, O. (2011). Increased propensity for aerial dispersal in
482   disturbed habitats due to intraspecific variation and species turnover. *Oikos*, 120, 1099–1109.

483   ESRI. (2011). *ArcGIS Desktop: Release 10*. Environmental Systems Research Institute, Redlands, CA.

484   Forstmeier, W., Wagenmakers, E.-J. & Parker, T.H. (2016). Detecting and avoiding likely false-positive
485   findings – a practical guide. *Biol. Rev.*, 92, 1941–1968.

486   Frank, T. & Reichhart, B. (2004). Staphylinidae and Carabidae overwintering in wheat and sown
487   wildflower areas of different age. *Bull. Entomol. Res.*, 94, 209–217.

488 Gagic, V., Bartomeus, I., Jonsson, T., Taylor, A., Winqvist, C., Fischer, C., et al. (2015). Functional
489   identity and diversity of animals predict ecosystem functioning better than species-based
490   indices. *Proc. R. Soc. Lond. B Biol. Sci.*, 282, 20142620.
491 Gardiner, M.M., Landis, D.A., Gratton, C., DiFonzo, C.D., O'Neal, M., Chacon, J.M., *et al.* (2009).
492   Landscape diversity enhances biological control of an introduced crop pest in the north-central
493   USA. *Ecol. Appl.*, 19, 143–154.
494 Gossner, M.M., Simons, N.K., Achtziger, R., Blick, T., Dorow, W.H., Dziock, F., *et al.* (2015). A
495   summary of eight traits of Coleoptera, Hemiptera, Orthoptera and Araneae, occurring in
496   grasslands in Germany. *Sci. Data*, 2, 150013.
497 Hao, N. & Zhang, H.H. (2017). A note on high-dimensional linear regression with interactions. *Am.*
498   *Stat.*, 71, 291–297.
499 Hainmueller, J., Mummolo, J. & Xu, Y. (2018). *How Much Should We Trust Estimates from*
500   *Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice* (SSRN
501   Scholarly Paper No. ID 2739221). Social Science Research Network, Rochester, NY.
502 Hastie, T., Tibshirani, R. & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining,
503   Inference, and Prediction, Second Edition. Springer Series in Statistics. 2nd edn. Springer-
504   Verlag, New York.
505 Homburg, K., Homburg, N., Schäfer, F., Schuldt, A. & Assmann, T. (2014). Carabids. org–a dynamic
506   online database of ground beetle species traits (Coleoptera, Carabidae). *Insect Conserv. Divers.*,
507   7, 195–205.
508 Kleyer, M., Dray, S., Bello, F., Lepš, J., Pakeman, R.J., Strauss, B., *et al.* (2012). Assessing species and
509   community functional responses to environmental gradients: which multivariate methods? *J.*
510   *Veg. Sci.*, 23, 805–821.
511 Kremen, C. & M'Gonigle, L.K. (2015). Small-scale restoration in intensive agricultural landscapes
512   supports more specialized and less mobile pollinator species. *J. Appl. Ecol.*, 52, 602–610.
513 Lavorel, S., McIntyre, S., Landsberg, J. & Forbes, T.D.A. (1997). Plant functional classifications: from
514   general groups to specific groups based on response to disturbance. *Trends Ecol. Evol.*, 12,
515   474–478.
516 Lavorel, S. & Garnier, E. (2002). Predicting changes in community composition and ecosystem
517   functioning from plant traits: revisiting the Holy Grail. *Funct. Ecol.*, 16, 545–556.
518 Leather, S.R. (1993). Overwintering in six arable aphid pests: a review with particular relevance to pest
519   management. *J. Appl. Entomol.*, 116, 217–233.
520 Ludwig, M., Schlinkert, H. & Meyhöfer, R. (2018). Wind-modulated landscape effects on colonization
521   of Brussels sprouts by insect pests and their syrphid antagonists. *Agric. For. Entomol.*, 20, 141–
522   149.
523 Mahr, T. (2017). polypoly: Helper Functions for Orthogonal Polynomials. R package version 0.0.2.
524   https://CRAN.R-project.org/package=polypoly.
525 McGarigal, K., Cushman, S.A., Neel, M.C. & Ene, E. (2002). *Fragstats: Spatial Pattern Analysis*
526   *Program   for   Categorical   Maps.*   Available   at:
527   http://www.umass.edu/landeco/research/fragstats/fragstats.html. Last accessed 3 July 2018.
528 Mestre, L., Hetz, J., Kolb, S., Pfister, S.C., Amato, M., Sutter, L., Jeanneret, P., Albrecht, M., Schirmel,
529   J. & Entling, M.H. (2018). Both woody and herbaceous semi-natural habitats are essential for
530   spider overwintering in European farmland. *Agric, Ecosyst. Environ.*, 267, 141-146.
531 Milne, A. & Laughlin, R. (1956). Biology and ecology of the garden chafer, Phyllopertha horticola (L.).
532   I.—The adult and egg production. *Bull. Entomol. Res.*, 47, 7–22.
533 Nakagawa, S. & Schielzeth, H. (2013). A general and simple method for obtaining R2 from
534   generalized linear mixed-effects models. *Methods Ecol. Evol.*, 4, 133–142.
535 NCBI. (2018). *National Center for Biotechnology Information, US National Library of Medicine.*
536   Available at: https://www.ncbi.nlm.nih.gov/. Last accessed 22 March 2018.
537 Nelder, J.A. (1994). The statistics of linear models: back to basics. *Stat. Comput.*, 4, 221–234.
538 Newbold, T., Hudson, L.N., Hill, S.L.L., Contu, S., Lysenko, I., Senior, R.A., *et al.* (2015). Global
539   effects of land use on local terrestrial biodiversity. *Nature*, 520, 45–50.
540 Nickel, H. & Remane, R. (2002). Check list of the planthoppers and leafhoppers of Germany, with notes
541   on food plants, diet width, life cycles, geographic range and conservation status (Hemiptera,
542   Fulgoromorpha and Cicadomorpha). *Beitr. Zur Zikadenkunde*, 5, 27–64.

543 Normandin, É., Vereecken, N.J., Buddle, C.M. & Fournier, V. (2017). Taxonomic and functional trait
544   diversity of wild bees in different urban settings. *PeerJ*, 5, e3051.
545 Öberg, S., Mayr, S. & Dauber, J. (2008). Landscape effects on recolonisation patterns of spiders in
546   arable fields. *Agric. Ecosyst. Environ.*, 123, 211–218.
547 O'Brien, R.M. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Qual.*
548   *Quant.*, 41, 673–690.
549 Olsen, L.R. (2017). groupdata2: Creating Groups from Data. R package version 1.0.0.
550 O'Neill, K.M. (2001). *Solitary Wasps: Behavior and Natural History*. Cornell University Press.
551 Pekár, S. & Toft, S. (2015). Trophic specialisation in a predatory group: the case of prey-specialised
552   spiders (Araneae). *Biol. Rev.*, 90, 744–761.
553 Pfiffner, L. & Luka, H. (2000). Overwintering of arthropods in soils of arable fields and adjacent semi-
554   natural habitats. *Agric. Ecosyst. Environ.*, 78, 215–222.
555 Raymond, L., Sarthou, J.-P., Plantegenest, M., Gauffre, B., Ladet, S. & Vialatte, A. (2014). Immature
556   hoverflies overwinter in cultivated fields and may significantly control aphid populations in
557   autumn. *Agric. Ecosyst. Environ.*, 185, 99–105.
558 Rusch, A., Bommarco, R., Jonsson, M., Smith, H.G. & Ekbom, B. (2013). Flow and stability of natural
559   pest control services depend on complexity and crop rotation at the landscape scale. *J. Appl.*
560   *Ecol.*, 50, 345–354.
561 Sarthou, J.-P., Badoz, A., Vaissière, B., Chevallier, A. & Rusch, A. (2014). Local more than landscape
562   parameters structure natural enemy communities during their overwintering in semi-natural
563   habitats. *Agric. Ecosyst. Environ.*, 194, 17–28.
564 Sawyer, A.J. & Haynes, D.L. (1986). Cereal leaf beetle spatial dynamics: simulations with a random
565   diffusion model. *Ecol. Model.*, 33, 89–99.
566 Skellern, M.P., Welham, S.J., Watts, N.P. & Cook, S.M. (2017). Meteorological and landscape
567   influences on pollen beetle immigration into oilseed rape crops. *Agric. Ecosyst. Environ.*, 241,
568   150–159.
569 Speight, M.C.D., Monteil, C., Castella, E. & Sarthou, J.P. (2010). StN_2010. *Syrph Net Publ. Dublin*.
570 Stireman, J.O., O'Hara, J.E. & Wood, D.M. (2006). TACHINIDAE: Evolution, Behavior, and Ecology.
571   *Annu. Rev. Entomol.*, 51, 525–555.
572 Sutter, L., Amato, M., Jeanneret, P. & Albrecht, M. (2018). Overwintering of pollen beetles and their
573   predators in oilseed rape and semi-natural habitats. Agric. Ecosyst. Environ., 265, 275–281.
574 Tolman, T. & Lewington, R. (2008). *Collins Butterfly Guide-the most comprehensive guide to the*
575   *butterflies of Britain and Europe*. HarperCollins Publishers, London.
576 Vázquez, X.A. (2002). *European Fauna of Oedemeridae: Coleoptera*. Argania editio.
577 Watanabe, H., Katayama, N., Yano, E., Sugiyama, R., Nishikawa, S., Endou, T., et al. (2014). Effects
578   of aphid honeydew sugars on the longevity and fecundity of the aphidophagous gall midge
579   Aphidoletes aphidimyza. *Biol. Control*, 78, 55–60.
580 Westbrook, J.K., Spurgeon, D.W., Eyster, R.S. & Schleider, P.G. (2000). Wind-directed dispersal of
581   boll weevils, Anthonomus grandis (Boh.). In: Proceedings Beltwide Cotton Conferences, San
582   Antonio, USA, 4-8 January, 2000: Volume 2. National Cotton Council, pp. 1261–1264.
583 Wheeler, A.G. (2001). *Biology of the Plant Bugs (Hemiptera: Miridae): Pests, Predators, Opportunists*.
584   Cornell University Press.
585 Wood, S.N. (2017). *Generalized additive models: an introduction with R (2nd edition)*. Chapman and
586   Hall/CRC.
587 World Spider Catalog. (2018). *World Spider Catalog. Natural History Museum Bern, version 19.0*.
588   Available at http://wsc.nmbe.ch. Last accessed 30 November 2018.