How far is enough? Prediction of the scale of effect for wild bees

James Desaegher, Annie Ouin, David Sheeren

ECOGRAPHY

Research

How far is enough? Prediction of the scale of effect for wild bees

James Desaegher, Annie Ouin and David Sheeren

J. Desaegher (https://orcid.org/0000-0001-9205-0184) ✉ (jamesdesaegher@gmail.com), A. Ouin and D. Sheeren, Univ. de Toulouse, INRAE, UMR DYNAFOR, Castanet-Tolosan, France and LTSER Zone Atelier 'PYRÉNÉES GARONNE', Auzeville-Tolosane, France.

A crucial issue for landscape ecologists is identifying the spatial extents at which a landscape affects species occurrence. Multi-scale analyses are usually conducted to identify the 'scale of effect', that is, the spatial extent associated with the best relationship between landscape variables and species occurrence, which is assumed to be related to species traits. However, few guidelines exist to determine the range of distances to be investigated.

Based on the foraging distances of wild bee species, our main goal was to estimate the maximum distance of effect, that is, the distance beyond which the scale of effect for wild bee species is unlikely to be detected.

Using the InVEST pollination model, we 1) modelled bee categories with distinct foraging distances and identified the scale of effect on their simulated abundance 2) defined an index, noted λ, that estimates the distance beyond which landscape composition has only negligible effects on simulated abundances. We validated our results by identifying the scale of effect on the abundances of 16 bee species collected in southwestern France.

We detected a significant positive relationship between the average foraging distance (α) of the modelled bees and their scale of effect. The λ index was linearly related to the average foraging distances of bees ($\lambda = 5.4\alpha + 253$) and was above the identified scale of effect for the modelled bees. The λ was also found to be above the scale of effect for 93% of the observed bee species.

Our results suggest that the λ index is a good estimator of the upper limit of the scale of effect for wild bees. The λ index could be used to identify the minimum distance between sampling sites before setting up an experiment and the maximum buffer size required in multi-scale analysis to detect the scale of effect.

Keywords: bee, InVEST, landscape, pollination, scale of effect

Introduction

'To understand the drama [in the 'ecological theatre' (Hutchinson 1965)], we must view it on the appropriate *scale*.' (Wiens 1989)

Landscape ecologists aim to study how landscape heterogeneity influences (the processes that affect) the abundance and distribution of organisms (Fahrig 2005). They study the relationships between biological responses collected at different locations (e.g. species abundance), and the environmental variables in the surrounding area (e.g. proportion of habitats). However, choosing the spatial extent at which to measure the landscape variables is crucial but complex (Levin 1992). Indeed, it depends not only on the organism under investigation, and its mobility, but also on habitat fragmentation, matrix quality or topography (Miguet et al. 2016). An inappropriate choice of spatial extent may result in failure to detect the ecological pattern or even in the detection of artefacts (Wiens 1989).

One common way to identify the 'appropriate' spatial extent of investigation related to the behavioural ecology of species is to measure landscape variables at multiple circular extents around sampling sites and then to choose the extent associated with the best statistical relationship between the landscape variables and the biological response (Fig. 1; Jackson and Fahrig 2012). This procedure is called 'multi-scale analysis' and the optimal spatial extent is often called the 'scale of effect' (Jackson and Fahrig 2012, 2014, Miguet et al. 2016). Here, the term 'scale' refers to the geographic scale and is used as a synonym for geographical extent.

Identifying the scale of effect is an essential step before formulating any explanations concerning a species–habitat relationship at landscape scale (Thornton and Fletcher 2014). However, a meta-analysis of the practices of ecologists performing multi-scale analyses revealed that the range of scales explored is too narrow, since 44% of studies reported a scale of effect at the smallest or the largest scale evaluated (Jackson and Fahrig 2014). To our knowledge, only the general guideline resulting from the simulation work of Jackson and Fahrig (2012) is available, which expects the scale of effect to be 4–9 times the median dispersal distance of the organism studied. There is a critical need to improve our ability to predict the scale of effect for a given species, and to validate the predictions with empirical data. Estimating a priori of the range in which the scale of effect would fall into, could help landscape ecologists choose sampling sites located sufficiently far away from one another to reduce the spatial pseudo-replication often identified a posteriori (Miguet et al. 2016). Of course, it would also improve the accuracy of the identification of the scale of effect associated with a species before analysing a species–habitat relationship. Better predictions of the scale of effect would help landscape managers and conservation biologists improve the efficiency of their actions, for example, scaling and designing nature reserves according to the focus organisms (Wiens 1989, Jackson and Fahrig 2012, Miguet et al. 2016, Moll et al. 2020).

An early considered approach to predict the scale of effect of an organism is using allometric relationships related to species mobility (Wiens 1989). However, predicting the scale of effect based on species traits, mainly tested on birds (Thornton and Fletcher 2014), often failed to produce significant results
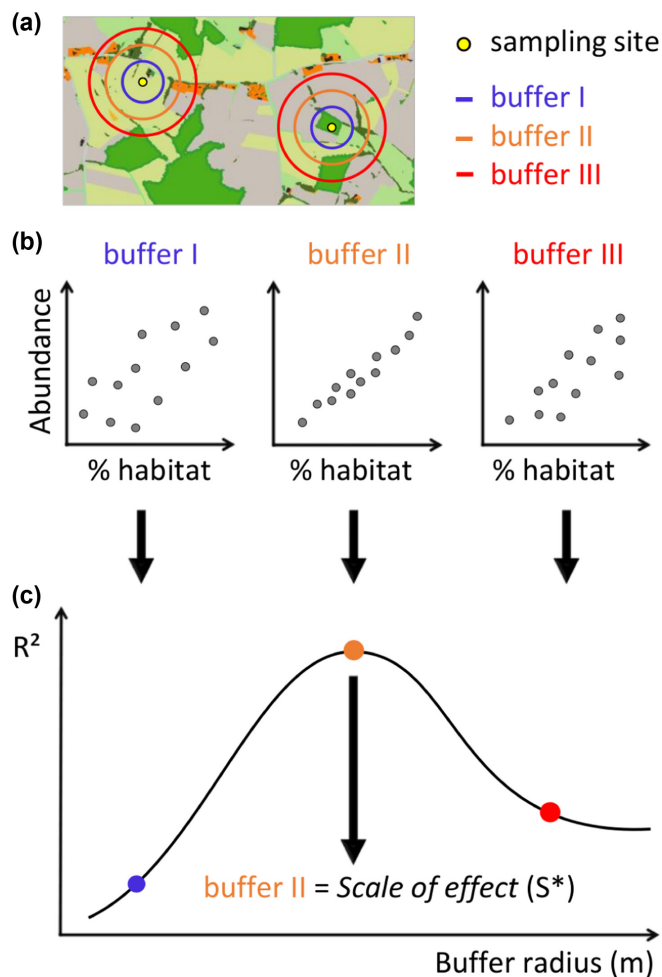

Figure 1. Design of a fictive multi-scale analysis (a) relating the abundance of a bee species to the proportion of a habitat (e.g. grassland) measured for different landscape extents (i.e. buffer radii) centred on the sampling sites (yellow dots). Based on the linear relationships between species abundance and the proportion of habitat calculated within the different landscape extents (b), the scale of effect (S*), is identified as the buffer radius with the best statistical relationship, for instance based on the R^2 (c).

(Stuber et al. 2018, Moll et al. 2020). Despite the ecological and economic importance of pollination (Gallai et al. 2009, Ollerton et al. 2011), quantitative relationships between bee species traits related to foraging distance (Gathmann and Tscharntke 2002, Greenleaf et al. 2007) and their specific scale of effect, have never been identified in wild bees.

One way to predict the scale of effect for a bee species could consist in restricting the upper limit of the range of candidate scales by identifying a distance beyond which the scale of effect for the species is biologically unlikely to be detected based on their estimated foraging distances. A conservative hypothesis would be that the scale of effect must necessarily be less than the distance beyond which any landscape composition (e.g. floral or nesting plots) only has negligible effects

on bee species abundance, hereafter called the maximum distance of landscape effect (abbreviated λ, Fig. 2). The λ distance can be estimated using spatially explicit models that predict pollinator abundances in landscapes.

In this study, we used the InVEST pollination model (Sharp et al. 2018) to predict the relative abundance of bees in a landscape with the aim of 1) defining and estimating the λ distances associated with different modelled bee categories characterised by different average foraging distances through simulations, 2) assessing if the scale of effect for the modelled bees are smaller than their λ distance and exploring the relationships between the scale of effect, average foraging distance and λ distance (Fig. 2) and 3) validating our simulation results by estimating the λ distance and the scale of effect associated with actual ground-nesting bee species collected in south-western France.

Material and methods

Study area

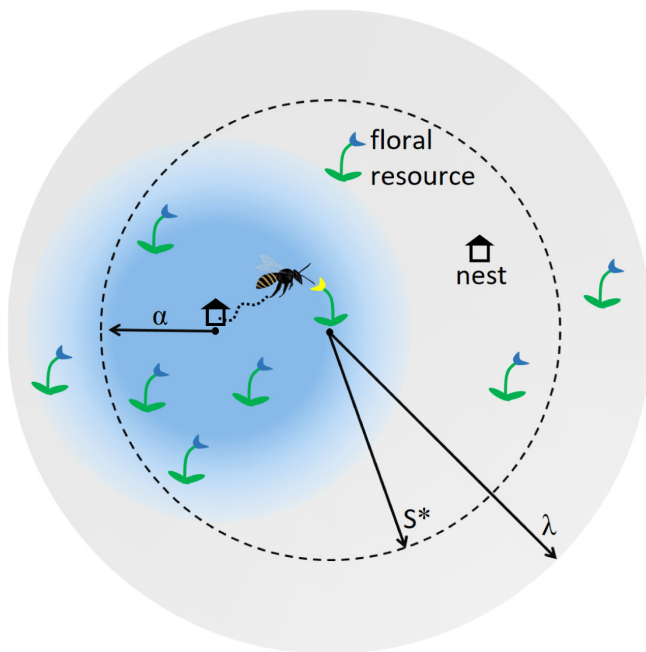The study area is located in south-western France in the Vallées et Coteaux de Gascogne site in the

Figure 2. Illustration of the distinction between the average foraging distance (α), the scale of effect (S*) and the λ distance and their hypothetical relationships. Bees are assumed to locate their nests in sites surrounded by floral resources within their foraging range (blue area). The scale of effect (S*), identified with a multi-scale analysis, is represented by the dashed line. The λ distance represents the distance beyond which landscape composition only has negligible effects on the abundance observed in the sampling site (yellow flower in the centre).

long term socio-ecological research area (LTSER ZA PYGAR) 'Pyrénées Garonne' (Ouin et al. 2022) and covers approximatively 2500 km² (50 × 50 km, Fig. 3). This agricultural area is characterised by hilly landscapes with a traditional cropping system combining woodlots and mixed crop-livestock farming. The majority of grasslands are located on hillsides while annual crops (cereals, oilseed rape and sunflower, maize) are grown in the valleys (Choisis et al. 2010).

The land use/land cover (LULC) map

For the purpose of our simulations and analyses, we produced a LULC map of the study area by applying a series of basic GIS operations to three different spatial datasets available in vector format: 1) the RPG (Registre Parcellaire Graphique) from the Common Agricultural Policy (CAP) declarations (<www.data.gouv.fr>), 2) the BDTOPO (2010) nationwide database, produced by the French national mapping agency (IGN) (<https://geoservices.ign.fr/documentation/diffusion/telechargement-donnees-libres.html>), 3) the OCS-GE (OCcupation du Sol à Grande Échelle) database also obtained from IGN. The three datasets were topologically corrected and combined in a single vector layer to produce a map with the highest possible accuracy. The resulting map was composed of 92 LULC categories with 81 agricultural land use categories (e.g. wheat, oilseed rape, sunflower) and 11 other LULC categories (e.g. buildings, hedgerows, water areas) (Supporting information). The vector layer was rasterized in a 4-m spatial resolution raster to take advantage of a good spatial resolution while keeping the volume of data manageable during the simulations.

The pollination model and parameterization

In this study, we used the InVEST pollination model (Sharp et al. 2018) mainly based on the equations of the Lonsdorf model (Lonsdorf et al. 2009). The model is spatially explicit and predicts the abundance of wild bees through their visitation rate on floral resources.

The visitation rate of a species s in a map pixel p is calculated in two steps. In a first step, the model identify where the species actually nests in the landscape, through the mapping of an index called 'pollinator supply', $Supply_{p,s}$. The pollinator supply in pixel p, depends on 1) the availability of suitable nesting substrate for the species in the pixel, $Nest_{p,s}$, 2) the accessible floral resources in the neighbourhood of pixel p, $AFR_{p,s}$, 3) the pollinator species abundance, $Poll_s$. The accessible floral resources, $AFR_{p,s}$, depends on 1) the average foraging distance of bees, α, 2) the abundance of floral resources in all the surrounding pixels p', during season j, $Flower_{p',j}$ and 3) the activity of species s in season j, $Activity_{s,j}$. In a second step, based on the pollinator supply map, the model calculates the 'visitation rates' to the flowering resources surrounding the original nests.

The model is defined as follows:

$$\text{Supply}_{p,s} = \text{Nest}_{p,s} \times \text{AFR}_{p,s} \times \text{Poll}_s = \text{Nest}_{p,s}$$

$$\times \left(\frac{\sum_{p'} e^{-\frac{D}{\alpha}} \times \sum_j \left(\text{Flower}_{p',j} \times \text{Activity}_{s,j} \right)}{\sum_{p'} e^{-\frac{D}{\alpha}}} \right) \quad (1)$$

$$\times \text{Poll}_s$$

$$\text{Visitation}_{p,j,s} = \left(\frac{\text{Flower}_{p,j} \times \text{Activity}_{s,j}}{\text{AFR}_{p,s}} \right)$$

$$\times \frac{\sum_{p'} \text{Supply}_{p',s} \times e^{-\frac{D}{\alpha}}}{\sum_{p'} e^{-\frac{D}{\alpha}}} \quad (2)$$

where D is the Euclidean distance between pixel p and p'.

The model requires three input data: 1) a LULC map in a raster format (section 'The land use/land cover (LULC) map'), 2) a table of LULC attributes (indices: Nest, Flower$_j$) and 3) a table of pollinator attributes (indices: α, Poll, Activity$_j$).

The table of LULC attributes combines indices ranging from 0 to 1 for the 92 LULC categories concerning 1) the relative availability of nests and 2) the relative availability of floral resources during the seasons or months studied. We decided to study bee species nesting in belowground nests since the majority of wild bee species are ground nesters (Cane 1991, Michener 2007). For each LULC category, the availability of belowground nests was attributed using expert knowledge. Concerning the availability of floral resources, we defined 12 indices associated with the 12 months of the year. For 70% of the LULC categories, the values of the floral resource indices were based on information reported in existing databases containing records of pollen and nectar production of plant taxa (mainly EFSA 2013, Baude et al. 2016). For the remaining LULC categories (30%) the floral resource indices were attributed using expert knowledge (see the Supporting information for details).

Concerning the table of pollinator attributes, we defined 16 modelled bee categories that only differed in their average foraging distance (α-index). The α-index ranged from 50 to 800 m with a 50-m increment. This range was representative of most bee species (Gathmann and Tscharntke 2002, Greenleaf et al. 2007). Since the 16 modelled bees were simulated separately, during one month, and were all belowground nesters, we arbitrarily set to 1, the relative abundance (Poll$_s$), the foraging activity during that month (Activity$_j$) and the preference for belowground nests (Nest$_{p,s}$).

The simulation

Our simulation using InVEST pollination model had two goals: 1) to quantify the distance at which the effects of the landscape on the visitation rate of modelled bees can be considered as negligible, according to the bees' average foraging distance; 2) to predict visitation rates for different sampling sites that can be used to explore the relationship between the scale of effect for the modelled bees and their average foraging distance.

We simulated the visitation rate of the 16 modelled bees in 50 target sunflower fields. We chose sunflower to have 1) a sufficient quantity of target plots for spatial sampling distribution as sunflower is the most frequently cultivated entomophilous crop in the study area, and 2) target plots with the same floral resource abundance, making visitation rates comparable among plots. We randomly selected fields between 10 000 m² and 20 000 m² in size, located at least 5 km from the border of the study area, and separated from each other by at least 4 km (Fig. 3). To assess the distance-weighted effects of landscape composition on the visitation rate of modelled bees, i.e. how the landscape effects decline with increasing distance from the sampling sites, we cropped the LULC input map around the centres of the 50 sunflower fields according to different buffer radii. For each sunflower field, 108 circular extents were defined with a buffer radius ranging from 10 to 90 m (10-m increment) and from 100 to 5000 m (50-m increment).

We ran the InVEST pollination model using Python ver. 2.7.16 programming language iteratively with the 5400 circular input maps (50 fields × 108 buffer radii) and for the 16 modelled bees (average foraging distance ranging from 50 to 800 m). Since sunflowers flower in July in the study area, we simulated the visitation rate in the 50 sunflower fields during July. From the 86 400 output maps of visitation rates (50 fields × 108 buffer radii × 16 modelled bees), we extracted the visitation rate for the pixel located closest to the centroid of the sunflower field (i.e. map centre).

Distance-weighted effects of landscape and λ distances

To assess how the InVEST pollination model weights the landscape effects on the visitation rates of modelled bees with increasing distance from the map centre, we calculated the percentage change in visitation rates according to the radius of the circular input maps.

For each combination of a modelled bee, a field and a buffer radius, we calculated the percentage change in visitation rates as follows:

$$\%\Delta\text{Visitation}(r) = \frac{\left| \text{Visitation}(r) - \text{Visitation}(r_{\max}) \right|}{\text{Visitation}(r_{\max})}$$

where Visitation(r) is the visitation rate for the pixel in the centre of the field with a buffer radius r, and Visitation($r_{\max}$) is the visitation rate for the pixel in the centre of the field with the maximum buffer radius (5000 m).

For each modelled bee (i.e. each α value), we calculated the %ΔVisitation(r) functions for the 50 target fields. Using
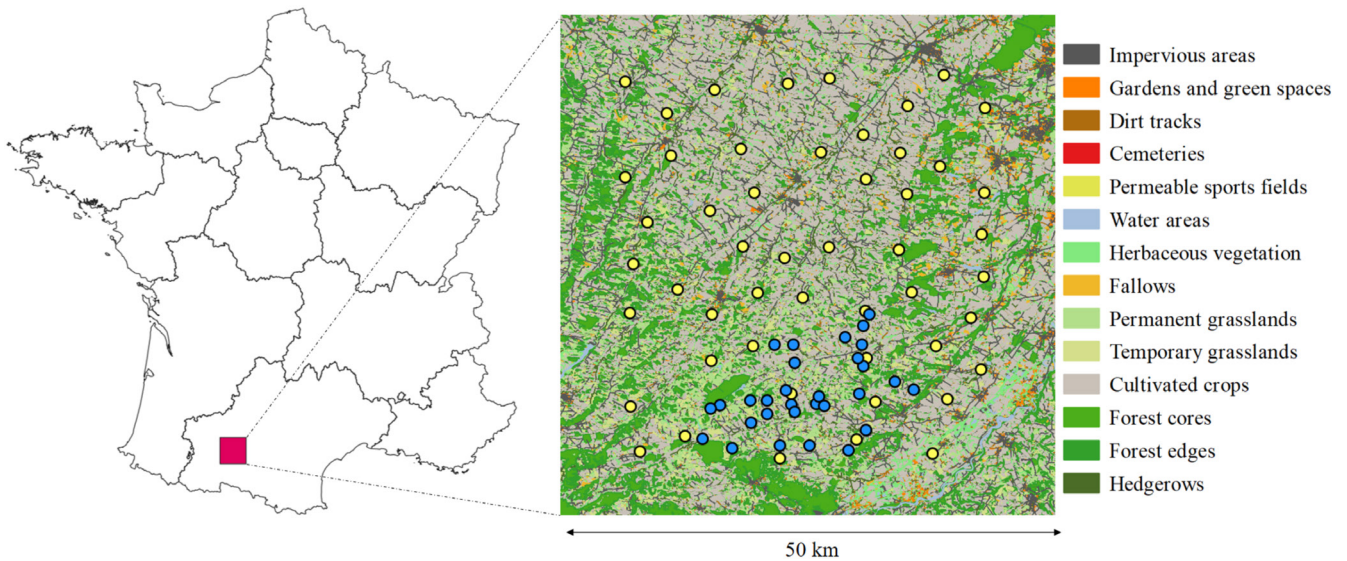
Figure 3. The left-hand map shows the French administrative regions and the right-hand map is the land use/land cover (LULC) map of the study area. The yellow dots represent the 50 target sunflower fields used in the simulation and the blue dots represent the 30 pan trapping sites.

the 50 %ΔVisitation(r) functions we identified the maximum percentage change for each radius, noted %ΔVisitation$_{MAX}(r)$. The %ΔVisitation$_{MAX}$ function is expected to decrease with increasing buffer radius and should reflect how the effects of the landscape decrease with distance. Finally, for the 16 modelled bees, we identified the radius abbreviated $\lambda_{1\%}$ thereby verifying the equation: %ΔVisitation$_{MAX}(\lambda_{1\%}) = 1\%$. Therefore, $\lambda_{1\%}$ represents the distance beyond which landscape composition (i.e. flower and nesting plots) is responsible for less than 1% change in the visitation rates of the modelled bee in the centre of the sunflower field.

We used a linear model to establish the statistical relationship between the average foraging distances of the 16 modelled bees (α) and their $\lambda_{1\%}$ distances.

Scale of effect for the modelled bees

To identify the scale of effect for the 16 modelled bees, we performed multi-scale analyses (Fig. 1) with two landscape variables calculated within 108 buffer radii (from 10 to 5000 m) as the explanatory variables, and the simulated visitation rate of modelled bees in the 50 sunflower fields as the response variable.

The two landscape variables were 1) the mean availability of floral resources in July, and 2) the mean availability of belowground nests. These two variables were weighted means obtained by summing the indices of floral resource or nesting resource availability in each LULC category occurring around the 50 fields within the 108 buffer radii (from 10 to 90 m by 10-m and from 100 to 5000 m by 50-m), weighted by the proportion of each LULC category in the different buffer zones.

Concerning the response variable, we used the visitation rate calculated with the largest circular map extent (buffer radius = 5000 m). We chose the 5000 m buffer radius

to have a 'scale-free' visitation rate estimation with no map border effects.

We fitted 1728 linear models (16 modelled bees × 108 buffer radii) to link the visitation rate of the modelled bees in the sunflower fields to the two landscape variables, after checking the assumptions for linear regressions (normality, homoscedasticity and independence of residuals) (Zuur et al. 2010). By construction in the pollination model, the mean floral and nesting resource availability are key variables for visitation predictions. Therefore, we only considered the two statistical models including both variables. Using a selection procedure based on the corrected Akaike information criterion (AICc) (Barton 2022), we selected the model most frequently identified as the best model across 108 buffer radii.

For each modelled bee, we compared the explanatory performance (goodness of fit) of the 108 models using the amount of explained variance (R^2). The buffer radius associated with the best R^2 was considered to be the scale of effect for the modelled bee. The scale of effect for the 16 modelled bees was related to their average foraging distance using linear models. An additional sensitivity analysis was performed for each scale of effect using a delete-five-observations Jackknife. This approach consisted in randomly removing five sampling sites from the data (10% of the data) before calculating the scale of effect. The procedure was repeated 1000 times.

Validation with empirical data

Pan trapping of bee species

In mid-April 2016, 30 different sampling sites were selected in the margins of cereal fields in the study areas. In each site, three coloured pan traps (blue, white, yellow), filled with soapy water were set up and left for four days. The collected bees were then counted and identified to the species level

by experienced entomologists (for protocol details see the Supporting information).

Scale of effect for the observed bee species

We performed our analyses on a subset of 16 bee species, selected according to the observed abundance (occurrence in at least three different sites), and the availability of information about their nesting preference and intertegular distance (ITD; distance between the wing's insertions) in Carrié et al. (2017, 2018). The foraging distance of the species was estimated using the allometric relationship between the intertegular distance and the typical homing distances (THD) proposed by Greenleaf et al. (2007, reported $R^2 = 0.72$). The THD is the distance where 50% of the bees return home when released at this distance. This is only a proxy measurement for foraging range. Hereafter, we distinguish the 'average foraging distance' associated with the modelled bees (α parameter), from the 'typical homing distance (THD)' associated with the observed bee species. We limited our analyses to bee species with a typical homing distances of less than 1000 m (80% of the observed bee species) to be consistent with the average foraging distances considered for the modelled bees.

To identify the scale of effect for the 16 observed bee species, we performed multi-scale analyses with the same landscape variables as those used for the modelled bees and calculated within the same 108 buffer radii (from 10 to 5000 m). The two explanatory variables were the mean availability of floral resources in April (month of pan trapping), and the mean availability of belowground nests. As response variable, we used the observed abundance of each species in pan traps at each sampling site.

We performed generalized linear models (GLM) since species abundances are count data. For each species, we selected

between the Poisson and the negative Binomial distributions according to the algorithm convergence warnings and the overdispersion of residuals (i.e residual deviance > residual degrees of freedom, following Hervé (2016)). For most species not all four models (abbreviated M1–4 in Table 1) were suitable for scale of effect analysis. For each species, we identified all suitable models for scale of effect analysis with two criteria 1) $\Delta AICc < 2$ compared to the model with the best AICc (i.e. $\Delta AICc_{inter-model}$, Supporting information) and 2) $\Delta AICc > 4$ between the 'worst' and the 'best' scale (i.e. $\Delta AICc_{inter-scale}$, Supporting information), in order to have enough AICc variability among scales to distinct the scale of effect. We also verified the existence of significant landscape effects at least at one of the 108 studied scales. Finally, among the suitable models, we selected the most complete model (see the Supporting information for detailed methods).

We tested for global spatial autocorrelation in model residuals (Kühn and Dormann 2012) using the Moran test (Sokal and Oden 1978). When global spatial autocorrelation was detected at least at one radius, we identified the sampling sites causing autocorrelation using the local indicators of spatial associations (LISA) (Anselin 1995), and then removed the sampling sites from the analysis. For each species, the explanatory quality of the model was evaluated across the 108 buffer radii using the AICc (Jackson and Fahrig 2012, 2014). The buffer radius associated with the lowest AIC was considered to be the scale of effect for the species. Finally, we used a delete-two-observations jackknife of the scale of effect to estimate sensitivity to subsampling.

Sensitivity analysis

The aim of this analysis was to assess how a variation of InVEST model parameters related to the floral and nesting

Table 1. Summary data associated with each bee species. The first column shows total abundance, all sampling sites combined, the second column correspond to the number of sites in which the species was found, the third column the intertegular distance (ITD) from Carrié et al. (2017, 2018), the fourth column correspond to the predicted typical homing distance (THD) according to Greenleaf et al. (2007), the fifth columns correspond to our index of maximum distance of effect $\lambda_{1\%}$. The sixth and seventh columns correspond to the statistical models used to study the scale of effect. 'P' and 'NB' stand for Poisson and Negative Binomial distributions, 'Ab', 'F' and 'N' respectively stand for 'Abundance', 'mean availability of floral resources' and 'mean availability of belowground nests'.

Species	Ab	Sites	ITD (mm)	THD (m)	$\lambda_{1\%}$ (m)	Distrib.	Selected model	Pseudo R^2 (McFadden)	Scale of effect (m)
1-*Andrena haemorrhoa*	7	6	2.64	529.5	311	P	M2: Ab ~ F+ N	0.36	1100
2-*Andrena nigroaenea*	6	6	2.84	670.9	3876	P	M2: Ab ~ F+ N	0.31	150
3-*Andrena nigroolivacea*	12	10	2.46	421.1	2527	P	M1: Ab ~ F × N	0.37	1800
4-*Andrena nitida*	7	5	2.97	775.7	4442	P	M1: Ab ~ F × N	0.66	2550
5-*Andrena ranunculi*	4	3	1.75	139.6	1007	P	M2: Ab ~ F+ N	0.32	750
6-*Andrena ventricosa*	5	5	1.63	110.9	852	P	M1: Ab ~ F × N	0.45	10
7-*Halictus maculatus*	3	3	1.55	94.2	762	P	M1: Ab ~ F × N	0.88	50
8-*Halictus scabiosae*	10	9	2.44	410.1	2468	P	M1: Ab ~ F × N	0.36	20
9-*Halictus simplex*	16	12	1.9	182.3	1237	NB	M3: Ab ~ F	0.11	NA
10-*Lasioglossum albipes*	5	4	1.64	113.1	863	P	M2: Ab ~ F+ N	0.39	650
11-*Lasioglossum malachurum*	82	23	1.79	150.2	1064	NB	M1: Ab ~ F × N	0.42	1550
12-*Lasioglossum marginatum*	48	11	1.59	102.3	805	NB	M3: Ab ~ F	0.43	10
13-*Lasioglossum pallens*	6	5	1.59	102.3	805	P	M2: Ab ~ F+ N	0.21	NA
14-*Lasioglossum pauxillum*	20	11	1.24	45.7	500	NB	M4: Ab ~ N	0.16	20
15-*Lasioglossum puncticolle*	24	15	1.49	82.9	701	NB	M2: Ab ~ F+ N	0.23	350
16-*Lasioglossum villosulum*	16	10	1.33	57.3	562	NB	M2: Ab ~ F+ N	0.24	100

resource availability affected the scale of effect for five different modelled bees ($\alpha = 50, 100, 200, 400, 800$ m), and for the 16 observed bees. Following the same methodology as the sensitivity analysis in Lonsdorf et al. (2009), for each parameter, we generated a random variation of ± 0.1 drawn from a uniform distribution. This variation appeared as a reasonable interval relative to the range in the floral and nesting resource availability. To keep the duration of simulation and the volume of data manageable, we generated 100 new parameter combinations and assessed for each combination how it affected the scale of effect for the modelled and the observed bees (see the Supporting information for details).

Results

λ distances and relationship with foraging distance for the modelled bees

For each modelled bee, the analyses of the percentage change in visitation rates in the target fields (%ΔVisitation) according to the increasing radius of the circular input maps revealed contrasted patterns depending on the landscape context of the fields (Fig. 4a). As expected, for each modelled bee, the maximum percentage change in visitation rates (%ΔVisitation$_{MAX}$) decreased with increasing radius of the circular input maps (Fig. 4a). Therefore, for each modelled bee, we identified the radius, abbreviated $\lambda_{1\%}$, representing the distance beyond which landscape composition is responsible for less than 1% change in the visitation rates in the target sunflower fields. We showed that $\lambda_{1\%}$ distances increased significantly with increasing average foraging distance of the modelled bees (p-value < 0.001). The $\lambda_{1\%}$ distances were linearly related to the average foraging distances (Fig. 4b) through the following equation: $\lambda_{1\%} = 5.4\alpha + 253$ ($r_{pearson} = 0.99$, $R^2 = 0.99$).

Scale of effect for the modelled bees

For all the modelled bees, we selected the linear model with no interaction between the mean availability of floral resources and the mean availability of belowground nests, as it was most frequently identified as the best model across 108 buffer radii. Using multi-scale analyses (Fig. 1), based on R^2 comparisons across buffer radii, we then detected the scale of effect for each modelled bee. The scale of effect increased significantly with increasing average foraging distance of the modelled bee (p-value < 0.001, Fig. 5a). The linear relationship between the average foraging distance (α) and the scale of effect (S^*), is given by the following equation: $S^* = 1.8\alpha - 35$ ($r_{pearson} = 0.99$, $R^2 = 0.97$). The results of the jackknife subsampling revealed that removing five observations moderately changed the scale of effect (Fig. 5a). The frequency distributions of the scale of effect values were unimodal and the interquartile intervals were less than 250 m (Supporting information). The variability of the scale of effect values increased with increasing average foraging distance of the

modelled bee. Importantly, no scale of effect values was found above the $\lambda_{1\%}$ equation. The sensitivity analysis on floral and nesting resource parameters revealed negligible variation in scale of effect (Supporting information).

Scale of effect for the observed bee species

For ten species, we selected the Poisson distribution, and for the six remaining species the negative Binomial distribution (Table 1). Among all suitable models, we selected a single model per species to identify the scale of effect (Table 1) but alternative models were also analysed (Supporting information). For two species (*H. simplex*, *L. pallens*), no models appeared suitable for scale of effect analysis (AICc$_{Model}$ > AICc$_{Null\ Model}$ and ΔAICc$_{inter-scale}$ < 4, respectively), thus these species were removed from the analysis (Table 1). Based on the selected models, we detected significant spatial autocorrelation in three species and therefore removed the sampling sites responsible for the autocorrelation (one sampling site for *A. nitida* and *H. scabiosae* and five sampling sites for *L. malachurum*). Although we detected significant landscape effects for all species, for some (e.g. *L. pauxillum*), we noticed low deviance explained (Table 1) with AICc close to the null models (Supporting information).

Using multi-scale analyses (Fig. 1), based on AICc comparisons across buffer radii, we detected the scale of effect for each bee species (Fig. 5b). The scale of effect increased significantly with increasing typical homing distance (p-value = 0.036). The linear relationship between the typical homing distance (THD) and the scale of effect (S^*) is given by the following equation: $S^* = 1.8\,THD + 162$ ($r_{pearson} = 0.56$, $R^2 = 0.32$). The results of the jackknife subsampling revealed that removing two observations drastically changed the scale of effect (Fig. 5b). For four species (*A. haemorrhoa*-1, *H. scabiosae*-8, *L. pauxillum*-14, *L. puncticolle*-15), we found a multimodal distribution in the frequency distribution of the scale of effect values, with modes separated by more than 1000 m (Supporting information).

Importantly, the scale of effect for 13 out of the 14 bee species were below the $\lambda_{1\%}$ equation. Only the scale of effect for *L. malachurum* (11), was 1.5 times above $\lambda_{1\%}$ (Fig. 5b). The sensitivity analysis on floral and nesting resource parameters revealed a very similar pattern of scale of effect variation compared to jackknife subsampling (Supporting information).

Discussion

Improving our ability to estimate, a priori, the scales at which species interact with their environment is fundamental to understand the patterns and processes underlying species–habitat relationships (Thornton and Fletcher 2014). Based on simulations using the InVEST pollination model plus empirical data, we showed that the average foraging distance (and the typical homing distance for observed bees) is positively correlated with the scale of effect
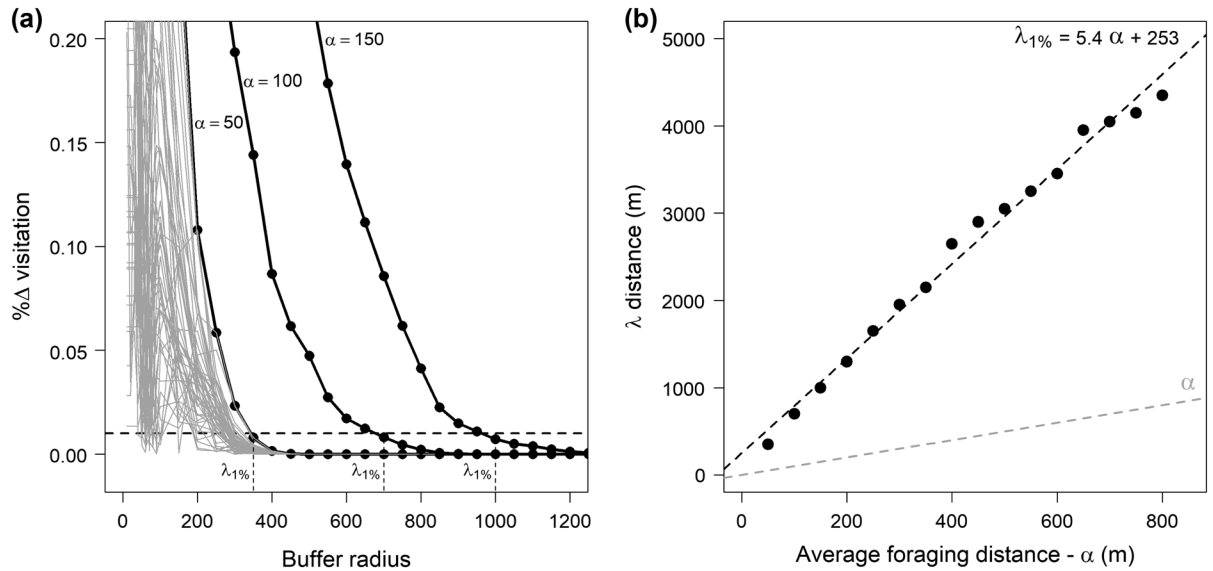
Figure 4. Left-hand plot (a), example for three modelled bees ($\alpha = 50$, 100, 150 m) of the relationships between the percentage change in visitation rates in the centre of the target sunflower fields according to the radius of the circular input maps. For the modelled bee with an average foraging distance of 50 m, the grey lines represent the percentage change in visitation rates for the 50 target sunflower fields. The black lines represent the maximum percentage change in visitation rates for three modelled bees. The horizontal dashed line is the 1% threshold and the vertical dashed lines the $\lambda_{1\%}$ distances associated with the modelled bees. Right-hand plot (b), linear relationship between the average foraging distance of the modelled bees (α), and their $\lambda_{1\%}$ distance. The grey dashed line represents the average foraging distance (x = y equation).

for wild bees. We defined and estimated a novel metric, the maximum distance of effect, noted $\lambda_{1\%}$, representing the distance beyond which landscape composition is responsible for less than 1% change in the simulated visitation rates of bees. We demonstrated that $\lambda_{1\%}$ is linearly related to the average foraging distance (α) of the modelled bees ($\lambda_{1\%} = 5.4\alpha + 253$) and that this index represents an upper limit distance beyond which the scale of effect for wild bees is unlikely to occur. The scale of effect for 13 out of the 14 observed bee species was below $\lambda_{1\%}$. Therefore, our study is the first successful attempt to refine the general guideline proposed by Jackson and Fahrig (2012) and to validate it with empirical data.

Factors affecting the scale of effect

A common explanation for the scale of effect is that it reflects the characteristic distance at which organisms interact with their environment (Holland et al. 2004). However, using a model that assumes that bee diversity is proportional to the size of bee habitats, Dormann and Seppelt (2007) showed that the scale of effect for bee diversity may simply reflects the distribution of habitats in the landscape and does not necessarily reflect a characteristic distance of biological interaction. The origin of the scale of effect has also been called into question because of the absence of empirical validation of the effect of species mobility on the scale of effect (Miguet et al. 2016, Stuber et al. 2018, Moll et al. 2020). Here, using a spatially explicit pollination model, our simulation results provide theoretical evidence that the scale of effect is correlated

with bees' foraging distances. Our simulation results are validated with empirical data as we observed a similarity between the linear equations for the modelled bees ($S^* = 1.8\alpha - 35$) and observed bee species ($S^* = 1.8THD + 162$). Our results support the hypothesis that allometric relationships related to species mobility can predict their scale of effect (Wiens 1989, Jackson and Fahrig 2012, Ricci et al. 2013).

Compared to our simulation results there was much variability in the estimated scale of effect for a given bee species, which hints that there are other factors at play. A first source of scale of effect variability is directly related to its definition. The scale of effect is defined as the single geographical extent exhibiting the best relationship between the landscape variables and the biological response studied (Jackson and Fahrig 2012). Focusing on a single scale occults that the related scales also explain the biological response (Miguet et al. 2016). The concept of scale of effect should rather be understood as a 'scale domain' of effect (Wiens 1989). Using resampling methods or sensitivity analyses, as we did (Moraga et al. 2019), helps in revealing these scale domains. Here, randomly removing a few sampling sites caused drastic shifts in the scale of effect for the observed bee species. Although other multi-scale studies used similar sized samples (Steffan-Dewenter et al. 2002, Krauss et al. 2003, Meyer et al. 2009, Moraga et al. 2019), scale of effect variations is probably due to our small sampling size relative to the biological variability in species abundance.

A second source of scale of effect variability is the possible existence of multiple scale domains of effect. For four observed bee species, we found multimodality in the
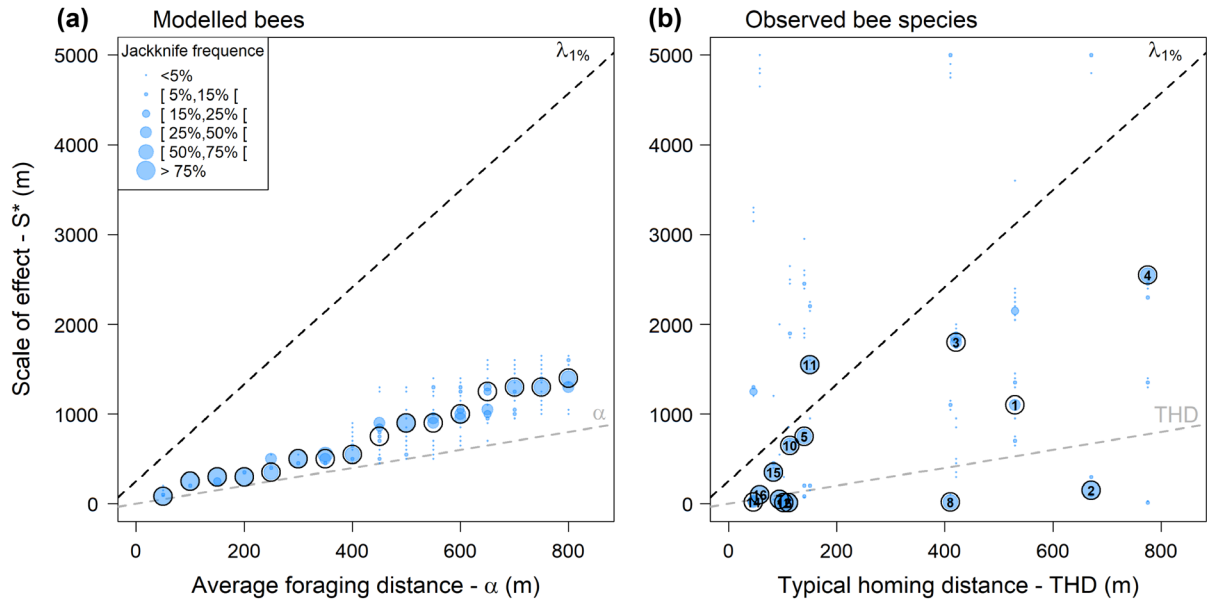
Figure 5. Left-hand plot (a), relationship between the scale of effect for the modelled bees and their average foraging distances (α). Right-hand plot (b), relationship between the scale of effect for the observed bee species and the prediction of their typical homing distance based on the intertegular distance according to Greenleaf et al. (2007). The numbers represent the species identity (Table 1). In both plots, the black circles represent the scale of effect based on the complete sample size, and the blue dots represent the scale of effect values obtained with jackknife subsampling. The size of the dots represents the occurrence frequency of the scale of effect values. The grey dashed line represents the x = y equation.

frequency distribution of the scale of effect, which is consistent with recent studies (Moraga et al. 2019, Stuber and Fontaine 2019). Multimodality may reflect artefacts related to the heterogeneous distribution of habitats (Dormann and Seppelt 2007) or that species abundance is shaped by different ecological processes acting at different spatial (and temporal) scales (e.g. foraging versus breeding activities or individual versus population processes; Stuber and Fontaine 2019, Desaegher et al. 2021). The multimodality creates ambiguity when investigating the scale of effect of a species, especially when results depend on the explanatory variables studied (Moraga et al. 2019, Stuber and Fontaine 2019, Moll et al. 2020) or on the statistical model selected. In our case, for the species *L. malachurum* (11), supplementary analyses performed on alternative models revealed drastically different results with a scale of effects at 200 m, compared to 1500 m with the selected model (Supporting information). Future studies should account for this intra-species scale of effect variability, when aiming to identify factors that affect the scale of effect (Miguet et al. 2016, Stuber and Fontaine 2019).

Strength and limits of the λ index

Our simulation results concerning maximum distance of effect $\lambda_{1\%}$, revealed that landscape effects can still be observed far beyond the scale of effect, and probably further than it is commonly acknowledged. The strength of this index, is that it restricts the upper limit of the range of candidate scales in multi-scale analyses while being conservative regarding scale of effect variability.

It has recently been recommended that the effects of landscape on biological responses should be investigated using distance weighting functions instead of a step function as is implicitly assumed in multi-scale analyses (Chandler and Hepinstall-Cymerman 2016, Miguet et al. 2017, Moll et al. 2020). The InVEST pollination model and the λ index do that as they are based on negative exponential weighting functions that account for the decreasing influence of landscape composition on pollinator supply and visitation rate with increasing distance (Sharp et al. 2018). The use of different weighting function probably accounts for the quantitative discrepancy with Jackson and Fahrig (2012) results ($S^* \approx$ 4–9 times median dispersal distance).

InVEST pollination model does not take into account the order of the seasons or phenological gaps in floral resources when predicting visitation rates (Sharp et al. 2018). Nevertheless, this should only have a limited impact on the estimation of the λ index and the scale of effect for the modelled bees. Indeed, there is no reason to think of a spatial bias in the distribution of these underestimated resources relative to our sampling sites. More generally, the same argument holds if we underestimate or overestimate the importance of some land use categories by attributing the floral or nesting resource availability and is supported by our sensitivity analysis.

The main limitation of our study is related to the validation dataset, as we have a small sampling size, sometimes

low explanatory power, and only included 16 ground-nesting bee species (three genera). Also we did not account that bee abundance can depend on floral resources available before the sampling month or more generally on the phenological continuity along the year as is growingly acknowledged (Timberlake et al. 2019, Guezen and Forrest 2021). Therefore, we call for future studies to test our equation with other species, in different seasons and accounting for other factors, for instance the aggregation of habitats that has been shown to affect the scale of effect (Ricci et al. 2013). However it should be reassuring that the model we used already showed its ability to explain the field estimates of bee abundance (Lonsdorf et al. 2009, Groff et al. 2016, Desaegher et al. 2021).

Applications of the λ index

We identified different applications to our proposed equation ($\lambda_{1\%} = 5.4\alpha + 253$) but in all cases predictions should be used with caution. First, the λ index should improve the efficacy of multi-scale analyses on bee species, as it 1) delimits the maximum buffer size of analysis and 2) help distinguish potential scale-dependant ecological mechanisms from artefacts related to sampling bias or to the heterogeneous distribution of habitats. Second, when setting up a new field experiment, the index can identify the minimum distance between sampling sites to avoid spatial pseudo-replication. For example, when applying the equation to the wild bee community observed in our study area, we obtained a $\lambda_{1\%} = 2.5$ km (THD weighted mean = 420 m). A safe distance between sites would be two times $\lambda_{1\%}$ (non-overlapping buffer radius of $\lambda_{1\%}$). Landscape sampling design is, of course, a tradeoff between the distancing and number of sites and the extent of the study area. Finally, as a quick estimate of the maximum distance of landscape effects on wild bees, λ index could help landscape managers and conservation biologists, in scaling or designing nature reserves.

Author contributions

James Desaegher: Conceptualization (lead); Data curation (lead); Formal analysis (lead); Investigation (lead); Methodology (lead); Resources (lead); Software (lead); Validation (lead); Visualization (lead); Writing – original draft (lead); Writing – review and editing (lead). **Annie**

Ouin: Conceptualization (supporting); Data curation (supporting); Formal analysis (supporting); Funding acquisition (lead); Investigation (supporting); Methodology (supporting); Project administration (supporting); Resources (supporting); Software (supporting); Supervision (lead); Validation (supporting); Visualization (supporting); Writing – original draft (supporting); Writing – review and editing (supporting). **David Sheeren**: Conceptualization (supporting); Data curation (supporting); Formal analysis (supporting); Funding acquisition (lead); Investigation (supporting); Methodology (supporting); Project administration (lead); Resources (equal); Software (equal); Supervision (lead); Validation (supporting); Visualization (supporting); Writing – original draft (supporting); Writing – review and editing (supporting).

Transparent Peer Review

The peer review history for this article is available at <https://publons.com/publon/10.1111/ecog.05758>.

Data availability statement

Data are available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.0rxwdbs2k> (Desaegher et al. 2022).

Supporting information

The supporting information associated with this article is available from the online version.

References

Anselin, L. 1995. Local indicators of spatial association – LISA. – Geogr. Anal. 27: 93–115.
Barton, K. 2022. MuMIn: multi-model inference. R package ver. 1.46.0. – <https://CRAN.R-project.org/package=MuMIn>.
Baude, M. et al. 2016. Historical nectar assessment reveals the fall and rise of floral resources in Britain. – Nature 530: 85.
Cane, J. H. 1991. Soils of ground-nesting bees (Hymenoptera: Apoidea): texture, moisture, cell depth and climate. – J. Kansas Entomol. Soc. 64: 406–413.
Carrié, R. et al. 2017. Relationships among ecological traits of wild bee communities along gradients of habitat amount and fragmentation. – Ecography 40: 85–97.
Carrié, R. et al. 2018. Bee diversity in crop fields is influenced by remotely-sensed nesting resources in surrounding permanent grasslands. – Ecol. Indic. 90: 606–614.
Chandler, R. and Hepinstall-Cymerman, J. 2016. Estimating the spatial scales of landscape effects on abundance. – Landscape Ecol. 31: 1383–1394.
Choisis, J.-P. et al. 2010. Understanding regional dynamics of mixed crop-livestock agricultural systems to support rural development in south-western France uplands. – Cah. Agric. 19: 97–103.
Desaegher, J. et al. 2021. Optimising spatial distribution of mass-flowering patches at the landscape scale to increase crop pollination. – J. Appl. Ecol. 58: 1876–1887.

Desaegher, J. et al. 2022. Data from: How far is enough? Prediction of the scale of effect for wild bees. – Dryad Digital Repository, <https://doi.org/10.5061/dryad.0rxwdbs2k>.

Dormann, C. and Seppelt, R. 2007. Do organisms use landscapes at certain spatial scales ? A null model for diversity pattern in relation to the spatial extend of landscapes. – In: Schröder, B. et al. (eds), Multiple scales in ecology. Peter Lang, Frankfurt, Germany, pp. 75–81.

EFSA 2013. EFSA Guidance document on the risk assessment of plant protection products on bees (*Apis mellifera*, *Bombus* spp. and solitary bees), Appendix D. – EFSA J. 11: 3295.

Fahrig, L. 2005. When is a landscape perspective important? – In: Wiens, J. A. and Moss, M. R. (eds), Issues and perspectives in landscape ecology. Cambridge Univ. Press, pp. 3–10.

Gallai, N. et al. 2009. Economic valuation of the vulnerability of world agriculture confronted with pollinator decline. – Ecol. Econ. 68: 810–821.

Gathmann, A. and Tscharntke, T. 2002. Foraging ranges of solitary bees. – J. Anim. Ecol. 71: 757–764.

Greenleaf, S. S. et al. 2007. Bee foraging ranges and their relationship to body size. – Oecologia 153: 589–596.

Groff, S. C. et al. 2016. Parameterization of the InVEST crop pollination model to spatially predict abundance of wild blueberry (*Vaccinium angustifolium* Aiton) native bee pollinators in Maine, USA. – Environ. Model. Softw. 79: 1–9.

Guezen, J. M. and Forrest, J. R. K. 2021. Seasonality of floral resources in relation to bee activity in agroecosystems. – Ecol. Evol. 11: 3130–3147.

Hervé, M. 2016. Aide-mémoire de statistique appliquée à la biologie: construire son étude et analyser les résultats à l'aide du logiciel R. – <https://cran.r-project.org/doc/contrib/Herve-Aide-memoire-statistique.pdf>.

Holland, J. D. et al. 2004. Determining the spatial scale of species' response to habitat. – Bioscience 54: 227–233.

Hutchinson, G. 1965. The ecological theatre and the evolutionnary play. – Yale Univ. Press.

Jackson, H. B. and Fahrig, L. 2012. What size is a biologically relevant landscape? – Landscape Ecol. 27: 929–941.

Jackson, H. B. and Fahrig, L. 2014. Are ecologists conducting research at the optimal scale? – Global Ecol. Biogeogr. 24: 52–63.

Krauss, J. et al. 2003. How does landscape context contribute to effects of habitat fragmentation on diversity and population density of butterflies? – J. Biogeogr. 30: 889–900.

Kühn, I. and Dormann, C. F. 2012. Less than eight (and a half) misconceptions of spatial analysis. – J. Biogeogr. 39: 995–998.

Levin, S. A. 1992. The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. – Ecology 73: 1943–1967.

Lonsdorf, E. et al. 2009. Modelling pollination services across agricultural landscapes. – Ann. Bot. 103: 1589–1600.

Meyer, B. et al. 2009. Contrasting resource-dependent responses of hoverfly richness and density to landscape structure. – Basic Appl. Ecol. 10: 178–186.

Michener, C. D. 2007. The bees of the world, 2nd edn. – Johns Hopkins Univ. Press.

Miguet, P. et al. 2016. What determines the spatial extent of landscape effects on species? – Landscape Ecol. 31: 1177–1194.

Miguet, P. et al. 2017. How to quantify a distance-dependent landscape effect on a biological response? – Methods Ecol. Evol. 8: 1717–1724.

Moll, R. J. et al. 2020. At what spatial scale(s) do mammals respond to urbanization? – Ecography 43: 171–183.

Moraga, A. D. et al. 2019. The scale of effect of landscape context varies with the species' response variable measured. – Landscape Ecol. 34: 703–715.

Ollerton, J. et al. 2011. How many flowering plants are pollinated by animals? – Oikos 120: 321–326.

Ouin, A. et al. 2022. Chapter Two – Building a shared vision of the future for multifunctional agricultural landscapes. Lessons from a long term socio-ecological research site in south-western France. – In: Bohan, D. A. et al. (eds), The future of agricultural landscapes, part III. Academic Press, pp. 57–106.

Ricci, B. et al. 2013. Do species population parameters and landscape characteristics affect the relationship between local population abundance and surrounding habitat amount? – Ecol. Complex. 15: 62–70.

Sharp, R. et al. 2018. InVEST 3.6.0 User's Guide. The Natural Capital Project, Stanford Univ., Univ. of Minnesota, The Nature Conservancy, and World Wildlife Fund. pp. 62–72.

Sokal, R. R. and Oden, N. L. 1978 Spatial autocorrelation in biology: 1. Methodology. – Biol. J. Linn. Soc. 10: 199–228.

Steffan-Dewenter, I. et al. 2002. Scale-dependant effects of landscape context on three pollinator guilds. – Ecology 83: 1421–1432.

Stuber, E. F. and Fontaine, J. J. 2019. How characteristic is the species characteristic selection scale? – Global Ecol. Biogeogr. 28: 1839–1854.

Stuber, E. F. et al. 2018. Predicting species–habitat relationships: does body size matter? – Landscape Ecol. 33: 1049–1060.

Thornton, D. H. and Fletcher, R. J. 2014. Body size and spatial scales in avian response to landscapes: a meta-analysis. – Ecography 37: 454–463.

Timberlake, T. P. et al. 2019. Phenology of farmland floral resources reveals seasonal gaps in nectar availability for bumblebees. – J. Appl. Ecol. 56: 1585–1596.

Wiens, J. A. 1989 Spatial scaling in ecology. – Funct. Ecol. 3: 385–397.

Zuur, A. F. et al. 2010. A protocol for data exploration to avoid common statistical problems. – Methods Ecol. Evol. 1: 3–14.