Appendix S4. Results of generalized additive mixed models (GAMM) for effects of landscape variables on arthropod abundance and ecosystem services.

Supporting Information to Martin, E. A. et al. The interplay of landscape composition and configuration: new pathways to manage functional biodiversity and agro-ecosystem services across Europe.

In order to assess the robustness of results to the statistical method applied, we ran an additional set of analyses examining effects of landscape variables at all spatial scales on arthropod abundance and ecosystem services using generalized additive mixed models (GAMM) instead of models with polynomial terms (see Methods and Appendix S1). For this, we used ti() tensors for the main explanatory variables % SNH (alternatively % arable land), edge density, and their interaction using function gamm() in R package mgcv v.1.8-24 (Wood 2017), with 'select=TRUE' to penalize smooths to zero and maximum likelihood estimation. Random effects for each response variable were selected identically to polynomial analyses as described in Appendix S1. For each term in the models, we set an initial value of 5 for the smoothness parameter 'k'. We then incrementally increased 'k' for each term and refit the model until effective degrees of freedom were stabilized (function choose.k() in R package mgcv). 'K' values used in final models were on average 9.3 ± 1.4 (mean ± SD) for composition variables (% SNH and % arable land), 8.7 ± 0.9 for edge density and 12.8 ± 2.4 for their interaction. Similarly to polynomial analyses, we present the results of full models (penalized by select=TRUE) to avoid parameter bias in the context of inference (not prediction). We ran GAMM models at all successive radii around sites (0.1, 0.25., 0.5, 1, 2, 3 km) and present effective degrees of freedom and p-values of terms for all radii in Appendix S4 Figs. 1-3. In contrast to quadratic models, confidence intervals of effects (regression estimates) are not obtainable in the GAMM context, and confidence intervals of predictions can only be visualized using 2-dimensional plots of predictions for each model (as opposed to 3-dimensional heatmaps). We thus here restrict the presentation of model outcomes at all radii to effective degrees of freedom and p-values of each term (Appendix S4 Figs. 1-3). However, we caution that p-values obtained with GAMM are explicitly approximate and mainly represent significance of the smoothness parameter. In Appendix S4 Figs. 1-3, effective degrees of freedom >1 indicate non-linearity of effects (Wood 2006). We present detailed results for each response variable at 1 km radius in Appendix S4 Figs. 4-6.

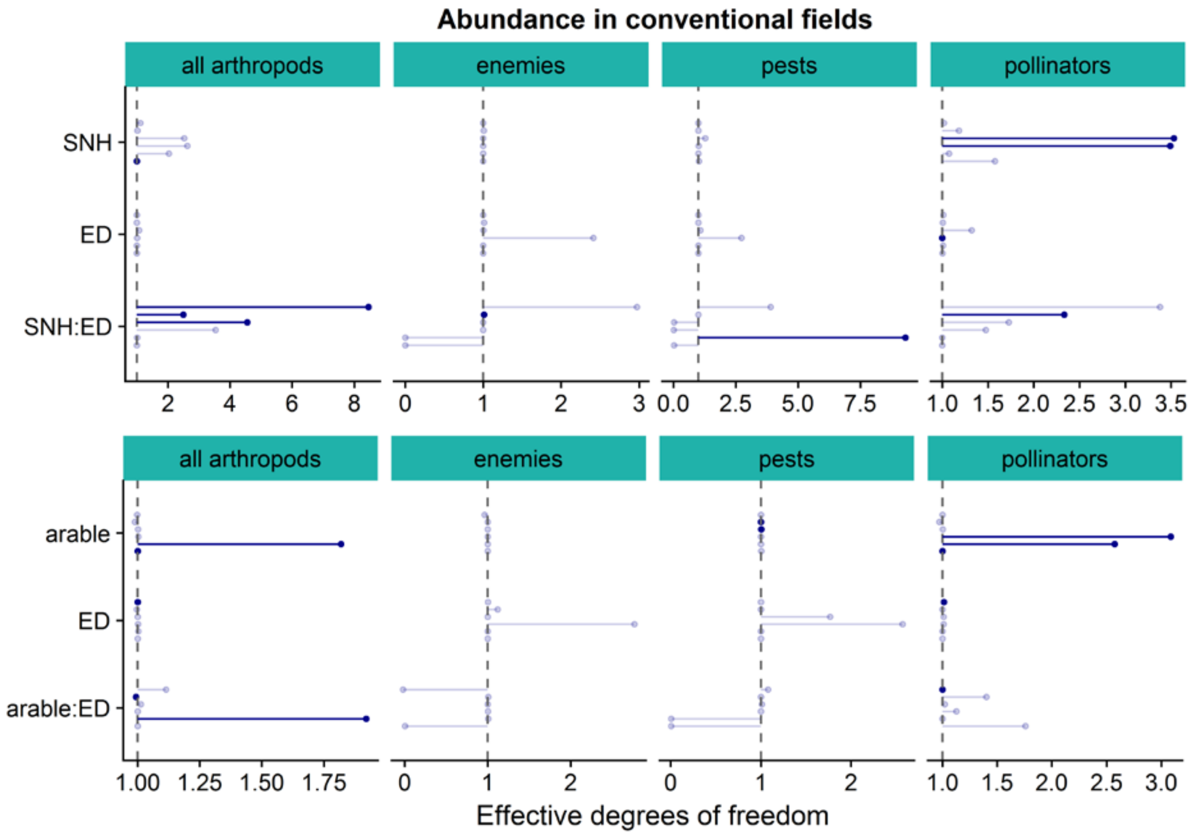
Results of GAMM analyses show strongly consistent results with polynomial analyses (Figs. 2-4, S5-S7), thus confirming the main patterns observed. In particular, we find highest abundances of all arthropods and pollinators at high values of both % SNH and ED (Appendix S4 Figs. 1, 4). Especially for pollinators, smaller increases are associated with ED at low values of % SNH, highlighting the non-linear (convex) effect of % SNH on pollinator abundance (Appendix S4 Fig. 1). One difference between results of GAMM and polynomial models is in the impact of edge density on pollinators. In Fig. S4 (polynomial model), interactions of % SNH and ED are significant at 2 radii, but a positive main effect of edge density at 1 km radius is found instead in Appendix S4 Fig. 1 (GAMM model).

Functional trait syndromes of enemies, pollinators and pests (Appendix S4 Figs. 2, 5) and ecosystem services of pest control, pollination and yields (Appendix S4 Figs. 3, 6) also show highly consistent results with polynomial analyses, with some differences in the approximate significance levels of the GAMM. We highlight for example the presence of regions of high yields per unit area throughout the gradients of edge density and % SNH (Appendix S4 Fig. 6). Due to the more complex interpretation of results at all radii using GAMM (effective degrees of freedom give no indication of the direction of effects in Appendix S4 Figs. 1-3), we base the discussion of results on polynomial analyses shown in Figs. 2-4, S5-S7 (see Methods).

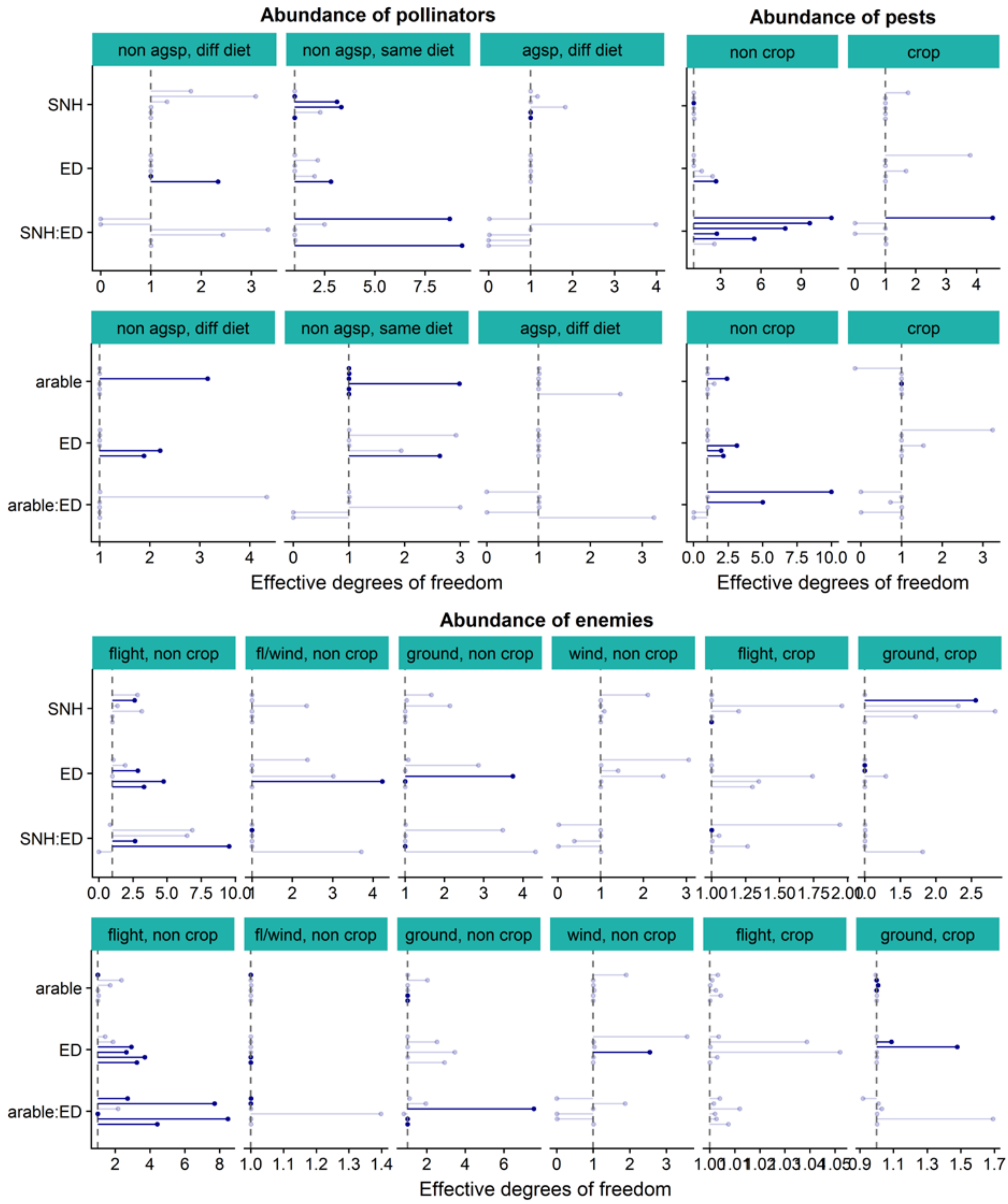
Supplementary references

Wood, S.N. (2006). *Generalized additive models: an introduction with R*. Chapman and Hall/CRC.

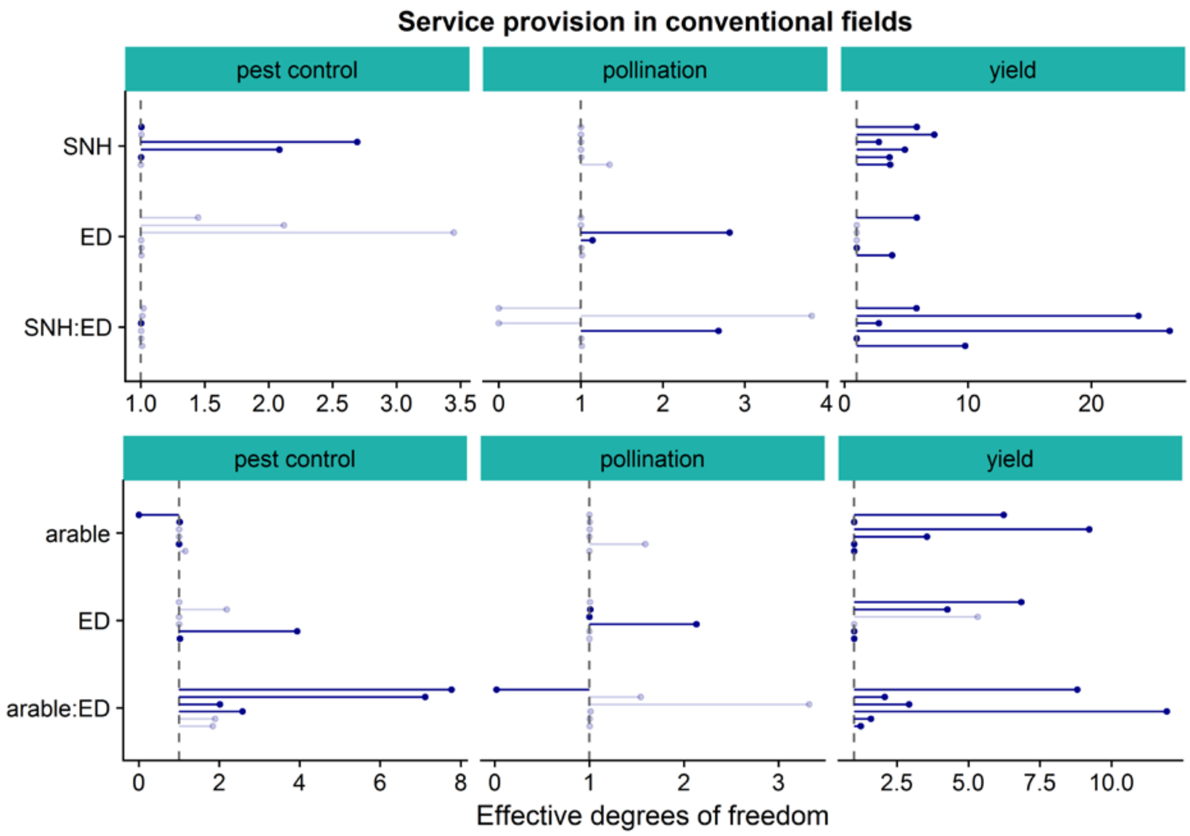
Wood, S.N. (2017). *Generalized additive models: an introduction with R (2nd edition)*. Chapman and Hall/CRC.



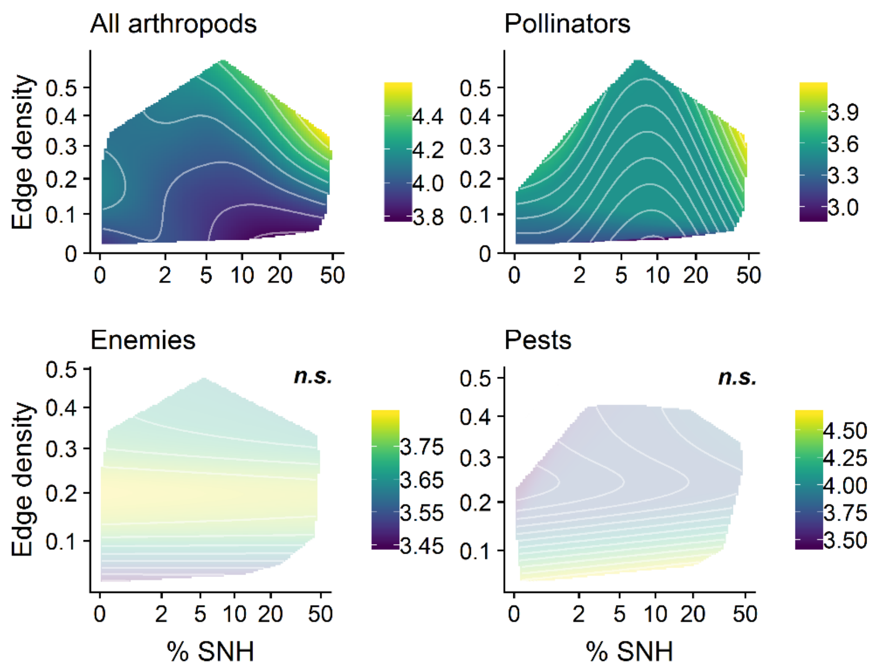
Appendix S4 Figure 1. Results of GAMM models explaining the abundance of all arthropods and subsets of enemies, pollinators and pests as a function of landscape composition (% SNH, top panel; % arable land, bottom panel) and configuration variables (ED, edge density of crop/crop and crop/non-crop boundaries) in conventional fields. Models were run at six successive spatial scales (0.1, 0.25, 0.5, 1, 2 and 3 km radii around sites) shown from top to bottom for each term. Radii at which *approximate* p-values were <0.05 are highlighted. Effective degrees of freedom are given for each response variable and radius. Effective degrees of freedom (edf) >1 indicate non-linearity of terms (we note that edf values give no indication of the direction of effects, in contrast to effect sizes of polynomial analyses; see Fig. S5). Separate models were run with either % SNH or % arable as composition variable. Intercept estimates are not shown. Only effects for which approximate p-values were <0.05 at more than one radius are considered for interpretation. See detailed plots of results at 1 km radius in Appendix S4 Fig. 4.



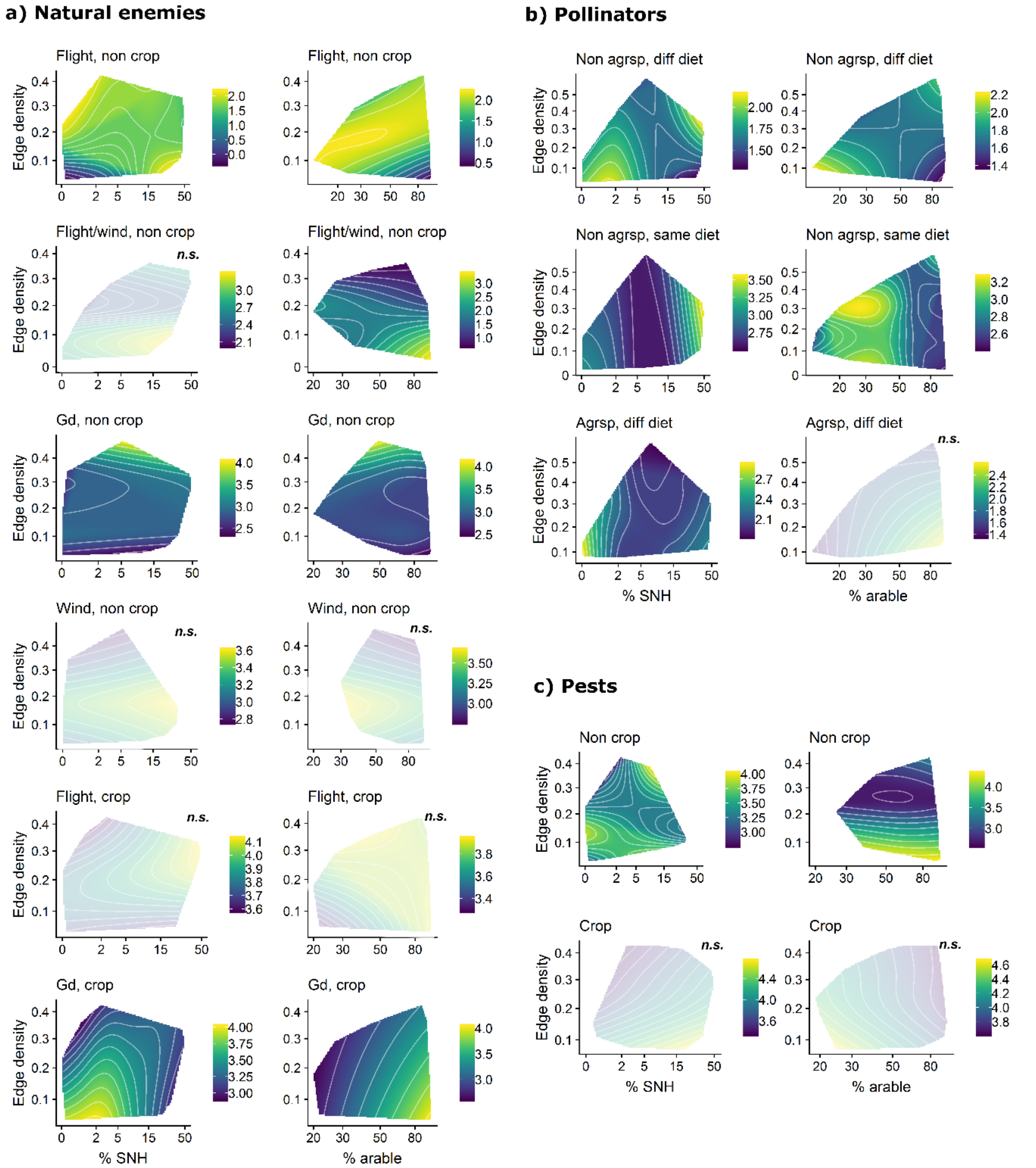
Appendix S4 Figure 2. Results of GAMM models explaining the abundance of functional trait syndromes of enemies, pollinators and pests as a function of landscape composition (% SNH, top panels; % arable land, bottom panels) and configuration variables (ED: edge density of crop/crop and crop/non-crop boundaries) in conventional fields. Trait syndromes were determined for each functional group by cluster regression of categorical traits. Models were run at six successive spatial scales (0.1, 0.25, 0.5, 1, 2 and 3 km radii around sites) shown from top to bottom for each term. Radii at which *approximate* p-values were <0.05 are highlighted. Effective degrees of freedom are given for each response variable and radius. Effective degrees of freedom >1 indicate non-linearity of terms (we note that edf values give no indication of the direction of effects, in contrast to effect sizes of polynomial analyses; see Fig. S6). Separate models were run with either % SNH or % arable as composition variable. Intercept estimates are not shown. Only effects for which approximate p-values were <0.05 at more than one radius are considered for interpretation. See detailed plots of results at 1 km radius in Appendix S4 Fig. 5.



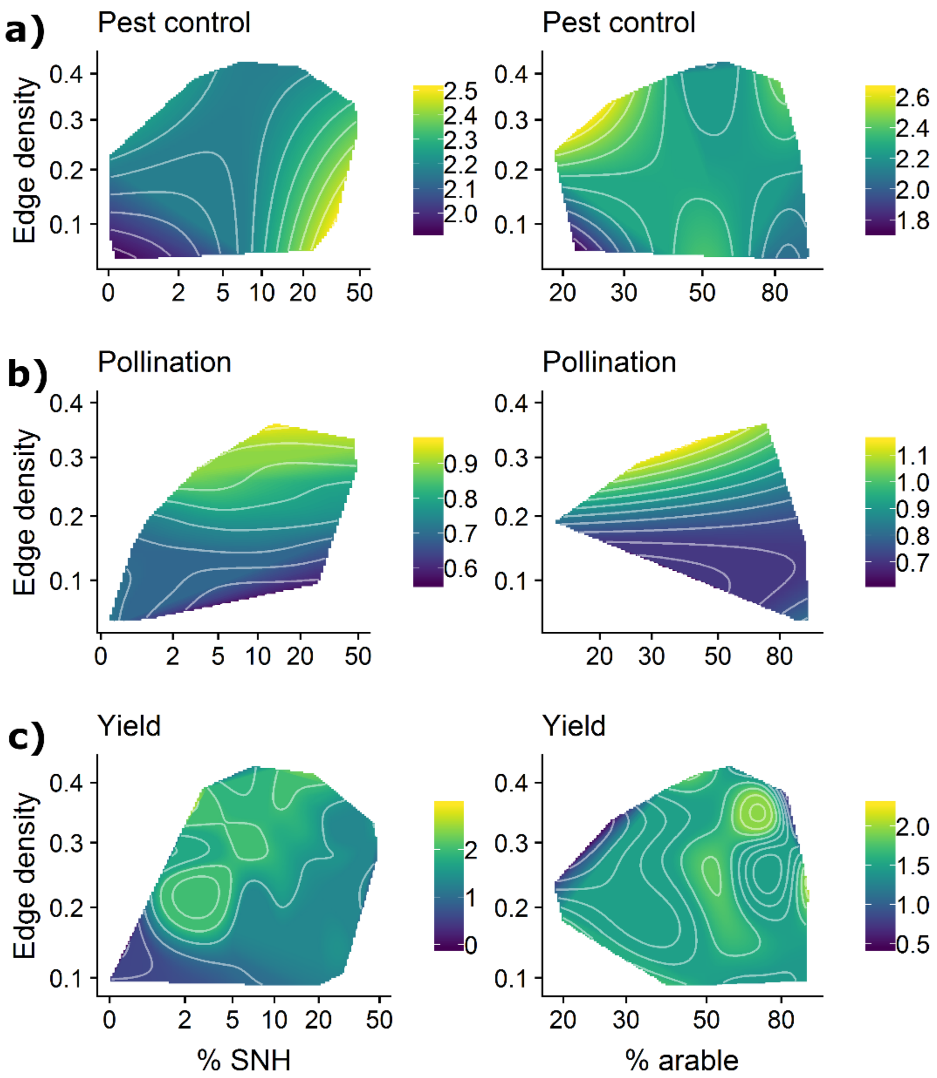
Appendix S4 Figure 3. Results of GAMM models explaining the provision of pest control, pollination and yields as a function of landscape composition (% SNH, top panels; % arable land, bottom panels) and configuration variables (ED: edge density of crop/crop and crop/non-crop boundaries) in conventional fields. Models were run at six successive spatial scales (0.1, 0.25, 0.5, 1, 2 and 3 km radii around sites) shown from top to bottom for each term. Radii at which *approximate* p-values were <0.05 are highlighted. Effective degrees of freedom are given for each response variable and radius. Effective degrees of freedom >1 indicate non-linearity of terms (we note that edf values give no indication of the direction of effects, in contrast to effect sizes of polynomial analyses; see Fig. S7). Separate models were run with either % SNH or % arable as composition variable. Intercept estimates are not shown. Only effects for which approximate p-values were <0.05 at more than one radius are considered for interpretation. See detailed plots of results at 1 km radius in Appendix S4 Fig. 6.



Appendix S4 Figure 4. Heatmaps of the effects of seminatural habitat amount (SNH; composition variable) and landscape edge density (ED in km/ha; configuration variable) on the abundance of arthropods (top left) and on functional groups of pollinators, natural enemies, and pests, using GAMM instead of polynomial models. Yellow indicates areas of highest abundance, blue areas of lowest abundance (see ln(x+1)-transformed abundance scale at the right of each panel). Effective degrees of freedom of effects are shown for all radii in Appendix S4 Fig. 1. Results at 1 km radius are shown here. Results are marked 'n.s.' and faded if p-values <0.05 were obtained at less than two out of six tested radii. Only the area covering the range of both variables for each response is plotted. Note a log-scale of predictor variables.



Appendix S4 Figure 5. Heatmaps of the effects of landscape composition (% SNH, left columns; % arable land, right columns) and landscape configuration (edge density in km/ha) on the abundance of functional response groups of a) natural enemies, b) pollinators, and c) pests, using GAMM instead of polynomial models. Functional groups were separated into trait syndromes based on cluster regression of six categorical traits (see abbreviations in Table 1; Figs. S2-3). Effective degrees of freedom of effects are shown at all radii in Appendix S4 Fig. 2; results are shown here at the 1 km radius. See further graph details in the legend of Appendix S4 Fig. 4.



Appendix S4 Figure 6. Heatmaps of the effects of landscape composition (% SNH, left columns; % arable land, right columns) and landscape configuration (edge density in km/ha) on a) pest control, b) pollination and c) crop yield in weight per unit area, using GAMM instead of polynomial models. Response variables represent an ecosystem service index accounting for differences in methods within and between studies (see Appendix S1). See Table S3 for detailed units and measurements per study. Blue: lowest service provision; yellow: highest service provision. Effective degrees of freedom of effects are shown at all radii in Appendix S4 Fig. 3; results are shown here at the 1 km radius. See further graph details in the legend of Appendix S4 Fig. 4.