ECOGRAPHY

Research article

Wild bees and landcover: bee species' body size does not predict the scale of effect, but bee phenology predicts association with landcover type

Dylan T. Simpson[1,2,3], **Colleen Smith**[1,4] **and Rachael Winfree**[5]

[1]Graduate Program in Ecology & Evolution, Rutgers University, New Brunswick, NJ, USA
[2]Plant Biology & Conservation, Northwestern University, Evanston, IL, USA
[3]Negaunee Institute for Plant Conservation Science and Action, Chicago Botanic Garden, Glencoe, IL, USA
[4]US Fish and Wildlife Service, Ventura, CA, USA
[5]Department of Ecology, Evolution & Natural Resources, Rutgers University, New Brunswick, NJ, USA

Correspondence: Dylan T. Simpson (dylan.tux.simpson@gmail.com)

Ecography
2025: e07982
doi: 10.1002/ecog.07982

Subject Editor: Carsten Dormann
Editor-in-Chief: Miguel Araújo
Accepted 24 June 2025

Habitat is a key aspect of any species' niche and can affect populations at multiple spatial scales. Basic ecology and effective conservation thus require an understanding of which habitats matter and at what scales. Yet, habitat studies are rarely scale-optimized, and what determines the scale(s) at which populations are affected by surrounding habitat (the 'scale of effect') is poorly understood. In this study, we test the 'mobility hypothesis,' which predicts that species with larger foraging ranges should have larger scales of effect. The mobility hypothesis is the most popular explanation of what determines species' scales of effect, but empirical support is mixed. We test the mobility hypothesis using wild bee species and, in doing so, also assess landscape-scale habitat associations of 84 bee species. We collected 30 376 specimens of 84 bee species from 165 sites in the northeastern USA and used linear models to determine landcover associations and scales of effect for each species. To test the mobility hypothesis, we asked whether scales of effect varied with two mobility-related traits – body size or sociality, which are the strongest known predictors of bee foraging ranges. Controlling the false discovery rate at 5%, we found 193 significant species–landcover associations across 60 (of 84) species. Scales of effect ranged from 100 to 8000 m (mode = 200 m; median = 1000 m) and, counter to the mobility hypothesis, were not associated with body size or sociality. As a result, we argue that ecologists should reconsider making assumptions about species' scales of effect and should instead explicitly measure scales of effect for their particular study organism and system. Considering the landcover associations themselves, we found these were broadly explained by phenology, with spring-flying bees being associated with forests and summer-flying bees being associated with more open, non-forested habitats.

Keywords: habitat associations, multi-scale analysis, phenology, scale of response, scale optimization

NORDIC SOCIETY OIKOS

www.ecography.org

Introduction

Habitat is a key aspect of any species' niche and is thus fundamental to both basic ecology and effective conservation and management (Grinnell 1924, Elton 1927, Hanski 2011). Understanding a species' habitat niche requires knowing not only which habitats a species responds to (James et al. 1984), but also the spatial extents at which habitat has an effect (Addicott et al. 1987, Brennan et al. 2002). The spatial extent at which habitat or related landcover variables have their strongest effect is referred to as their 'scale of effect' (Martin and Fahrig 2012) or sometimes 'scale of response' (Holland et al. 2004). Scales of effect are critical for study design and conservation planning (Jackson and Fahrig 2015, Gallo et al. 2018, Yeiser et al. 2021) but, despite their importance, scales of effect remain poorly understood and hard to predict.

There are many hypotheses as to what determines a species' scales of effect, but most have little empirical support (Miguet et al. 2016). The most popular hypothesis, which in this paper we call the 'mobility hypothesis,' posits that species with larger home ranges and/or greater dispersal distances should have larger scales of effect. The reasoning is that such species interact with their environment across larger spatial scales (Jackson and Fahrig 2012, Miguet et al. 2016). This general hypothesis has also led to an expectation that larger animals will tend to have larger scales of effect, because larger animals are typically more mobile (McCauley et al. 2015, Miguet et al. 2016). The mobility hypothesis has clear support from simulation studies (Jackson and Fahrig 2012, Ricci et al. 2013, Desaegher et al. 2022), but real-world tests have been inconsistent (Thornton and Fletcher 2014, Miguet et al. 2016, Stuber et al. 2018, Moll et al. 2020, Desaegher et al. 2022, Arroyo-Rodríguez et al. 2023). And yet, despite the equivocal support, the mobility hypothesis is often the default assumption for studies of how species respond to habitat.

In this study, we test the mobility hypothesis using data on wild bees. Specifically, we relate our focal species' scales of effect to mobility-related traits. In doing so, we also assess bee habitat associations, which themselves are also poorly understood.

Due to concern over bee declines, landscape ecology studies of bees have become increasingly common. Such studies, however, typically do not measure scales of effect. Instead, studies tend to rely on two simplifying assumptions. First, studies typically choose a single scale a priori and assume it applies to their entire bee community (Pardee and Philpott 2014, Carrié et al. 2017, Bartomeus et al. 2018). If a study is focused on an aggregate metric like diversity, this might be a necessary assumption, but this assumption is also used when species are analyzed individually (Collado et al. 2019, Novotny et al. 2021, Smith et al. 2021). This is a problem because, if scales of effect vary among species, this method would miss species–habitat relationships occurring at different scales (McGarigal et al. 2016). Second, to choose a scale of effect a priori, most studies rely on the mobility hypothesis and assume bees' scales of effect correspond to their foraging range. In practice, because bee foraging distances are usually unknown, scales of effect are estimated from body size, which is correlated with flight capacity (Greenleaf et al. 2007, Kendall et al. 2022). The single chosen scale of effect is meant to represent an average or maximum for the community. Ultimately, this use of the mobility hypothesis may be a good one, but we consider it premature because direct evidence for the mobility hypothesis in bees is limited. A direct test would estimate scales of effect for multiple bee species and regress them against proxies of mobility. We are aware of only one such study, which did find modest support for the mobility hypothesis ($R^2 = 0.32$, $p = 0.036$) but tested only 14 bee species (Desaegher et al. 2022). Other studies found landscape responses to differ between 'large' and 'small' bees (Klein et al. 2008, Benjamin et al. 2014), but it is difficult from such studies to characterize the shape or strength of the relationship between bees' body size and scales of effect. Thus, while there is some evidence for the mobility hypothesis in bees, it is limited in scope. It is also unclear whether to expect a one-to-one relationship between the scale of effect and the foraging range. Further, even assuming mobility is what determines scales of effect, body size alone may be insufficient to describe mobility. For example, sociality may have a stronger influence on foraging distance than body size (Kendall et al. 2022), and so may also help predict scales of effect. Altogether, it remains unclear whether scales of effect can be reliably predicted using mobility-related traits, as is often assumed.

Beyond the broader question of the mobility hypothesis, we also use the opportunity to determine habitat associations for our focal bee species. While bees are well studied in some respects, little is known about individual species' habitat associations. This is surprising because many bee species are declining, and habitat loss is considered one of their greatest threats (Goulson et al. 2015, Cameron and Sadd 2020). The problem is that most studies of how bees respond to habitat focus on the aggregate abundance or diversity of bee communities, without much attention to individual species (Winfree et al. 2011, but see Collado et al. 2019, Novotny et al. 2021, Smith et al. 2021). Further, the focus is often on contrasting natural and anthropogenic habitats, sometimes with little regard to distinctions among the varied habitat types within these broad groups. As a result, there is little understanding of which specific habitats matter to which bee species.

In our study region of the eastern United States, we expect bee habitat requirements to be associated with phenology. In this region, bees are more active in forests in the spring and more active in open habitats in the summer (Harrison et al. 2018). This is because bees depend on flowers, and spring-blooming flowers in this region are primarily found in forests, while summer-blooming flowers are primarily found in open habitats. As a result, we predict that spring-flying bees are more likely to be forest-associated and summer-flying bees are more likely to be open habitat-associated.

Using bee abundance data from 165 sites in the eastern United States, we assess scales of effect and habitat associations for 84 bee species. In doing so, we ask: 1) are scales of effect associated with body size or sociality, as expected under the mobility hypothesis? 2) How is our ability to detect habitat associations affected by a priori assumptions regarding species' scales of effect? 3a) With which habitats are each species associated? And finally, 3b) do habitat associations differ among species with different phenologies?

Material and methods

Bee abundance and phenology

To assess bee habitat associations and scales of effect, we used bee abundance data collected as part of five studies conducted by the Winfree Lab Group between 2003 and 2018. Four of these datasets were previously published (Winfree et al. 2007, 2014, Harrison et al. 2018, Smith et al. 2021) and the fifth we publish here for the first time (details in the Supporting information). In each study, bee communities were sampled at 44, 36, 37, 32, and 16 sites, respectively (165 sites in total), across New Jersey, Pennsylvania, and New York, USA (Fig. 1, Supporting information). These studies each investigated the effects of landcover or land use on bee communities, so sites varied widely in their landscape contexts, including agriculture, pasture, human development, and different types of forest. Each site was visited one to 11 times across one to three years. On each site visit, bees were collected by pan or vane trap and later identified to species. The number of traps and how long they were out (i.e. sampling effort) varied among studies, but were consistent within a study. Differences in sampling effort among studies were accounted for in our statistical models using random intercept terms.

To reduce data limitation in our analyses, we only included bee species that were represented by at least 30 specimens across at least 10 sites. After this filtering step, our dataset included 84 bee species represented by a total of 30 376 specimens. Within this subset, sample sizes for each of the original studies were: 1470 specimens from 44 sites (Winfree et al. 2007), 2279 specimens from 16 sites (Winfree et al. 2014), 11 923 specimens from 36 sites (Harrison et al. 2018), 10 550 specimens from 32 sites (Smith et al. 2021), and 4154 specimens from 37 sites (Winfree et al. unpubl.; see the Supporting information for details).

To characterize each bee species' phenology, we estimated each species' flight window from recorded collection dates (Supporting information). We used a combination of datasets collected by the Winfree Lab Group, including those used in this study, and collections data from the American Museum of Natural History from New Jersey, New York, and Pennsylvania (Bartomeus et al. 2011). For the 84 focal species, this phenology dataset included a total of 84 400 specimens. We defined the flight window for each species as the 5–95th percentiles of observation dates for that species.

To compare habitat associations among bees with different phenologies, we defined species as spring, summer, or long-season bees (Wood et al. 2018). These distinctions were based on patterns in the start and end dates of bees' flight windows (Supporting information): spring bees started early and ended early, summer bees started late and ended late, and long-season bees started early but ended late. Cutoffs for 'early' and 'late' were 13 May and 10 July, respectively, and were chosen on natural clustering of start and end dates (see the Supporting information for more details).

To estimate each bee species' average body size, we measured intertegular distance of female specimens (using workers, for social species) collected by the Winfree Lab Group (this study; Cariveau et al. 2016; Supporting information). Intertegular distance is the distance between a bee's wings and is the standard measure used to associate body size with foraging distance (Greenleaf et al. 2007, Kendall et al. 2022).

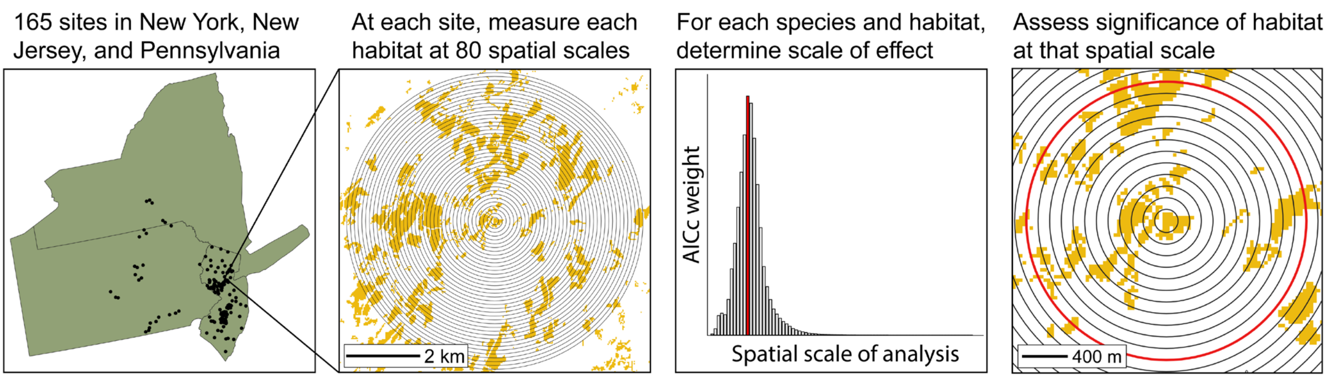


Figure 1. Summary of our analytical workflow: at each site (first panel), we measured habitat proportion within each of 80 spatial extents, from 100 to 8000 m (second panel, though for clarity this panel only shows up to 4000 m). To determine each bee species' scale of effect for each habitat, we modeled each species' abundance as a function of landcover proportion at each extent and compared model efficacy using AICc (third panel). These models also include covariates we determined to be relevant to each species. Finally, to test the significance of each bee species–landcover association, we used likelihood ratio tests to compare the best model (i.e. that in which landcover measured at that species' scale of effect; fourth panel) to a null model that included covariates but no landcover predictor. After completing this scale optimization for every species–landcover pair, we calculated p-values based on a randomization-based null distribution designed to account for our particular data structure and scale-selection process, then assessed significance while controlling our false discovery rate at 5% (following Benjamini et al. 2006).

We used at least five specimens per species to estimate average IT distance (n = 5–25).

Habitat data

We used landcover as a proxy for habitat. While habitat can be more precisely defined, especially at small scales, we are focused here on broad-scale habitat types for which landcover is the best available measure. We measured landcover proportion around each site using National Landcover Database classifications (NLCD; Wickham et al. 2023). The NLCD classifications provide continuous, standardized coverage at 30-m resolution, with seven releases from 2001 to 2016.

We used the NLCD data to measure nine landcover variables around each site. These variables were: 1) deciduous, 2) evergreen and 3) mixed deciduous and evergreen forests, 4) forested wetland, 5) pasture, and 6) crops, which were all landcover classes taken directly from the NLCD; and 7) low-intensity development, which combined the NLCD open-developed (e.g. mowed parks, golf courses) and low-intensity development classes; 8) high-intensity development, which combined the NLCD medium- and high-intensity development classes; and 9) forest edge, which we measured as the length of forest to non-forest edge per unit area (m ha^{-1}, using the 'landscapemetrics' package in R; Hesselbarth et al. 2019). We excluded the grassland landcover class, despite grasslands being known habitat for many bee species, because it was very rare in our region (≤ 3% of the landscape surrounding our sites). Most open green space was instead classified as pasture or open-developed. To minimize inaccuracies due to landcover change, we paired each dataset with the most recent NLCD classification prior to that survey (i.e. data collected in 2013 would be paired with the 2011 NLCD classification and data collected in 2017–2018 would be paired with the 2016 classification).

To help determine the scale of effect of each landcover association, we measured landcover proportion within 80 radii around each site, from 100 to 8000 m. We increased scale in 100-m steps. Our goal here was to cover as broad a range as feasible, aiming to include scales larger and smaller than our species' actual scales of response (Jackson and Fahrig 2015). We chose our minimum scale (100 m) as the smallest we felt comfortable measuring with 30-m resolution landcover data. We chose our maximum scale (8000 m) to be much larger than we expected any bee to respond. This expectation is based norms in the literature (which typically measure the landscape within ≤ 2000 m, e.g. Pardee and Philpott 2014, Carrié et al. 2017, Bartomeus et al. 2018) and the expected 'typical homing distance' of our focal species based on body size, which are all also under 2000 m (calculated following Greenleaf et al. 2007).

Collinearity among landcover classes was generally low, being lowest at small scales and relatively higher at large scales. At the 100 m-scale, the average absolute correlation among landcover classes was $|r| = 0.14$ and the maximum was $|r| = 0.31$. At the 8000-m scale, the average $|r| = 0.29$ and the maximum $|r| = 0.71$. At that largest scale, only four of 72 pairwise correlations were above $|r| = 0.5$.

Analysis

To determine landcover associations and scales of effect for each bee species, we used negative binomial mixed-effects models to describe bee abundance as a function of landcover proportion. We ran models separately for each species and landcover, because scales of effect could vary among landcover types, even for the same species (Addicott et al. 1987, Stuber and Fontaine 2019). Before modeling each species, we filtered the data to include only those collection events within the phenological windows of the focal bee species. This was to reduce the number of zeros in the dataset due to reasons other than landcover.

Our analytical workflow for each species, which we describe further below, was as follows:

1. Assess potential covariates to define that species' null model.
2. For each landcover, compare the performance of models in which the landcover predictor was measured at each of 80 scales.
3. Choose the best-performing model for each landcover and test it for significance by comparing it to that species' null model.
4. If that model proves significant, the scale of that model is taken as the scale of effect for that species–landcover association.

All our models took the same general form of $n \sim BX + (1|study) + (1|site{:}study)$, where n is the number of bees of the focal species observed during one site visit, site and study are nested random intercept terms to account for effects of study design and for repeated site visits, and *X* are predictors including landcover proportion and covariates. We ran models using the package and function 'glmmTMB' (Brooks et al. 2017), and compared models using the function *model.sel* in the package 'MuMIn' (Barton 2022), both in R ver. 4.1.3 (www.r-project.org).

For each species, the significance of any landcover predictor was tested against that species' null model, which is the same model with all the same covariates but without the landcover predictor. Thus, the first step for each species is to determine the appropriate null model. Potential covariates were latitude and day of year, each as a linear or quadratic predictor (Table 1). Latitude was meant to control for broad geographic trends and day of year to control for phenology. To determine the appropriate combination of covariates for each species, we compared nine candidate models using each combination of potential covariates but no landcover variables (Table 1). From these, we retained that with the lowest AICc as the null model for that species.

With the appropriate null model as a baseline, we then determined the best scale for each landcover predictor. To do so, we used AICc to compare models with landcover measured at each of 80 scales from 100 to 8000 m, each identical to the null but with the additional landcover predictor (Fig. 1, Supporting information). For each landcover, the best performing model was then tested for significance.

To test landcover predictors for significance, we compared the best-performing, or 'scale-optimized,' model for that

Table 1. The number of times each candidate null model was selected. Possible covariates were day of year (day) and latitude (lat), either as a linear or quadratic predictor. Every model included site nested within study as a random intercept. For each species, we used the model with the lowest AICc as the null against which we compared the landcover models. Landcover models used the same covariates as the null, but with the additional landcover term.

Candidate model	Frequency (no. species)
Intercept only	6
day	6
$day + day^2$	12
lat	4
$lat + lat^2$	8
day + lat	14
$day + day^2 + lat$	4
$day + lat + lat^2$	11
$day + day^2 + lat + lat^2$	19

landcover to that species' null model using a likelihood ratio test. In doing so, we adjusted our p-values and significance level to keep our false discovery rate at 5% (next section). If a landcover predictor was significant, the optimal scale for that landcover is taken as the scale of effect for that species–landcover association. If a species is a broad generalist, with little or no preference for one landcover type or another, it would be expected to have no significant landcover associations.

To test for an association between species' scales of effect and mobility-related traits, we used a linear mixed-effects model. First, we filtered our landcover associations to only those that were deemed 'scale sensitive,' which we defined as cases where models at the best scale outperformed the worst scale by ΔAICc $\geq$ 7 (Jackson and Fahrig 2012). We then modeled species' scales of effect (natural-log transformed) as a function of body size (natural-log transformed) and sociality (eusocial, facultatively social, or solitary), using Gaussian error and including species as a random intercept term. This model describes species' mean scale of effect as a function of the mobility-related traits, while the random intercept accounts for variation within species, which we include because some species have more than one scale of effect. However, because it is not ideal to estimate random effects with only one or few observations per group (species, in this case), we also ran an analogous simple linear model using the mean scale of effect per species. The results were qualitatively identical, so we refer in Results to the mixed-effects model.

To assess the consequences of relying on unsupported, a priori assumptions of species' scales of effect, we repeated our analyses using two alternate approaches. First, we assessed species' landcover associations while assuming a universal 1000-m scale of effect (a common approach; e.g. Collado et al. 2019). Second, we used an explicit but naïve application of the mobility hypothesis, and assumed scales of effect equal to a species' estimated 'typical homing distance,' estimated from body size following Greenleaf et al. (2007). In both cases, we maintained false discovery rates at 5%. To assess the consequence of taking these naïve approaches, we compared the number of landcover associations detected here versus in our primary analysis in which scales of effect were determined for each species–landcover association separately. See the Supporting information for more details.

To compare landcover associations among species with different phenologies, we tallied the frequency of positive and negative associations between spring, summer, and long-season bees and different landcover types. Given broad patterns in flower phenology in our study region, we focused on differences between forested versus open landcovers, in which flowers bloom primarily in spring versus summer, respectively. To test if each group of bees (spring, summer, or long-season bees) was more likely associated with one landcover type or the other, we tested for differences in the frequency of positive or negative responses to forest and open habitats using two-sided binomial tests. For example, we tested whether spring bees had positive associations more often with forested versus open landcovers. Here, the number of 'successes' were the number of positive associations with forest, and the number of 'trials' were the total number of positive associations.

Controlling the false discovery rate

Because our analysis involved so many models (84 species × 9 landcovers × 80 scales = 60 480 models), we would expect to find a large number of spurious species–landcover associations (i.e. type I errors, or false discoveries) just by chance. We consider false discoveries at the level of the species–landcover association. The false discovery rate (γ) is the proportion of significant landcover associations that are actually type I errors (Pike 2011).

In testing for significance, we adjusted our critical p-value to control our false discovery rate at $\gamma = 0.05$. To do so, we followed a two-step process. First, we calculated p-values for each species–landcover association using a randomization-based null distribution. We used a randomization-based null because, given our scale optimization process, we expected unimportant/null predictors to outperform the typical χ^2 null distribution (Supporting information). This null distribution describes the expected performance of landscape predictors under the null hypothesis of no association, given the structure of our data and the scale-optimization procedure. Second, we determined the critical p-value following the step-up procedure of Benjamini et al. (2006), which resulted in a critical p-value of 0.015. That is, we considered models as significant if $p \leq 0.015$.

Results

In total, we found 193 significant species–landcover associations across 60 (of 84) species (Fig. 2). These models outperformed their respective nulls by ΔAICc = 7.45 to 38.6 (Supporting information). Of the 193 significant models, 183 were scale sensitive, meaning the optimal scale outperformed the worst scale by ΔAICc $\geq$ 7. Across these 183 scale-sensitive landcover associations, the distribution of scales of effect was very right skewed (Supporting information). The most common scale of effect (the mode) was 200 m

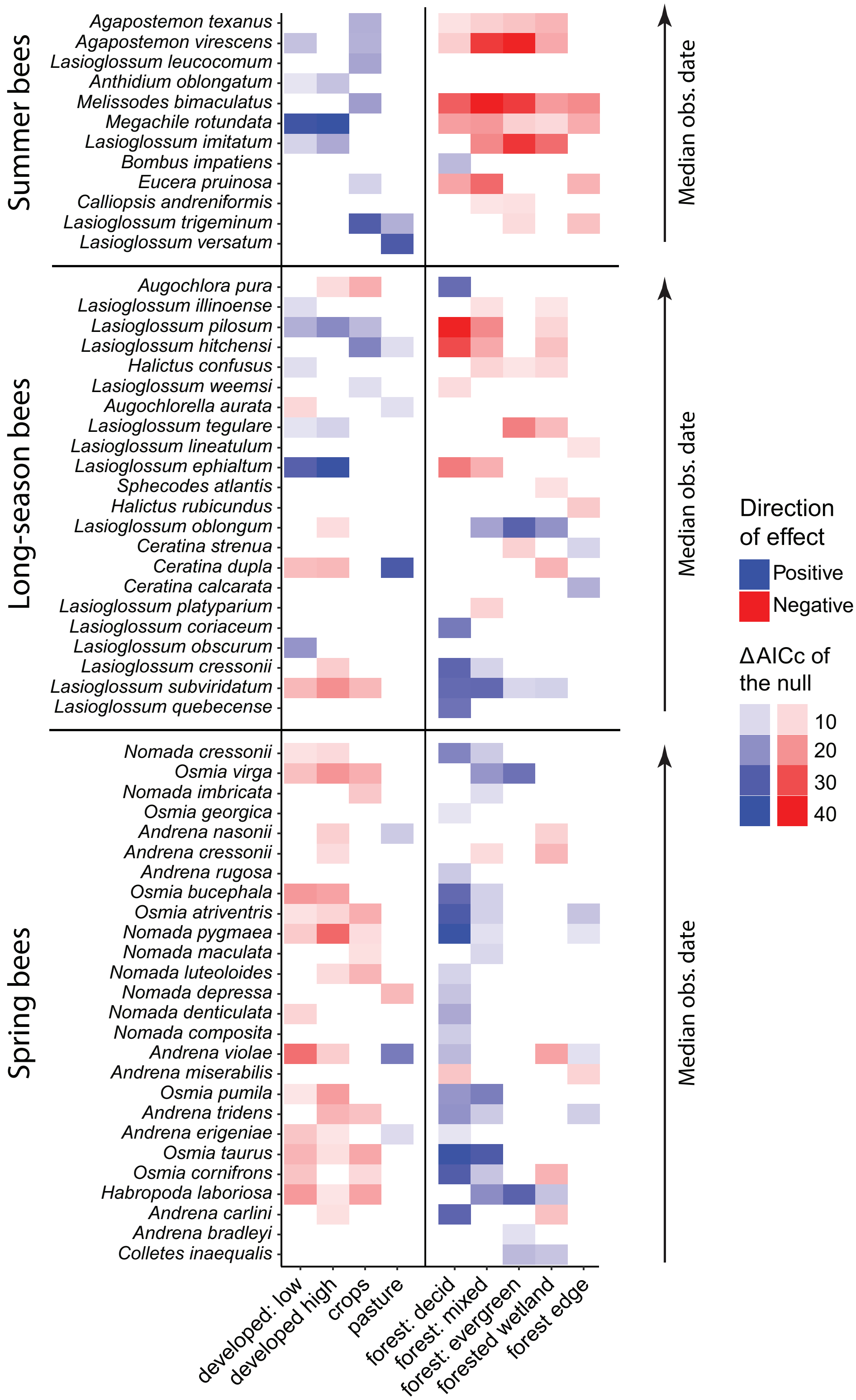


Figure 2. Species–landcover associations detected in our analyses. Colored cells denote significant associations. Species are arranged first by phenological group (spring, summer, or long-season) then, within each group, by median observation date; species at the bottom fly earlier, species towards the top fly later. This arrangement makes the phenological pattern we observed clear: spring bees are much more likely to be associated positively with forest and negatively with open habitats, while summer bees show the opposite. Notably, within long-season bees, who showed no preference as a group, earlier flying bees appeared more like spring bees and later flying bees appeared more like summer bees.

and the median was 1000 m. Unexpectedly, though, we did detect scales of effect up to the maximum measured distance of 8000 m, even for small bodied species like *Ceratina dupla*. Also noteworthy, species often appeared to have different scales of effect for different landcover types (Fig. 3). For example, *Agapostemon virescens* was associated (negatively) with deciduous forest at 300 m but (positively) with development at 4000 m.

Q1: The mobility hypothesis

Scales of effect were not associated with mobility-related traits (Fig. 4). Among the 183 scale-sensitive landcover associations, scales of effect were not significantly associated with body size nor sociality (Wald χ^2 tests, $p \geq 0.52$, $R^2 = 0.025$). Our results thus do not support the mobility hypothesis.

Q2: Consequences of a priori assumptions

Relying on a priori assumptions, rather than explicitly estimating species' scales of effect would have resulted in many undetected species–landcover interactions. Assuming a 1000-m scale of effect resulted in only 64 significant species–landcover associations across 31 (of 84) species. Assuming scale of effect equal to estimated 'typical homing distance' resulted in 61 significant models across 29 (of 84) species. These are compared to our scale-optimized approach that yielded 193 significant models across 60 (of 84) species.

Q3: Bee–landcover associations

While no landcover had universally positive or negative associations with bees, some tended to be positive or negative more often than others (Supporting information). Among our nine tested landcovers, two had significantly more positive than negative responses: deciduous forest (23 positive versus 10 negative; binomial test $p = 0.035$) and pasture (8 positive versus 1 negative, $p = 0.039$), while another two had significantly more negative than positive responses: forested wetland (17 negative versus 4 positive; $p = 0.007$) and high intensity development (20 negative versus 6 positive; $p = 0.009$).

Species' landcover associations were strongly associated with phenology. Spring bees were significantly more likely to be positively associated with forests and negatively associated with open, non-forested landcovers, while summer bees were more likely to be positively associated with open, non-forested landcovers and negatively associated with forests (binomial tests, all $p < 0.001$; Supporting information). Long-season bees, at least as a group, did not differ in the frequency of their associations with forested versus open landcovers ($p > 0.05$).

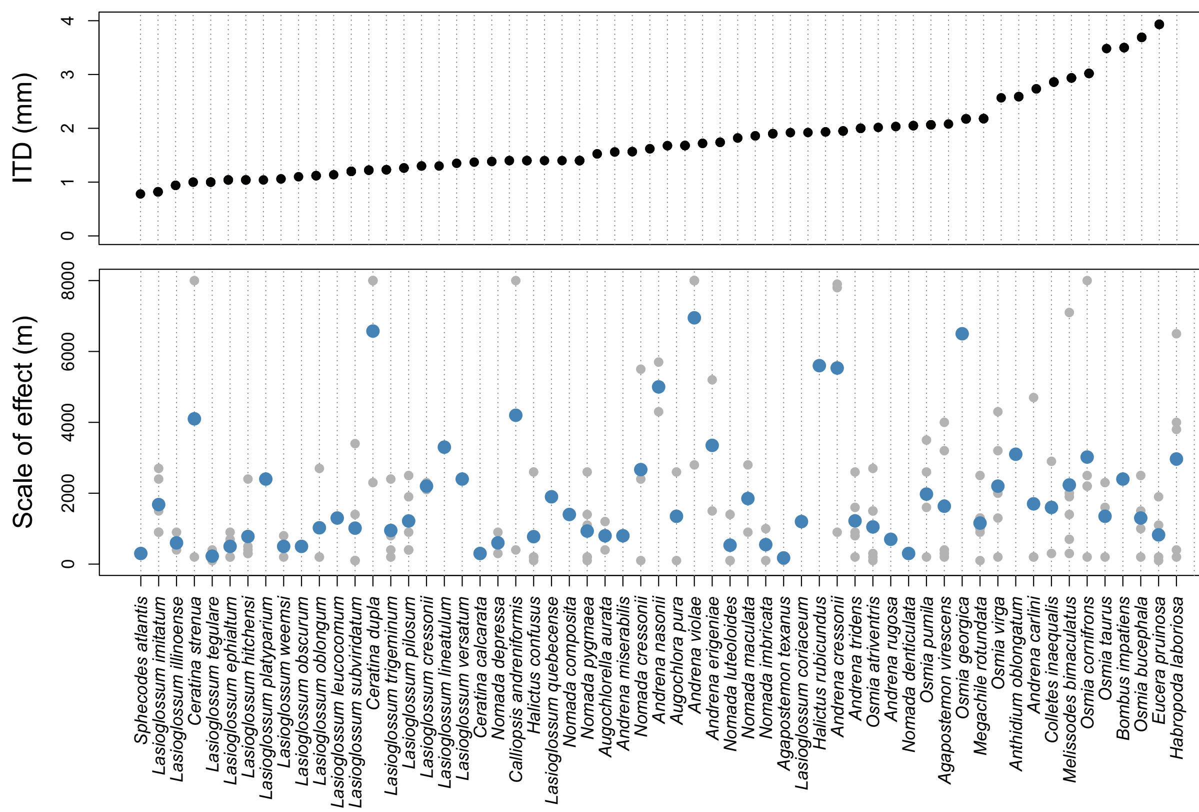


Figure 3. Body size (as intertegular distance, or ITD) and scales of response for the 60 bee species with significant landcover associations. Each species–landcover association was assessed independently, so species could be assigned different scales of response to different landcovers. In the bottom panel, the grey dots represent scales of effect for each species' individual landcover associations, and the larger, blue dots represent mean scale of effect across landcovers.

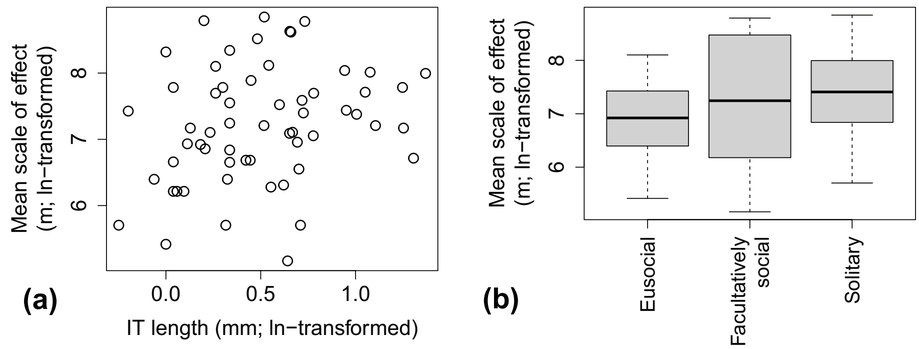


Figure 4. (a) Mean scale of effect for each bee species (n = 60) against body size, measured as intertegular distance. (b) Mean scales of response by sociality. Species' scales of effect were not significantly associated with either of these mobility-related traits.

Discussion

Despite the importance of understanding species' scales of effect to effective study design or effective conservation, what determines scales of effect remains poorly understood (Miguet et al. 2016, Arroyo-Rodríguez et al. 2023). Here, we tested the most popular explanation for what determines species' scales of effect, which we call the 'mobility hypothesis.' In doing so, we also assess the landscape-scale habitat associations of 84 wild bee species, another subject about which surprisingly little is known. Here, we identified 193 landcover associations across 60 bee species, with scales of effect ranging from 100 to 8000 m – across the entire range of scales that we tested. We report two important results. First, we found no support for the mobility hypothesis. Instead, bees' scales of effect were idiosyncratic, varying among landcovers even within species, and being unrelated to body size or sociality, as predicted by the mobility hypothesis. Second, no landcover was universally good or bad for bees; instead, which landcovers are most important varies by bee species and across seasons.

Scales of effect and the mobility hypothesis

We used wild bees to test the 'mobility hypothesis,' which predicts that species with greater dispersal or foraging distances will have larger scales of effect. Despite its equivocal support in the literature, the mobility hypothesis has been the default assumption for many studies, including our own (Winfree and Dushoff 2005, Harrison et al. 2018, Smith et al. 2021). Specifically, our assumption has been that bee species' scales of effect correspond to their flight capacity.

We found that bee species' scales of effect were not associated with body size or sociality (Fig. 3), which are the strongest known predictors of bee species' flight capacity (Greenleaf et al. 2007, Kendall et al. 2022). Our results thus join Arroyo-Rodriguez et al.'s (2023) review in rebuke of the mobility hypothesis as a meaningful explanation for species' scales of effect.

This rebuke of the mobility hypothesis has important implications for how landcover studies are conducted. To illustrate, consider that our alternate methods, relying on a priori assumptions of species' scales of effect, each led to roughly one-third the number of detected landcover associations. This demonstrates that not optimizing scales of effect can lead to massive (in our case, ~ 66%) type II error – that is, failing to detect landcover associations. This is critical: across taxa, a review found less than 5% of landcover selection studies were scale optimized (McGarigal et al. 2016). Our findings suggest that such an approach could lead to researchers missing many important results.

We see two possible explanations for why the mobility hypothesis was not supported, which need not be mutually exclusive, nor exclusive to bees. First, the mobility hypothesis could be incorrect, meaning a species' mobility does not determine its scales of effect. Instead, scales of effect may depend on different species- or population-level traits. For example, scales of effect may decrease with increasing population growth rate or population density, both of which should increase the importance of local population processes relative to immigration from adjacent areas (Miguet et al. 2016). Second, it could be that the mobility hypothesis *is* correct, meaning a species' mobility *does* determine its scales of effect, but body size and sociality are insufficient proxies for mobility. These traits are meant as predictors of foraging distance, which is only one aspect of mobility. Scales of effect may depend instead – or additionally – on other mobility-related traits like intergenerational dispersal (Jackson and Fahrig 2012) about which little is known in bees. Or it could be that mobility itself is context dependent. That is, populations may differ in their capacity or propensity to move across the landscape due to trait variation or landscape context (Arroyo-Rodríguez et al. 2023). In particular, we think dispersal and foraging distances could both depend on the arrangement or quality of surrounding habitats. For example, bees are likely to forage further in more fragmented or resource-poor landscapes because resources are spread across a broader area (Minckley et al. 1994, Wood et al. 2015, Miguet et al. 2016). Conversely, bees in resource-rich areas may have no reason to fly long distances even if they have the ability (Seeley 1985, Cresswell et al. 2000, Bartomeus and Winfree 2011). This sort of informed movement could weaken or remove the relationship between trait averages and scales of effect. This idea is supported by the simulation of Jackson and Fahrig (2012), who found that allowing informed dispersal greatly reduced the effect of dispersal distance on the simulated species' scale of effect.

We should also acknowledge the variability we found in species' scales of effect. This variability occurred at two levels. First, for individual species–landcover associations, there was often uncertainty in the strongest scale of effect. This uncertainty appeared as broad and sometimes bimodal distributions of AICc weight across competing scales (Supporting information). Second, species' strongest scales of effect usually varied among landcover types (Fig. 3). While it is not usually discussed, both these types of variation are actually typical of this sort of analysis (Findlay and Houlahan 1997, Holland et al. 2004, Stuber et al. 2018, Stuber and Fontaine 2019, Moll et al. 2020). Some of this variation could result from the correlation of landcover proportion across scales, which leads to imprecision and thus statistical uncertainty. But it could also be due to true biological variation. For example, context dependency (e.g. a local population's scale of effect depending on characteristics of its particular landscape) could create uncertainty in choosing a single scale of effect for populations spread across a large area. Similarly, if a habitat affects a population through multiple processes occurring at different scales (Addicott et al. 1987), this could create a bimodal pattern in which a particular species appears to respond to one habitat type at multiple scales. Regardless of its source, such variation and uncertainty in estimated scales of effect make it difficult to precisely characterize or accurately predict species' scales of effect. It also makes us hesitant to over-interpret small differences in scales of effect between species or between landcover associations within species.

Bee–landcover associations

To date, most studies of bee–landcover associations have been limited in at least one of two ways. First, most such studies have focused on entire bee communities, measuring total abundance or diversity, rather than assessing the needs or responses of individual species (Kennedy et al. 2013, Baldock et al. 2015). Second, bee landcover studies tend to focus on contrasting natural and anthropogenic landcovers (Steffan-Dewenter 2003, Benjamin et al. 2014, Baldock et al. 2015, Smith et al. 2021), which usually means grouping natural or semi-natural landcovers together in contrast to agriculture and/or urban development. As a result, the distinctions within these broader groups are lost.

Here, we use nine natural and anthropogenic landcover types to measure landcover associations of 84 species. We found significant landcover associations for 60 of 84 (71%) of species (Fig. 2). For the remaining 24 species, the lack of significant landcover associations could be seen as indication of broad generalization. We expect, however, that it is due at least as much to sample size (and thus lack of statistical power); species with high observed abundance or site occupancy *all* had significant associations and only species with low abundance or occupancy had no significant associations (Supporting information).

Our results highlight important distinctions among landcover types that are often grouped together. In particular, among bees positively associated with forest, most were associated with deciduous forest while few were positively associated with mixed or evergreen forest and many bees had negative responses to forested wetland. Similarly, among bees associated with open landcovers, species were typically associated with urbanization or agriculture, but not both.

Our results show that bee–landcover associations are themselves associated with phenology. Spring bees are significantly more likely to be associated with forests, while summer bees are more likely to be associated with open landcovers. This builds on prior work in the northeastern USA showing that the forest bee community flies earlier in the year than the community inhabiting urban and agricultural landscapes (Harrison et al. 2018), and that about a third of bee species in this region are significantly more abundant in forest than anthropogenic habitats, broadly (Smith et al. 2021). But, by using a larger dataset and more finely distinguished landcover types, we were able to better identify species' specific landcover associations and relate these directly to phenology.

Both these results have important implications for bee conservation. At least until recently, bee habitat conservation has focused almost exclusively on open, meadow- or prairie-like habitats with summer-blooming flowers (Durant and Otto 2019, Baldock 2020, Quinlan et al. 2021). This focus on meadows is the result of bee habitat studies being conducted primarily in the summertime, when open meadows would indeed be the best bee habitat. Our results show, however, that by overlooking forests, current conservation practices have overlooked the entire spring bee community (Wood et al. 2018, Mola et al. 2021, Urban-Mead et al. 2021, 2023). Critically, this spring-forest bee community is just as abundant and diverse as the communities found in more open habitats in the summer (Harrison et al. 2018). Thus, bee conservation must expand its scope.

Conclusions

We see two big takeaways from our study. First, the scales at which species respond to their environment are idiosyncratic and hard to predict. We thus argue that scales of effect should not be assumed; this approach risks missing important results. Instead, scales of effect should be explicitly and empirically estimated for the focal species and landscape.

Second, while open habitats support many bee species, these are but a subset of the broader bee community. Our results thus add to a nascent but growing literature demonstrating the importance of forests especially, but also habitat diversity generally, for supporting bee diversity (Urban-Mead et al. 2021, Eckerter et al. 2022, Chase et al. 2023, Ulyshen et al. 2023, Urban-Mead et al. 2023).

Acknowledgements – We thank Tina Harrison, who collected many of the data included in our analysis. We also thank James Reilly and other members of the Winfree Lab Group for their input throughout. Data collection was supported by the USDA McIntire Stennis program (Accession 1014396) and the US Fish and Wildlife Service.

Author contributions

Dylan T. Simpson: Conceptualization (lead); Formal analysis (lead); Visualization (lead); Writing – original draft (lead). **Colleen Smith**: Methodology (equal); Writing – review and editing (supporting). **Rachael Winfree**: Conceptualization (supporting); Methodology (equal); Supervision (lead); Writing – review and editing (supporting).

Transparent peer review

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/ecog.07982.

Data availability statement

Data are available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.ngf1vhj55 (Simpson et al. 2025).

Supporting information

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

References

Addicott, J. F., Aho, J. M., Antolin, M. F., Padilla, D. K., Richardson, J. S. and Soluk, D. A. 1987. Ecological neighborhoods: scaling environmental patterns. – Oikos 49: 340.

Arroyo-Rodríguez, V., Martínez-Ruiz, M., Bezerra, J. S., Galán-Acedo, C., San-José, M. and Fahrig, L. 2023. Does a species' mobility determine the scale at which it is influenced by the surrounding landscape pattern? – Curr. Landscape Ecol. Rep. 8: 23–33.

Baldock, K. C. 2020. Opportunities and threats for pollinator conservation in global towns and cities. – Curr. Opin. Insect Sci. 38: 63–71.

Baldock, K. C. R., Goddard, M. A., Hicks, D. M., Kunin, W. E., Mitschunas, N., Osgathorpe, L. M., Potts, S. G., Robertson, K. M., Scott, A. V., Stone, G. N., Vaughan, I. P. and Memmott, J. 2015. Where is the UK's pollinator biodiversity? The importance of urban areas for flower-visiting insects. – Proc. R. Soc. B 282: 20142849.

Bartomeus, I. and Winfree, R. 2011. The Circe principle: are pollinators waylaid by attractive habitats? – Curr. Biol. 21: R652–R654.

Bartomeus, I., Ascher, J. S., Wagner, D., Danforth, B. N., Colla, S., Kornbluth, S. and Winfree, R. 2011. Climate-associated phenological advances in bee pollinators and bee-pollinated plants. – Proc. Natl Acad. Sci. USA 108: 20645–20649.

Bartomeus, I., Cariveau, D. P., Harrison, T. and Winfree, R. 2018. On the inconsistency of pollinator species traits for predicting either response to land-use change or functional contribution. – Oikos 127: 306–315.

Barton, K. 2022. MuMIn: multi-model inference. – https://CRAN.R-project.org/package=MuMIn.

Benjamin, F. E., Reilly, J. R. and Winfree, R. 2014. Pollinator body size mediates the scale at which land use drives crop pollination services. – J. Appl. Ecol. 51: 440–449.

Benjamini, Y., Krieger, A. M. and Yekutieli, D. 2006. Adaptive linear step-up procedures that control the false discovery rate. – Biometrika 93: 491–507.

Brennan, J. M., Bender, D. J., Contreras, T. A. and Fahrig, L. 2002. Focal patch landscape studies for wildlife management: optimizing sampling effort across scales. – In: Liu, J. and Taylor, W. W. (eds), Integrating landscape ecology into natural resource management, 1st edn. – Cambridge Univ. Press, pp. 68–91.

Brooks, M. E., Kristensen, K., Benthem, K. J. van, Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Mächler, M. and Bolker, B. M. 2017. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. – R J. 9: 378–400.

Cameron, S. A. and Sadd, B. M. 2020. Global trends in bumble bee health. – Annu. Rev. Entomol. 65: 209–232.

Cariveau, D. P., Nayak, G. K., Bartomeus, I., Zientek, J., Ascher, J. S., Gibbs, J. and Winfree, R. 2016. The allometry of bee proboscis length and its uses in ecology. – PLoS One 11: e0151482.

Carrié, R., Andrieu, E., Cunningham, S. A., Lentini, P. E., Loreau, M. and Ouin, A. 2017. Relationships among ecological traits of wild bee communities along gradients of habitat amount and fragmentation. – Ecography 40: 85–97.

Chase, M. H., Charles, B., Harmon-Threatt, A. and Fraterrigo, J. M. 2023. Diverse forest management strategies support functionally and temporally distinct bee communities. – J. Appl. Ecol. 60: 2375–2388.

Collado, M. Á., Sol, D. and Bartomeus, I. 2019. Bees use anthropogenic habitats despite strong natural habitat preferences. – Divers. Distrib. 25: 924–935.

Cresswell, J. E., Osborne, J. L. and Goulson, D. 2000. An economic model of the limits to foraging range in central place foragers with numerical solutions for bumblebees. – Ecol. Entomol. 25: 249–255.

Desaegher, J., Ouin, A. and Sheeren, D. 2022. How far is enough? Prediction of the scale of effect for wild bees. – Ecography 2022: e05758.

Durant, J. L. and Otto, C. R. V. 2019. Feeling the sting? Addressing land-use changes can mitigate bee declines. – Land Use Policy 87: 104005.

Eckerter, T., Braunisch, V., Buse, J. and Klein, A. M. 2022. Open forest successional stages and landscape heterogeneity promote wild bee diversity in temperate forests. – Conservat Sci. Pract. 4: e12843.

Elton, C. 1927. Animal ecology. – Sigwick and Jackson Limited.

Findlay, C. S. and Houlahan, J. 1997. Anthropogenic correlates of species richness in southeastern Ontario wetlands: correlativos antropogénicos de la riqueza de especies en humedales del sureste de Ontario. – Conserv. Biol. 11: 1000–1009.

Gallo, T., Lehrer, E. W., Fidino, M., Kilgour, R. J., Wolff, P. J. and Magle, S. B. 2018. Need for multiscale planning for conservation of urban bats. – Conserv. Biol. 32: 638–647.

Goulson, D., Nicholls, E., Botías, C. and Rotheray, E. L. 2015. Bee declines driven by combined stress from parasites, pesticides, and lack of flowers. – Science 347: 1255957.

Greenleaf, S. S., Williams, N. M., Winfree, R. and Kremen, C. 2007. Bee foraging ranges and their relationship to body size. – Oecologia 153: 589–596.

Grinnell, J. 1924. Geography and evolution. – Ecology 5: 225–229.

Hanski, I. 2011. Habitat loss, the dynamics of biodiversity, and a perspective on conservation. – Ambio 40: 248–255.

Harrison, T., Gibbs, J. and Winfree, R. 2018. Forest bees are replaced in agricultural and urban landscapes by native species with different phenologies and life-history traits. – Global Change Biol. 24: 287–296.

Hesselbarth, M. H. K., Sciaini, M., With, K. A., Wiegand, K. and Nowosad, J. 2019. landscapemetrics: an open-source R tool to calculate landscape metrics. – Ecography 42: 1648–1657.
Holland, J. D., Bert, D. G. and Fahrig, L. 2004. Determining the spatial scale of species' response to habitat. – BioScience 54: 227.
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. 2015. Are ecologists conducting research at the optimal scale? Is research conducted at optimal scales? – Global Ecol. Biogeogr. 24: 52–63.
James, F. C., Johnston, R. F., Wamer, N. O., Niemi, G. J. and Boecklen, W. J. 1984. The Grinnellian niche of the wood thrush. – Am. Nat. 124: 17–47.
Kendall, L. K., Mola, J. M., Portman, Z. M., Cariveau, D. P., Smith, H. G. and Bartomeus, I. 2022. The potential and realized foraging movements of bees are differentially determined by body size and sociality. – Ecology 103: e3809.
Kennedy, C. M. et al. 2013. A global quantitative synthesis of local and landscape effects on wild bee pollinators in agroecosystems. – Ecol. Lett. 16: 584–599.
Klein, A.-M., Cunningham, S. A., Bos, M. and Steffan-Dewenter, I. 2008. Advances in pollination ecology from tropical plantation crops. – Ecology 89: 935–943.
Martin, A. E. and Fahrig, L. 2012. Measuring and selecting scales of effect for landscape predictors in species–habitat models. – Ecol. Appl. 22: 2277–2292.
McCauley, D. J., Pinsky, M. L., Palumbi, S. R., Estes, J. A., Joyce, F. H. and Warner, R. R. 2015. Marine defaunation: animal loss in the global ocean. – Science 347: 1255641.
McGarigal, K., Wan, H. Y., Zeller, K. A., Timm, B. C. and Cushman, S. A. 2016. Multi-scale habitat selection modeling: a review and outlook. – Landscape Ecol. 31: 1161–1175.
Miguet, P., Jackson, H. B., Jackson, N. D., Martin, A. E. and Fahrig, L. 2016. What determines the spatial extent of landscape effects on species? – Landscape Ecol. 31: 1177–1194.
Minckley, R. L., Wcislo, W. T., Yanega, D. and Buchmann, S. L. 1994. Behavior and phenology of a specialist bee (*Dieunomia*) and sunflower (*Helianthus*) pollen availability. – Ecology 75: 1406–1419.
Mola, J. M., Hemberger, J., Kochanski, J., Richardson, L. L. and Pearse, I. S. 2021. The importance of forests in bumble bee biology and conservation. – BioScience 71: 1234–1248.
Moll, R. J., Cepek, J. D., Lorch, P. D., Dennis, P. M., Robison, T. and Montgomery, R. A. 2020. At what spatial scale(s) do mammals respond to urbanization? – Ecography 43: 171–183.
Novotny, J. L., Reeher, P., Varvaro, M., Lybbert, A., Smith, J., Mitchell, R. J. and Goodell, K. 2021. Bumble bee species distributions and habitat associations in the midwestern USA, a region of declining diversity. – Biodivers. Conserv. 30: 865–887.
Pardee, G. L. and Philpott, S. M. 2014. Native plants are the bee's knees: local and landscape predictors of bee richness and abundance in backyard gardens. – Urban Ecosyst. 17: 641–659.
Pike, N. 2011. Using false discovery rates for multiple comparisons in ecology and evolution. – Methods Ecol. Evol. 2: 278–282.
Quinlan, G. M., Milbrath, M. O., Otto, C. R. V. and Isaacs, R. 2021. Farmland in U.S. conservation reserve program has unique floral composition that promotes bee summer foraging. – Basic Appl. Ecol. 56: 358–368.
Ricci, B., Franck, P., Valantin-Morison, M., Bohan, D. A. and Lavigne, C. 2013. Do species population parameters and landscape characteristics affect the relationship between local population abundance and surrounding habitat amount? – Ecol. Complexity 15: 62–70.
Seeley, T. D. 1985. Honeybee ecology. – Princeton Univ. Press.
Simpson, D. T., Smith, C. and Winfree, R. 2025. Data from: Wild bees and landcover: bee species' body size does not predict the scale of effect, but bee phenology predicts association with landcover type. – Dryad Digital Repository, https://doi.org/10.5061/dryad.ngf1vhj55.
Smith, C., Harrison, T., Gardner, J. and Winfree, R. 2021. Forest-associated bee species persist amid forest loss and regrowth in eastern North America. – Biol. Conserv. 260: 109202.
Steffan-Dewenter, I. 2003. Importance of habitat area and landscape context for species richness of bees and wasps in fragmented orchard meadows. – Conserv. Biol. 17: 1036–1044.
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., Gruber, F. and Fontaine, J. J. 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.
Ulyshen, M., Urban-Mead, K. R., Dorey, J. B. and Rivers, J. W. 2023. Forests are critically important to global pollinator diversity and enhance pollination in adjacent crops. – Biol. Rev. 98: 1118–1141.
Urban-Mead, K. R., Muñiz, P., Gillung, J., Espinoza, A., Fordyce, R., Van Dyke, M., McArt, S. H. and Danforth, B. N. 2021. Bees in the trees: diverse spring fauna in temperate forest edge canopies. – For. Ecol. Manage. 482: 118903.
Urban-Mead, K. R., Van Dyke, M., Muñiz, P., Young, A. D., Danforth, B. N. and McArt, S. H. 2023. Early spring orchard pollinators spill over from resource-rich adjacent forest patches. – J. Appl. Ecol. 60: 553–564.
Wickham, J., Stehman, S. V., Sorenson, D. G., Gass, L. and Dewitz, J. A. 2023. Thematic accuracy assessment of the NLCD 2019 land cover for the conterminous United States. – GIsci. Remote Sens. 60: 2181143.
Winfree, R. and Dushoff, J. 2005. Testing simple indices of habitat proximity. – Am. Nat. 165: 707–717.
Winfree, R., Griswold, T. and Kremen, C. 2007. Effect of human disturbance on bee communities in a forested ecosystem. – Conserv. Biol. 21: 213–223.
Winfree, R., Bartomeus, I. and Cariveau, D. P. 2011. Native pollinators in anthropogenic habitats. – Annu. Rev. Ecol. Evol. Syst. 42: 1–22.
Winfree, R., Williams, N. M., Dushoff, J. and Kremen, C. 2014. Species abundance, not diet breadth, drives the persistence of the most linked pollinators as plant–pollinator networks disassemble. – Am. Nat. 183: 600–611.
Wood, T. J., Holland, J. M., Hughes, W. O. H. and Goulson, D. 2015. Targeted agri-environment schemes significantly improve the population size of common farmland bumblebee species. – Mol. Ecol. 24: 1668–1680.
Wood, T. J., Gibbs, J., Rothwell, N., Wilson, J. K., Gut, L., Brokaw, J. and Isaacs, R. 2018. Limited phenological and dietary overlap between bee communities in spring flowering crops and herbaceous enhancements. – Ecol. Appl. 28: 1924–1934.
Yeiser, J. M., Morgan, J. J., Baxley, D. L., Chandler, R. B. and Martin, J. A. 2021. Optimizing conservation in species-specific agricultural landscapes. – Conserv. Biol. 35: 1871–1881.

ECOGRAPHY Open Access
A JOURNAL OF SPACE AND TIME IN ECOLOGY

Supporting Information

Habitat associations of wild bees: on the importance of phenology and the idiosyncrasy of scale

Dylan Simpson, Colleen Smith, and Rachael Winfree

Contents

Bee abundance data

This analysis used a dataset amalgamated from five previous studies, each of which collected bees by pan or vane trap. Following is a brief summary of each of these studies. Data from three of these studies were previously published, and the publications are noted below, while data from the other two is being published here for the first time.

Dataset 1: Pinelands (Winfree et al. 2007)

This study was designed to ask after the effects of human land use on bee communities. The study region was the Pine Barrens of southern New Jersey. There were 44 sites placed along a human land use gradient. In this region, the natural landcover is predominantly forested ericaceous heath, and human landcover is predominantly agriculture (blueberry, cranberry) and suburbs.

The 44 sites were visited 2-5 times each in 2003, for a total of 167 site-visits. At each site visit, 44 pan traps (plastic bowls painted with white or fluorescent blue or yellow) were placed along a 110-m transect and left for 8 hours between 07:00 and 17:00. Data were only collected on sunny or partly sunny days.

Of our 84 focal species, 59 were detected in this study, represented by 1470 specimens.

Dataset 2: Biotic homogenization (Harrison et al. 2018a, b)

This study was designed to assess the role of human land use in biotic homogenization across space and ecoregion. The study region was New Jersey, New York, and Pennsylvania. There were 36 sites in a nested block design, with three blocks of three sites each nested within four ecoregions. Each block had one site embedded in each of three dominant landcover types - agriculture, (sub)urban, and forest – while local habitat was standardized as mown grass.

Sites were visited 1-6 times per year, from spring to autumn, in the years 2013-2015. Across sites and years, there were a total of 377 site visits. At each site visit, 36 pan traps and two vane traps were set out and left for 24 hours.

Of our 84 focal species, 83 were detected in this study, represented by 11923 specimens.

Dataset 3: SWG (previously unpublished)

This study was designed to examine differences in bee abundance, diversity, and community composition among different habitat types in New Jersey. There were 37 sites visited a total of 79 times between March and September of 2016. Sites were haphazardly located across the state, with representation of many major landcover and habitat types, including different forest types, crops, sub/urbanization, and near wetlands. Each site was visited at least twice, once during the spring and once during the summer. At each site visit, 35 white, blue, and yellow pan traps were placed along a 50-m transect from for 5-7 hours between 08:00 and 15:00.

Of our 84 focal species, 77 were detected in this study, represented by 4154 specimens.

Dataset 4: Forests 1 (Smith et al. 2021)

This study was designed to assess the effects of forest age, area, and fragmentation on bee communities. The study region was the Piedmont ecoregion of New Jersey. There were 32 sites, all embedded within forests but with varying landscape contexts.

Sites were visited 2-4 times in 2017 and 2018, except 5 sites that were only visited in 2017. At each site visit, 39 pan traps were placed in a 40 m x 100 m grid. In 2018, four vane traps were also placed. Traps were left for ca. 8 hours between 05:30 and 20:00.

Of our 84 focal species, 64 were detected in this study, represented by 10550 specimens.

Dataset 5: Forests 2 (Winfree et al. 2014)

This study was designed to assess the effects of forest vs. non-forest “matrix” habitat on bee communities. The study region was the Piedmont region of New Jersey. There were 16 forest sites embedded within forest fragments of differing size, or within sub/urban or agricultural matrix.

Sites were visited 4 times each in April and May of 2006. On each site visit, an array of 39 white, yellow, and blue pan traps were placed and left for 4 hours.

Of our 84 focal species, 55 were detected in this study, represented by 2279 specimens.

Bee flight windows and phenological classifications

To estimate flight windows of the 84 focal species, we took the 5th and 95th percentiles of observation date across all specimens in a large, amalgamated dataset. These data included all specimens collected in this study region by the Winfree Lab group at Rutgers University, as well as natural history collections from the American Museum of Natural History that were collected in New Jersey, Pennsylvania and New York. We took the inner 90th percentile of observation dates because our goal was to define the period in which these bees are most typically active and likely to be observed. Including all collection dates would mean including outliers that are atypical and not representative of species’ general ecology.

Using these flight windows, we classified species as either spring-, summer-, or long-season bees. Spring bees started early and ended early, summer bees started late and ended late, and long-season bees started early but ended late. We decided cutoff dates for each category based on natural clustering in species’ start and end dates (red lines in Figure S3): spring bees’ flight window starts before May 13 and ends before July 10; summer bees’ start after May 13 and end after July 10; long-season bees start before May 13 and end after July 10. After initially choosing these dates based on the location of points in Figure S3, we checked whether species’ classifications followed our ecological intuition. In particular, we checked whether known early-flying bees like *Osmia* were indeed classified as spring bees, and whether known social (e.g. many *Lasioglossum* species) and multivoltine (e.g., *Augochlora pura*) species were classified as long-season bees. The only notable exception was *Bombus impatiens*, which we would have intuitively classified as a long-season bee, insofar as it is well known to emerge as nest-founding queens in the spring (our observations peak in April) but was instead classified by our statistical method as a summer bee. We would have classified it as a long season bee because its colonies are founded in the spring and are active until the fall. It was classified as a summer bee, though, because of its very high peak worker abundance and the fact that the vast majority of observations of workers, not queens. As a result, the inner 90th percentile will not include founding queens.

Most of the data for this phenology analysis were previously published; the primary references for these data are Bartomeus et al. 2011, Winfree et al. 2007, 2018, Winfree and Kremen 2009, Cariveau et al. 2013, Benjamin et al. 2014, MacLeod et al. 2016, Harrison et al. 2018a, Roswell et al. 2019, Smith et al. 2019, 2021, Genung et al. 2023.

Controlling the False Discovery Rate

In assessing model significance, we adjusted our critical p-value (i.e., our α value) to maintain a false discovery rate of γ = 0.05. To do so, there are two broad steps. First is to calculate observed p-values. Second is to use the distribution of observed p-values to calculate the critical p-value. Due to reasons explained below, we calculated observed p-values based on a numerically-derived null distribution. Then we determined the critical p-value following established procedures to control the false discovery rate (Benjamini et al. 2006).

In a typical analysis, calculating observed p-values is trivial because p-values are based on parametric expectations (e.g., the expected *t*, *F*, or χ^2 distributions). These distributions represent the expected distribution of model performance – as represented by the relevant test statistic – under the null hypothesis. In a linear model, this means the performance of a model in which the predictor is actually unrelated to the outcome. Here, we used likelihood ratio tests to compare models with and without landcover predictors. Typically, such a test would rely on the χ^2 distribution with one degree of freedom but, as we explain below, we did not think this appropriate in our analysis here.

Traditional parametric statistics were not appropriate for our analysis because our observed test statistics were not the result of a single independent statistical test. This is because the models we test for significance were first scale optimized. The scale optimization process is to compare 80 related models, in which the focal landcover predictor is measured at each of 80 scales. This means our tested models do not represent a single, independent statistical test but, instead, the best-performing of 80 tests. As a result, we expect models with unrelated predictors to "outperform" the typical χ^2 distribution (Figure S5). That said, these models are not fully independent – landcover proportion at 100 m is highly correlated with landcover proportion at 200 m, that at 200 m is highly correlated with that at 300 m, and so on. Thus, we also would not expect our models to perform as well as the best of 80 truly independent tests. The question then is how well *should* we expect our models to perform under the null hypothesis? Answering this question allows us to robustly assess model significance.

To determine how well we should expect our landcover models to perform under the null hypothesis, we generated fake landcover data that resembled our real landcover data in meaningful ways, then used these as predictors of bee abundance, just as in our real analysis. We did this repeatedly and used the distribution of model performance across iterations as our null distribution. We did this using 12 bee species as case studies. We did not perform this analysis on all 84 species because it was too computationally intensive. The species we used as case studies were: *Augochlora pura*, a species for which we had a strong *a priori* expectation as to its landcover associations and that we regularly used to test our code[1]; 10 species chosen at random using the slice_sample() function in the dplyr package in R; and the most abundant species in our dataset (*Ceratina calcarata*), which we added to help assess whether sample size affected performance under the null.

The key aspect of our landcover data we wanted to recreate was their cross-scale correlational structure. That is, the degree to which landcover data at each scale are correlated with data at each other scale. This correlation structure is what determines the relative independence vs. redundancy of the models run in the scale optimization procedure.

To create data that have the same cross-scale correlational structure as our observed landcover data, we used a multivariate random-normal[2] distributions (using the mvrnorm function in the MASS package in R). We then performed the same scale-optimization analysis with these fake data as in our main analysis, and used the performance of these models to define the null distribution against which to test our real models. We did this separately for each species because the sites included in each species' analysis could vary. This is because some sites may be excluded if they were only surveyed outside the phenological window of the focal species.

To measure the cross-scale correlations of our landcover data, we took the pair-wise correlation between a focal landcover measured at each of 80 scales to create a 'cross-scale correlation matrix.' We did this for each of our nine landcover classes. These correlation matrices were very similar across landcover types, and so each iteration we randomly chose one landcover to use as a template. We then used the correlation matrix for that template landcover as the covariance matrix for the multivariate-random normal landcover predictors. The result is multivariate random normal data with the same correlation structure as the template landcover. We then used these random data as predictors in our scale optimization procedure.

For each of the 12 bee species, we repeated this procedure 100 times. For each iteration, we compared the scale-optimized "landcover" model to the null model selected for that species and measured the performance of the model as natural log of their likelihood ratio (as in our main analysis). Null distributions were similar among species and did not appear to vary with sample size or occupancy rate. We therefore combined the null results for all 12 species and used the resulting 1,200 null likelihood ratios as our null distribution. We calculated p-values by comparing our observed, empirical likelihood ratios to this null distribution.

[1] *Augochlora pura* is a species known to nest in rotting logs of deciduous trees, so we had a strong *a priori* expectation that *A. pura* would respond positively to deciduous forests. We tested our code with a species with strong *a priori* expectations as a reality check that our analysis produced results that make sense. We used the same species repeatedly so that we could see if or how our analytical decisions affected results, but without seeing the full results so we could not bias ourselves towards an approach that produced more or particular significant results. Notably, during code and method development, the results for *A. pura* never qualitatively changed.

[2] Admittedly, our habitat data were not normally distributed. Instead, they tended to be right skewed. We did experiment with using random-log normal data for this analysis, but it seemed to make no difference in the expected performance of our models. For our final analysis, we kept the normal, rather than log normal, predictors because we were able to more closely control the cross-scale correlation structure to match our empirical data.

Finally, to control the false discovery rate at $\gamma = 0.05$, determined the critical p-value at which only an estimated 5% of significant results would be type I errors. To do so, we used the step up procedure described by Benjamini *et al.* (2006). This method is based on the observed distribution of p-values across many tests, and the deviation of this distribution from what would be expected if all models were actually null. We determined our critical p-value was 0.015. We thus considered models significant if their $p \leq 0.015$.

Results under different analytical approaches

To convey the importance of measuring species' scales of response instead of relying on assumptions, we demonstrate the consequences of using two alternate approaches. Both approaches are based on choosing a scale of effect *a priori*, which is a common approach in habitat selection studies generally (McGarigal et al. 2016). We expect the same problems that arise here will arise for other taxa besides bees. The first alternate approach we tested was to assess the landcover associations of all 84 bee species while measuring landcover proportion at a 1000-m radius only. One thousand meters is a common radius used in bee landcover studies, but is not special or unique and we chose it somewhat arbitrarily as an example. The second alternate approach was to estimate scales of effect for each species based on body size. Here, we use an equation that relates body size (as intertegular distance in mm) to 'typical homing range' (in m). The equation was originally published in Greenleaf et al. (2007).

To test each scenario, we used the same models we had already run in our first analysis. In the first scenario, we used the models in which landcover had been measured at 1000 m. In the second scenario, we calculated the scale of effect as the estimated 'typical homing distance' based on the equation in Greenleaf et al. (2007), then chose the model that most closely matched that scale. While we did not re-measure landcover proportion using the exact scale calculated from body size, these calculated scales were extremely well correlated with the scales we did use (Pearsons r = 0.99). In each case, we selected a single model and tested it against the null model. As in our main analysis, we tested for significance using likelihood ratio tests and adjusted our critical p-value to maintain a false discovery rate of 5%. However, because we were no longer performing scale optimization, we calculated p-values using the traditional χ^2 distribution, rather than our randomization-based null model.

References

Bartomeus, I., Ascher, J. S., Wagner, D., Danforth, B. N., Colla, S., Kornbluth, S. and Winfree, R. 2011. Climate-associated phenological advances in bee pollinators and bee-pollinated plants. - Proc. Natl. Acad. Sci. U.S.A. 108: 20645–20649.

Benjamin, F. E., Reilly, J. R. and Winfree, R. 2014. Pollinator body size mediates the scale at which land use drives crop pollination services (J Osborne, Ed.). - J Appl Ecol 51: 440–449.

Benjamini, Y., Krieger, A. M. and Yekutieli, D. 2006. Adaptive linear step-up procedures that control the false discovery rate. - Biometrika 93: 491–507.

Cariveau, D. P., Williams, N. M., Benjamin, F. E. and Winfree, R. 2013. Response diversity to land use occurs but does not consistently stabilise ecosystem services provided by native pollinators (D Kleijn, Ed.). - Ecol Lett 16: 903–911.

Cariveau, D. P., Nayak, G. K., Bartomeus, I., Zientek, J., Ascher, J. S., Gibbs, J. and Winfree, R. 2016. The Allometry of Bee Proboscis Length and Its Uses in Ecology (O Rueppell, Ed.). - PLoS ONE 11: e0151482.

Desaegher, J., Ouin, A. and Sheeren, D. 2022. How far is enough? Prediction of the scale of effect for wild bees. - Ecography 2022: e05758.

Genung, M. A., Reilly, J., Williams, N. M., Buderi, A., Gardner, J. and Winfree, R. 2023. Rare and declining bee species are key to consistent pollination of wildflowers and crops across large spatial scales. - Ecology 104: e3899.

Greenleaf, S. S., Williams, N. M., Winfree, R. and Kremen, C. 2007. Bee foraging ranges and their relationship to body size. - Oecologia 153: 589–596.

Harrison, T., Gibbs, J. and Winfree, R. 2018a. Forest bees are replaced in agricultural and urban landscapes by native species with different phenologies and life-history traits. - Glob Change Biol 24: 287–296.

Harrison, T., Gibbs, J. and Winfree, R. 2018b. Phylogenetic homogenization of bee communities across ecoregions. - Global Ecol Biogeogr 27: 1457–1466.

MacLeod, M., Genung, M. A., Ascher, J. S. and Winfree, R. 2016. Measuring partner choice in plant-pollinator networks: using null models to separate rewiring and fidelity from chance. - Ecology 97: 2925–2931.

McGarigal, K., Wan, H. Y., Zeller, K. A., Timm, B. C. and Cushman, S. A. 2016. Multi-scale habitat selection modeling: a review and outlook. - Landscape Ecol 31: 1161–1175.

Roswell, M., Dushoff, J. and Winfree, R. 2019. Male and female bees show large differences in floral preference (ME Saunders, Ed.). - PLoS ONE 14: e0214909.

Smith, C., Weinman, L., Gibbs, J. and Winfree, R. 2019. Specialist foragers in forest bee communities are small, social or emerge early (J Resasco, Ed.). - J Anim Ecol 88: 1158–1167.

Smith, C., Harrison, T., Gardner, J. and Winfree, R. 2021. Forest-associated bee species persist amid forest loss and regrowth in eastern North America. - Biological Conservation 260: 109202.

Winfree, R. and Kremen, C. 2009. Are ecosystem services stabilized by differences among species? A test using crop pollination. - Proc. R. Soc. B. 276: 229–237.

Winfree, R., Griswold, T. and Kremen, C. 2007. Effect of Human Disturbance on Bee Communities in a Forested Ecosystem. - Conservation Biology 21: 213–223.

Winfree, R., Williams, N. M., Dushoff, J. and Kremen, C. 2014. Species Abundance, Not Diet Breadth, Drives the Persistence of the Most Linked Pollinators as Plant-Pollinator Networks Disassemble. - The American Naturalist 183: 600–611.

ECOGRAPHY Open Access
A JOURNAL OF SPACE AND TIME IN ECOLOGY

Winfree, R., Reilly, J. R., Bartomeus, I., Cariveau, D. P., Williams, N. M. and Gibbs, J. 2018. Species turnover promotes the importance of bee diversity for crop pollination at regional scales. - Science 359: 791–793.

Supplemental Tables

Table S1 Trait information for the 60 species with significant landcover associations. Season was determined by the 5th and 95th percentile of observation dates (see *Bee phenology data and estimating flight windows*, above). Intertegular (IT) distances were measured on specimens from the Winfree lab group collection, either for this study or a previous study (Cariveau et al. 2016).

Species	Season	Median obs day	5th percentile obs day	95th percentile obs day	Mean IT distance (mm)	IT sample size	IT standard deviation (mm)	Sociality
Agapostemon texanus	summer	233	140	275	1.90	6	0.11	Fac. social
Agapostemon virescens	summer	223	158	264	2.06	14	0.15	Fac. social
Andrena bradleyi	spring	113	90	131	2.02	5	0.11	Solitary
Andrena carlini	spring	113	99	141	2.59	9	0.12	Solitary
Andrena cressonii	spring	126	105	166	1.93	6	0.05	Solitary
Andrena erigeniae	spring	117	99	137	1.72	10	0.08	Solitary
Andrena miserabilis	spring	118	91	150	1.56	5	0.09	Solitary
Andrena nasonii	spring	126	107	165	1.62	5	0.08	Solitary
Andrena rugosa	spring	125	86	161	2.02	6	0.13	Solitary
Andrena tridens	spring	117	105	126	1.95	6	0.14	Solitary
Andrena violae	spring	118	101	137	1.68	5	0.27	Solitary
Anthidium oblongatum	summer	207	162	258	2.57	9	0.12	Solitary
Augochlora pura	long season	210	118	240	1.68	22	0.12	Solitary
Augochlorella aurata	long season	195	125	258	1.53	20	0.14	Eusocial
Bombus impatiens	summer	203	170	260	3.48	11	0.45	Eusocial
Calliopsis andreniformis	summer	201	162	236	1.40	5	0.10	Solitary
Ceratina calcarata	long season	171	106	229	1.37	21	0.14	Fac. social
Ceratina dupla	long season	171	118	233	1.22	9	0.21	Fac. social
Ceratina strenua	long season	171	118	227	1.00	7	0.10	Fac. social
Colletes inaequalis	spring	102	79	130	2.73	6	0.19	Solitary
Eucera pruinosa	summer	202	186	223	3.69	10	0.32	Solitary

Species	Season	Median obs day	5th percentile obs day	95th percentile obs day	Mean IT distance (mm)	IT sample size	IT standard deviation (mm)	Sociality
Habropoda laboriosa	spring	116	104	157	3.93	25	0.16	Solitary
Halictus confusus	long season	197	122	240	1.40	5	0.07	Fac. social
Halictus rubicundus	long season	182	109	217	1.92	9	0.21	Fac. social
Lasioglossum coriaceum	long season	166	117	261	1.92	5	0.08	Fac. social
Lasioglossum cressonii	long season	146	108	258	1.30	5	0.14	Eusocial
Lasioglossum ephialtum	long season	184	113	258	1.04	5	0.05	Eusocial
Lasioglossum hitchensi	long season	197	118	258	1.04	5	0.09	Eusocial
Lasioglossum illinoense	long season	206	117	237	0.94	5	0.05	Eusocial
Lasioglossum imitatum	summer	203	137	239	0.82	5	0.04	Eusocial
Lasioglossum leucocomum	summer	207	135	263	1.14	8	0.12	Eusocial
Lasioglossum lineatulum	long season	187	118	242	1.30	5	.12	Eusocial
Lasioglossum oblongum	long season	173	105	254	1.12	5	0.08	Eusocial
Lasioglossum obscurum	long season	158	108	230	1.10	5	0.00	Eusocial
Lasioglossum pilosum	long season	197	111	260	1.26	16	0.09	Eusocial
Lasioglossum platyparium	long season	170	101	266	1.04	5	0.09	Solitary
Lasioglossum quebecense	long season	114	84	217	1.40	5	0.12	Solitary
Lasioglossum subviridatum	long season	118	100	199	1.20	16	0.10	Eusocial
Lasioglossum tegulare	long season	189	114	251	1.00	5	0.07	Eusocial
Lasioglossum trigeminum	summer	197	147	246	1.23	10	0.07	Eusocial
Lasioglossum versatum	summer	196	139	228	1.35	6	0.14	Eusocial
Lasioglossum weemsi	long season	196	112	228	1.06	5	0.11	Eusocial
Megachile rotundata	summer	203	162	261	2.18	8	0.09	Solitary
Melissodes bimaculatus	summer	205	193	224	2.86	5	0.22	Solitary
Nomada composita	spring	118	103	128	1.40	5	0.07	Solitary
Nomada cressonii	spring	141	113	170	1.57	6	0.20	Solitary
Nomada denticulata	spring	118	108	163	2.03	6	0.14	Solitary
Nomada depressa	spring	118	103	130	1.38	6	0.16	Solitary

Species	Season	Median obs day	5^{th} percentile obs day	95^{th} percentile obs day	Mean IT distance (mm)	IT sample size	IT standard deviation (mm)	Sociality
Nomada imbricata	spring	129	107	158	1.86	5	0.05	Solitary
Nomada luteoloides	spring	118	101	138	1.74	5	0.13	Solitary
Nomada maculata	spring	118	114	149	1.82	5	0.14	Solitary
Nomada pygmaea	spring	118	105	161	1.40	7	0.29	Solitary
Osmia atriventris	spring	118	101	165	2.00	5	0.10	Solitary
Osmia bucephala	spring	121	106	171	3.50	10	0.21	Solitary
Osmia cornifrons	spring	116	99	141	2.94	14	0.16	Solitary
Osmia pumila	spring	117	100	163	2.05	6	0.22	Solitary
Osmia taurus	spring	116	92	128	3.02	5	0.15	Solitary
Osmia virga	spring	133	105	182	2.18	5	0.08	Solitary
Sphecodes atlantis	long season	183	128	241	0.78	5	0.04	Solitary

Table S2 The frequency with which spring, summer, and long-season bees had significantly positive and negative associations with forest and open landcover. Comparisons within cells (positive vs. negative for a bee-landcover combination) and within rows (forest vs. open for positive or negative associations with a particular group) were all significant for spring and summer bees (binomial tests, all $p < 0.001$), but not long-season bees ($p > 0.05$).

	Forested landcovers	Open landcovers
Spring bees	39 positive 8 negative	3 positive 38 negative
Summer bees	1 positive 28 negative	15 positive 0 negative
Long-season bees	14 positive 22 negative	15 positive 10 negative

Table S3 Frequency of associations – total, positive, and negative – of bees with each landcover type. We tested whether each landcover had more associations in one or the other direction with a two-tailed binomial test. Those with significantly more positive or negative responses are shown in bold text and with an asterisk (*).

Landcover	Significant associations	Positive	Negative	p
Developed – low	25	10	15	0.42
Developed – high*	26	6	20	**0.0094**
Crops	21	9	12	0.66
Pasture*	9	8	1	**0.039**
Forest – deciduous*	33	23	10	**0.035**
Forest – evergreen	16	6	10	0.45
Forest – mixed	29	15	14	1.0
Foreseted wetland*	21	4	17	**0.0072**
Forest edge	13	6	7	1.0

Supplemental Figures

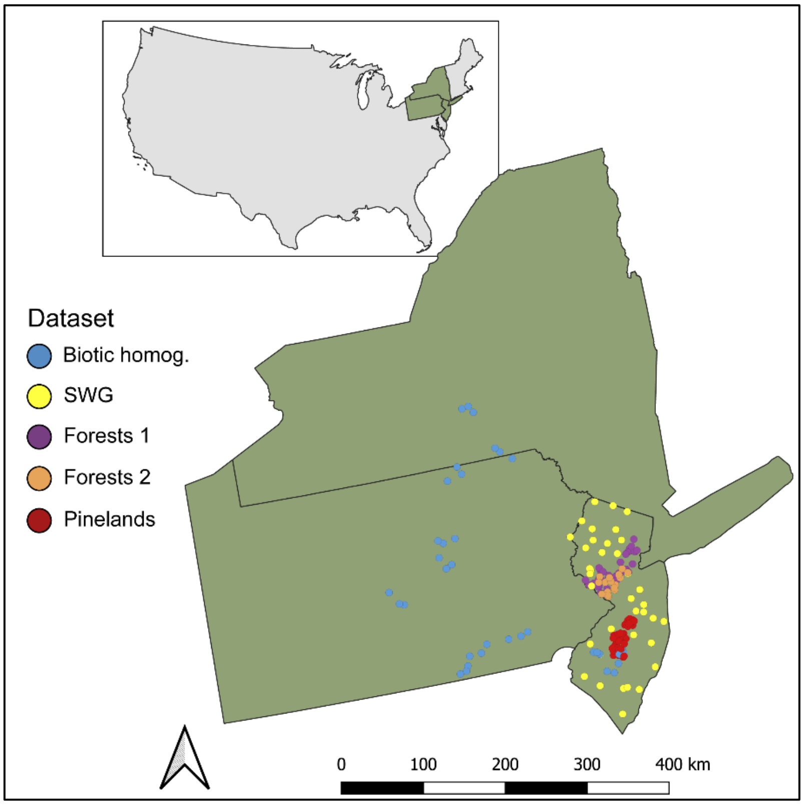
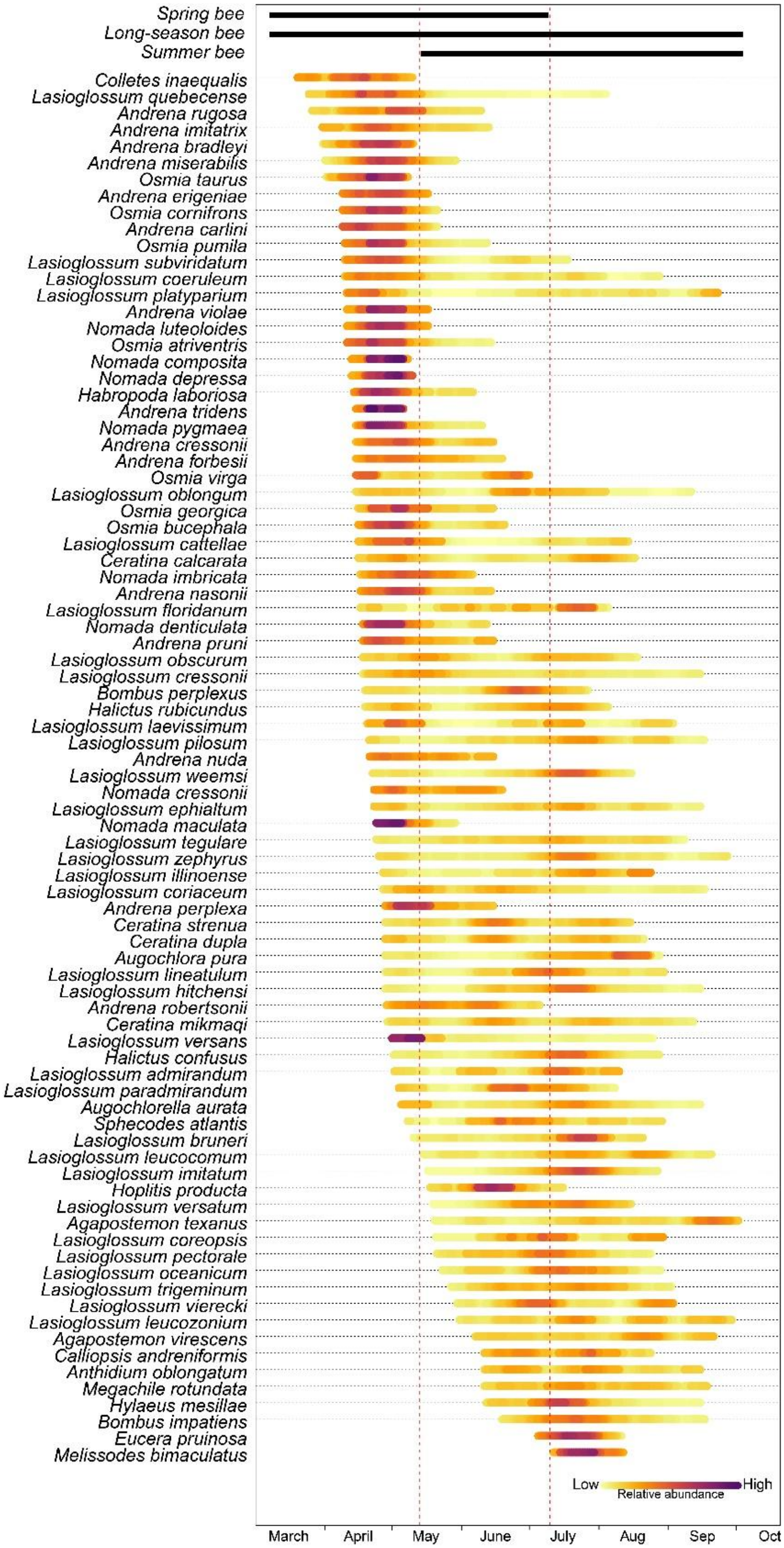


Figure S1 Location of 165 study sites from the 5 datasets used in this analysis. Sites were located primarily in New Jersey, but also Pennsylvania and New York, all three being states in the northeastern United States (inset).

Figure S2 (next page) Flight windows of the 84 species analyzed in this study. The start and end of each species' flight window was defined as the 5th and 95th percentile of observation dates across all specimens of that species in the phenology dataset (see Main Text and Supplemental Methods, above). The relative abundance for each day was calculated as the proportion of all specimens of that species that were collected during in a 15-day window around that date. The vertical dashed lines and bars at the top show our operational definition of spring, long-season, and summer bees: spring bees start their flight window before May 13 and end before July 10; long-season bees start before May 13 but end after July 10; and summer bees start after May 10 (Figure S1).



Spring bee
Long-season bee
Summer bee
Colletes inaequalis
Lasioglossum quebecense
Andrena rugosa
Andrena imitatrix
Andrena bradleyi
Andrena miserabilis
Osmia taurus
Andrena erigeniae
Osmia cornifrons
Andrena carlini
Osmia pumila
Lasioglossum subviridatum
Lasioglossum coeruleum
Lasioglossum platyparium
Andrena violae
Nomada luteoloides
Osmia atriventris
Nomada composita
Nomada depressa
Habropoda laboriosa
Andrena tridens
Nomada pygmaea
Andrena cressonii
Andrena forbesii
Osmia virga
Lasioglossum oblongum
Osmia georgica
Osmia bucephala
Lasioglossum cattellae
Ceratina calcarata
Nomada imbricata
Andrena nasonii
Lasioglossum floridanum
Nomada denticulata
Andrena pruni
Lasioglossum obscurum
Lasioglossum cressonii
Bombus perplexus
Halictus rubicundus
Lasioglossum laevissimum
Lasioglossum pilosum
Andrena nuda
Lasioglossum weemsi
Nomada cressonii
Lasioglossum ephialtum
Nomada maculata
Lasioglossum tegulare
Lasioglossum zephyrus
Lasioglossum illinoense
Lasioglossum coriaceum
Andrena perplexa
Ceratina strenua
Ceratina dupla
Augochlora pura
Lasioglossum lineatulum
Lasioglossum hitchensi
Andrena robertsonii
Ceratina mikmaqi
Lasioglossum versans
Halictus confusus
Lasioglossum admirandum
Lasioglossum paradmirandum
Augochlorella aurata
Sphecodes atlantis
Lasioglossum bruneri
Lasioglossum leucocomum
Lasioglossum imitatum
Hoplitis producta
Lasioglossum versatum
Agapostemon texanus
Lasioglossum coreopsis
Lasioglossum pectorale
Lasioglossum oceanicum
Lasioglossum trigeminum
Lasioglossum vierecki
Lasioglossum leucozonium
Agapostemon virescens
Calliopsis andreniformis
Anthidium oblongatum
Megachile rotundata
Hylaeus mesillae
Bombus impatiens
Eucera pruinosa
Melissodes bimaculatus
Low
High
Relative abundance
March
April
May
June
July
Aug
Sep
Oct

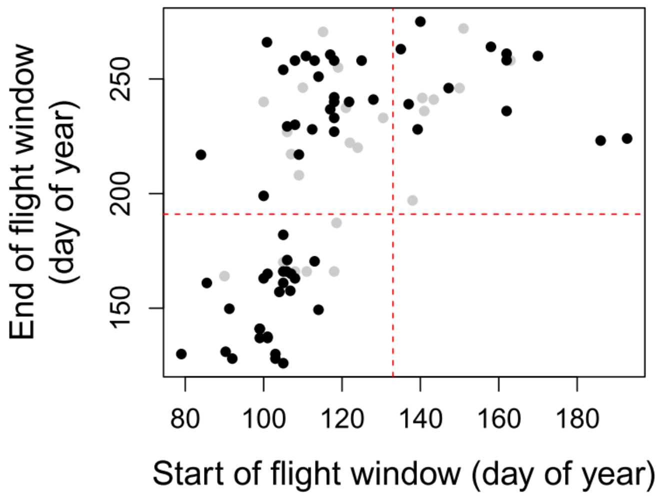


Figure S3 Phenological window start and end dates for each bee species in our analysis. Black dots are species that had significant landcover associations; grey dots are those that did not. We categorized bees as spring, summer, or long-season species based on the start and end of their flight window. Spring bees, in the lower left, emerge early and end early. Summer bees, in the top right, start late and end late. Long season bees, in the top left, start early and end late. We chose the cutoff dates to define these groups based on based on visual inspection of this plot and what seemed to be natural gaps in the start or end dates. We then confirmed that bees with known phenologies landed in the expected groups (e.g., *Osmia* and *Andrena* being spring bees).

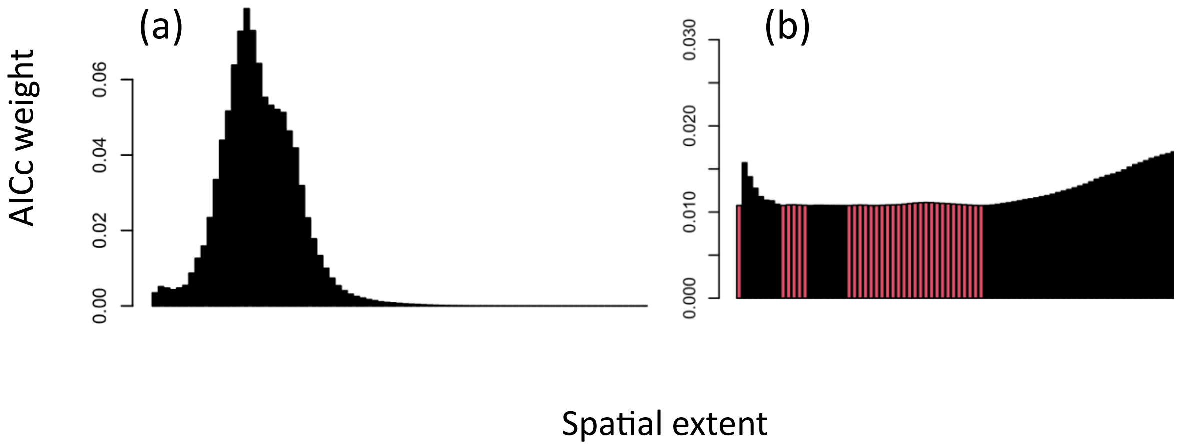


Figure S4 An illustration of our scale selection process, using the response by the bee *Andrena tridens* to (a) deciduous forest and (b) low intensity development. Each bar along the horizontal axis represents a spatial scale (extent) at which landcover proportion was measured (black = positive effect, red = negative effect), and the vertical axis shows the AICc weight of the model in which landcover was measured at that scale. *Andrena tridens* shows a positive response to deciduous forest, with a scale of effect that is strongest at 1600 m (ΔAICc of the null = 18.1) and does not show any response to low intensity development (ΔAICc = 0).

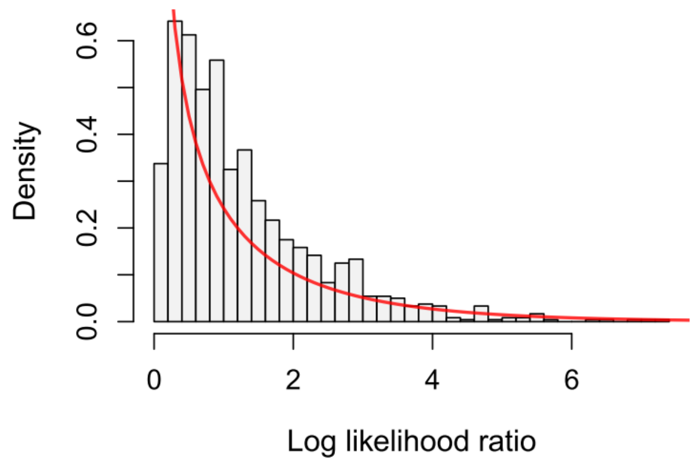


Figure S5 Distribution of the log likelihood ratios (grey histogram) generated by our randomization-based simulations (see Supplemental Methods, above), overlaid with the χ^2 distribution (*df* = 1) that would be expected when testing a series of independent models (red line). Here, because we tested models for significance after performing scale selection, even null models tended to outperform the typical parametric expectation. As a result, calculating p-values from the typical χ^2 distribution would have been anti-conservative and led to inflated type I error and false discovery rates.

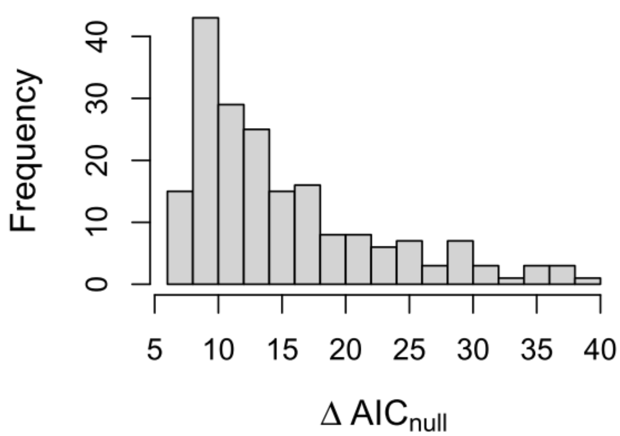


Figure S6 Frequency distribution of model performance as measured by the ΔAICc of the null model. This shows the performance of only the 203 significant models. Range is ΔAICc = 7.45 to 38.6.

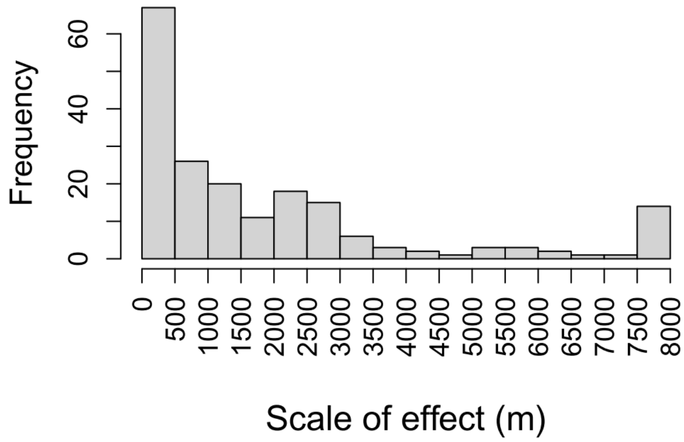


Figure S7 Distribution of the 203 scales of response, across all species and landcovers. Mode = 200 m; median = 1000 m; mean = 1994 m.

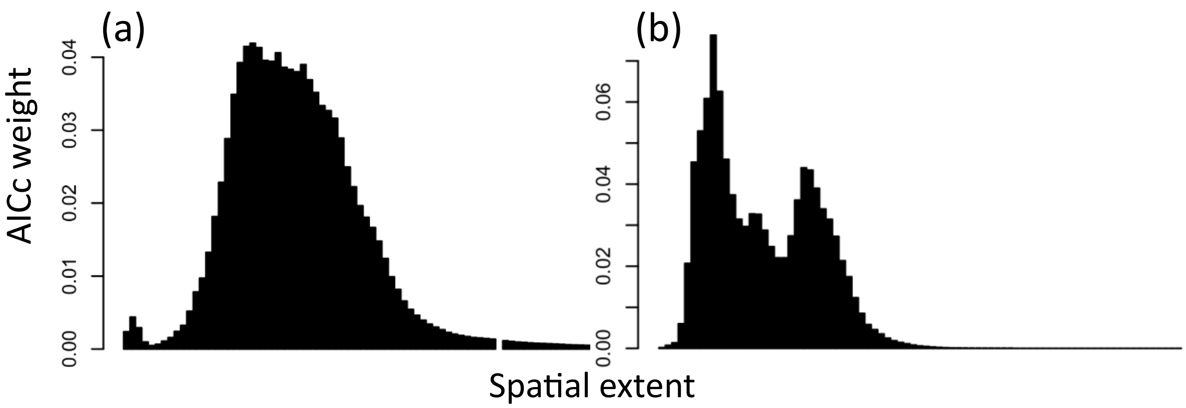


Figure S8 Results from scale selection were often imprecise, with a broad distribution of AICc weight across putative scales of response (e.g., panel a), resulting in an imprecise estimates of a species' scale of effect, and was sometimes bimodal (e.g., panel b), suggesting effects occurring at two distinct scales. In the latter case, the taller peak was used as that species' scale of effect.

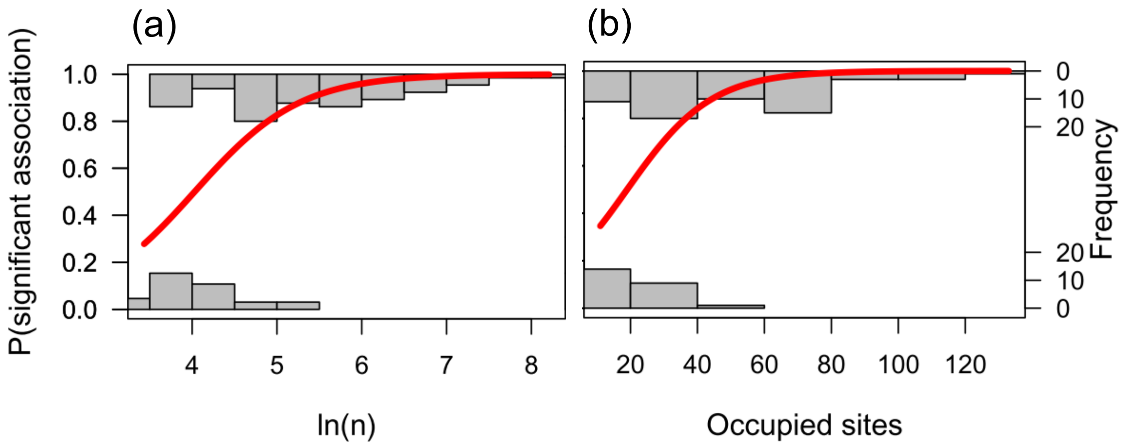


Figure S9 The effects of (a) bee abundance and (b) site occupancy on the probability of a species having one or more significant landcover associations. Effects were determined via logistic regression ($p < 0.001$), and red lines represent predicted probabilities for a given value of abundance or occupancy. The histograms on the top and bottom axes of each panel show the frequency of species at different abundances or occupancies. The takeaway is that species with high abundance or occupancy always had at least one significant landcover association, whereas species with low abundance or occupancy had a split chance of having a significant landcover association or not. We take this as evidence that non-significant results could be due to statistical power rather than evidence for those species' habitat generality.