



Greater bee diversity is needed to maintain crop pollination over time

Natalie J. Lemanski¹✉, Neal M. Williams² and Rachael Winfree¹

The current biodiversity crisis underscores the need to understand how biodiversity loss affects ecosystem function in real-world ecosystems. At any one place and time, a few highly abundant species often provide the majority of function, suggesting that function could be maintained with relatively little biodiversity. However, biodiversity may be critical to ecosystem function at longer timescales if different species are needed to provide function at different times. Here we show that the number of wild bee species needed to maintain a threshold level of crop pollination increased steeply with the timescale examined: two to three times as many bee species were needed over a growing season compared to on a single day and twice as many species were needed over six years compared to during a single year. Our results demonstrate the importance of pollinator biodiversity to maintaining pollination services across time and thus to stable agricultural output.

The Earth is experiencing rapid biodiversity loss as a result of human activity¹. This crisis has created an urgent need to understand the consequences of biodiversity loss for the ecosystem functions and services on which humans depend^{1,2}. One vital ecosystem service that is likely to be threatened by the biodiversity crisis is crop pollination by wild bees, especially given current evidence of widespread declines in bee communities^{3–5}. A wealth of experimental^{1,6–11} and theoretical^{12–15} studies showed strong evidence that biodiversity loss can adversely affect ecosystem functions like pollination. However, it is not fully understood how these experimental and theoretical results translate to degradation of ecosystem function in real-world systems^{16,17}, such as a loss of crop pollination¹⁸, carbon storage⁶, pest control¹⁹ or other important services currently provided by wild species, where there are direct consequences for human well-being.

One way real-world communities differ from experiments is that they have highly uneven species abundance distributions, such that most of the individuals in the community belong to only a few species²⁰. As a result, at any one place and time, the majority of function is commonly provided by a small number of highly abundant species^{21–27}. Previous studies of wild bee communities found that a few species provide the majority of pollination at any one place and time^{22,23,28,29}. Because many species in the community are too rare to contribute notably to function, species richness (the most widely used measure of biodiversity^{6,30,31}) is a weaker predictor of function in communities with uneven species abundances than in communities where abundances are more even^{25,28}.

Another way biodiversity experiments typically differ from real-world systems is their smaller spatial and temporal scale. Many species may be needed to maintain ecosystem function at larger scales^{17,32} due to different species providing function in different places^{33–35}. Similarly, if different species provide function at different times, more species will be required to maintain a consistent level of ecosystem function over a longer timescale. Therefore, biodiversity may be important to the stability of function across time by providing insurance against fluctuations in individual species^{36–39}. Within a year, phenological differences among species mean that the identity of the most important providers of ecosystem function will likely change over time in a predictable pattern as different

species peak in abundance at different parts of the season^{40–42}. For example, for plants that flower over many weeks within a growing season, phenological turnover in the bee community may mean that different bee species pollinate the same plant at different times of the year¹². Across years, the relative abundance of different species may change over time due to differences among species in their response to environmental fluctuations^{17,43,44}. Indeed, large stochastic fluctuations in population size are the norm for insects and many other important providers of ecosystem function^{40,42,45–47}. In addition, flowering phenology may fluctuate across years, resulting in the bloom period overlapping different pollinator assemblages that are active at different times. As a result, plants may be pollinated by different bee species in different years. Lastly, fluctuations in total bee abundance over time, regardless of species-specific abundances, may increase the number of species needed to provide consistent function over longer timescales. In low-abundance years, a site may need all of its species, including rare ones, to achieve a function threshold⁴⁸. These phenomena could mean that the number of bee species needed to maintain pollination over time is much greater than has been predicted by ‘snapshot’ studies conducted thus far.

There is experimental, but thus far limited real-world, evidence to support the idea that the importance of diversity can depend on temporal scale. Studies of experimental communities found that higher diversity increases the temporal stability of ecosystem function^{15,39} due to different species contributing to function at different times^{49,50} or under different environmental conditions^{51–53}. Similarly, in biodiversity experiments, positive effects of diversity on function generally increase with longer spatial and temporal scales^{8,33,50,54}. However, it is not well understood how the need for biodiversity changes with temporal scale in real-world communities. This issue is particularly important for pollinator-dependent crops because reliable yields are critical to farmers’ livelihoods and human nutrition. Understanding the role of pollinator diversity across a crop flowering season and from one year to the next is fundamental to safeguarding pollinator diversity and its potential role in ensuring sustainable crop yields. The alternative is increasing reliance on managed pollinator populations, which is increasingly fraught (for example, Perry et al.⁵⁵) and, in many contexts, does not fully compensate for the contributions of wild pollinator species¹⁸. Thus, it is

¹Rutgers University, Department of Ecology, Evolution & Natural Resources, New Brunswick, NJ, USA. ²University of California Davis, Department of Entomology & Nematology, Davis, CA, USA. ✉e-mail: Natalie.Lemanski@rutgers.edu

Table 1 | GLMM results for the change in the number of species needed with the number of sampling days

Crop	Effect size (slope)	Factor increase	Intercept	Marginal R^2	Conditional R^2
Eastern watermelon	0.283 ± 0.053	1.33	0.825 ± 0.127	0.097	0.307
Western watermelon	0.141 ± 0.025	1.15	0.479 ± 0.088	0.124	0.157
Blueberry	0.283 ± 0.074	1.33	0.253 ± 0.178	0.105	0.261

Effect sizes are reported as the mean slope of the GLMM $\pm$ s.e. Because the GLMM uses a log link, exponentiating the effect size gives the factor increase in the number of bee species needed for each additional date across the growing season. Marginal R^2 values indicate how much of the variance in the number of bee species needed was explained by the main effect (the duration of the growing season considered), while the conditional R^2 also includes variance explained by differences among sites.

important to examine how temporal scale affects the relationship between pollinator diversity and pollination services in real-world systems.

In this study, we asked how the number of species needed for ecosystem function changes with the temporal scale examined, that is, with the length of time over which the function is required to be maintained. We addressed this question using three datasets that measured crop pollination by wild bees, both multiple times over the course of the flowering season and for multiple years. The study systems included watermelon farms in the western United States (hereafter, western watermelon), watermelon farms in the eastern United States (eastern watermelon) and blueberry farms in the eastern United States (eastern blueberry). Specifically, we asked how the number of wild bee species needed to maintain a threshold level of pollination changes as we increased the timescale of the analysis across (1) an increasing number of days within the growing season and (2) an increasing number of years. To control statistically for stochastic effects related to sampling bee communities over time, we developed a null model that resampled from a fixed species pool. We then compared our observed results to the null model to determine how much of the increase in the number of species needed over time was due to biological effects (species turnover) versus sampling effects.

Results

Bee diversity needed to maintain pollination across the growing season. In all three study systems (western watermelon, eastern watermelon and eastern blueberry), increasing the length of the season over which the pollination threshold must be met increased the number of bee species needed (Table 1). For the western watermelon farms, a mean of 1.64 species were needed to meet the threshold at a single date but the number of species needed increased by 15% for each added sampling date, with 3 times as many species required over the entire 12-week season (a 209% increase; Fig. 1a). For the eastern watermelon farms, the mean number of species needed was 3 for a single date, but increased by 33% for each additional sampling date, increasing 76% over the entire 7-week season (Fig. 1b). For the blueberry farms, the mean number of species needed was 1.73 for a single date but increased 33% for each additional date sampled and 76% over the entire 3-week season (Fig. 1c). The null model showed that the majority of this increase was due to actual turnover rather than sampling effects (Extended Data Fig. 1). Specifically, species turnover accounted for 95% of the increase for western watermelon, 65% for eastern watermelon and 77% for eastern blueberry (Supplementary Table 7). When only the first 3 time points are considered for the western watermelon system, to make its time series the same as the other two crops, the percentage of the increase attributable to turnover was similar (79%), suggesting that the strength of biological effects observed increases with the length of the time series.

A sensitivity analysis showed that the number of species needed increased with season length over a range of function thresholds. However, the choice of threshold influenced the magnitude of the increase, with the number of species needed increasing more steeply

when higher function thresholds were used (Extended Data Fig. 5). For the western watermelon farms, the number of species needed more than doubled over the entire season even for the lowest threshold of 10% and increased over fourfold for the highest threshold of 90% (Supplementary Table 9). For the eastern watermelon farms, the percentage increase in species needed over the entire season ranged from 57% at the lowest (10%) threshold to 88% at the highest (90%) threshold (Supplementary Table 9). Similarly, for the eastern blueberry farms, the percentage increase in species needed over the season ranged from 54% at the 10% threshold to 112% at the 90% threshold (Supplementary Table 9).

Bee diversity needed to maintain pollination across multiple years. The number of bee species needed to meet the pollination threshold in all years also increased with the number of years included (Table 2). For the eastern watermelon farms, the number of bee species needed increased 26% for each additional year, with 3 times as many species required for a 6-year period compared to a single year (a 219% increase; Fig. 2a). Similarly, for the western watermelon farms, the number of bee species needed increased by 27% per year, with 62% more species required for a 3-year period (Fig. 2b). For the blueberry system, the number of species needed increased by 21% for each additional year, with 47% more species needed for a 3-year period (Fig. 2c). The null model indicated that the majority of the increase in species needed was attributable to bee species turnover (a biological effect) rather than sampling effects (Extended Data Fig. 2). Specifically, species turnover accounted for 57%, 63% and 75% of the increase for western watermelon, eastern watermelon and blueberry, respectively (Supplementary Table 7).

A sensitivity analysis showed that the increase in species needed with the number of years was relatively robust to the choice of threshold (Extended Data Fig. 6). For the eastern watermelon and western watermelon farms, the slope of the increase changed very little with threshold (Extended Data Fig. 6). For the eastern watermelon farms, the percentage increase in species needed across years was actually higher with a lower choice of threshold because lowering the function threshold also decreased the number of species needed for a single year and thus made the relative increase to subsequent years steeper (Supplementary Table 10). For the western watermelon farms, the choice of function threshold had essentially no effect, with the number of species needed increasing 59–66% for all thresholds tested (Supplementary Table 10). In contrast, for the eastern blueberry farms, the function threshold influenced the results: for the highest threshold (90%), the number of species needed doubled for the full 3-year period compared to a single year, while for the lowest threshold (10%), there was little change in the number of species needed with the number of years (Extended Data Fig. 6). However, given evidence that blueberry fruit set is often limited by insufficient pollination^{56,57}, requiring only 10% of observed pollination as a threshold is probably unreasonably conservative; indeed even the 90% threshold (for which the number of species needed doubles over time) is likely insufficient to provide full pollination.

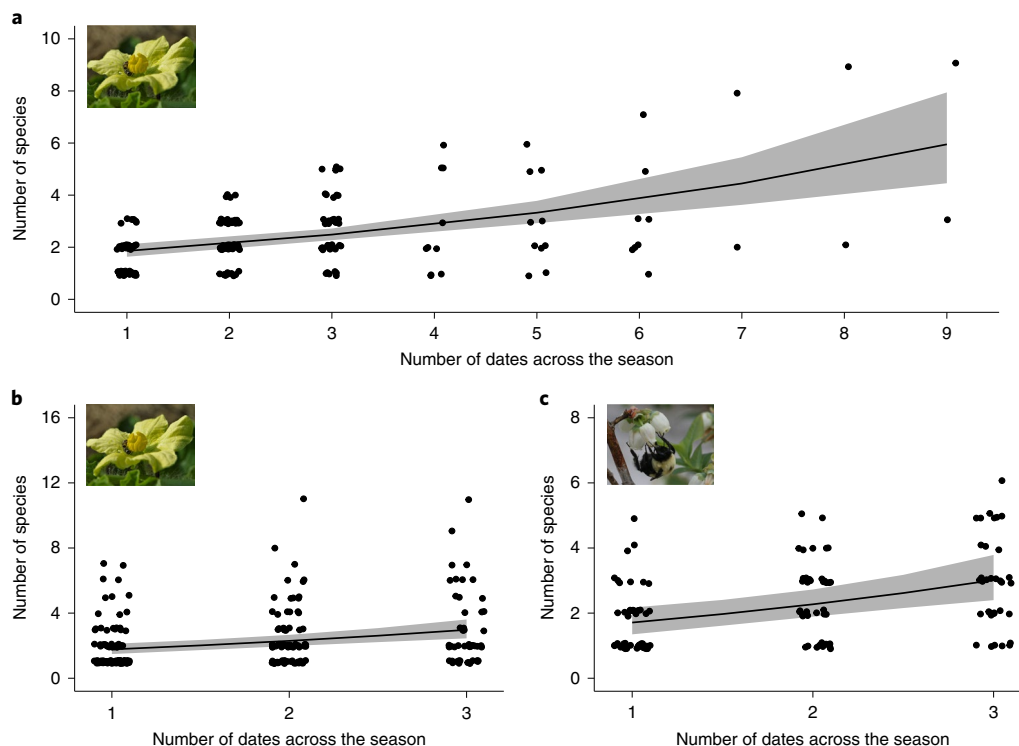


Fig. 1 | Number of bee species needed for pollination across multiple dates within a flowering season. a–c, Number of bee species needed to meet the pollination threshold (50% of mean observed pollination) at every date sampled over the course of a single year for western watermelon farms (a), eastern watermelon farms (b) and blueberry farms (c). Each point represents one site-year. The solid line represents the slope and intercept estimated from the GLMM with random intercepts for site. The shaded area represents the 95% CI around the GLMM estimates.

Table 2 | GLMM results for the change in the number of species needed with the number of years

Crop	Effect size (slope)	Factor increase	Intercept	Marginal R^2	Conditional R^2
Eastern watermelon	0.232 ± 0.033	1.26	0.943 ± 0.191	0.258	0.761
Western watermelon	0.240 ± 0.112	1.27	0.910 ± 0.193	0.093	0.093
Blueberry	0.191 ± 0.103	1.21	1.12 ± 0.219	0.085	0.266

Effect sizes are reported as the mean slope of the GLMM $\pm$ s.e. Because the GLMM uses a log link, exponentiating the effect size gives the factor increase in the number of bee species needed for each additional year considered. Marginal R^2 values indicate how much of the variance in the number of bee species needed was explained by the main effect (the number of years), while the conditional R^2 also includes variance due to differences among sites.

Discussion

A wealth of experimental evidence suggests that biodiversity loss adversely affects ecosystem function^{6,7,58} but uncertainty remains over how these results translate to the larger scales at which biodiversity losses occur in real-world systems. The insurance hypothesis suggests that diversity will be more important for function when measured over a longer time scale because having more species provides functional redundancy, which buffers ecosystem function against fluctuations in species abundances^{36,54,59}. Our results provide some of the strongest support to date for insurance effects in real-world communities. We found that 2–3 times as many bee species were needed to meet a target level of crop pollination over the course of a growing season compared to a single date (Fig. 1). Similarly, twice as many species were needed to provide pollination over the course of 6 years compared to a single year (Fig. 2). These effects were most apparent for our longer time series and did not show signs of levelling off, suggesting that looking at a longer time series may reveal even stronger effects of timescale.

The main mechanism by which the need for diversity has been hypothesized to increase with timescale is through the turnover in the identity of functionally important species^{17,37,49}. Because natural

communities are often characterized by highly uneven species abundance distributions, relatively few species are abundant enough to be important to function at any one place or time^{7,21,28,60}. However, the insurance hypothesis predicts that the number of species needed will increase with timescale because of changes over time in species relative abundances^{36,37,61}. We found that the number of bee species needed increased 76–209% over the course of the growing season due to changes in abundances (Fig. 1). Phenological data for the bee species in our studies supports the conclusion that these changes in abundance are due to phenological difference among species (Extended Data Figs. 3 and 4). In addition, our null model confirms that species turnover, not merely sampling effects, generates the need for greater diversity, especially for our longest time series (Extended Data Fig. 1). This result suggests that pollinator diversity may be especially important for crops with a long flowering season and for farms growing a variety of crops that bloom at different times of the year. Of course, pollinator diversity will be even more important in natural (non-agricultural) systems where there are even more plant species that co-occur in one place⁶². Our results demonstrate that a ‘snapshot’ view can underestimate the need for biodiversity when communities change in either composition or total abundance over time.

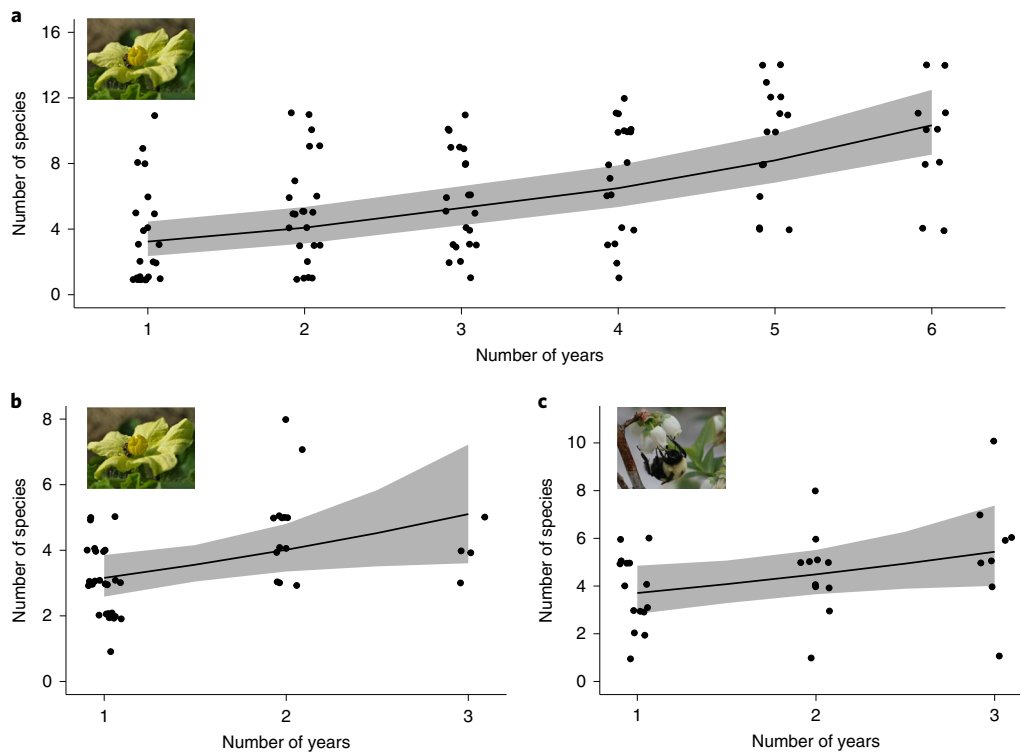


Fig. 2 | Number of bee species needed for pollination across multiple years. a–c, Number of bee species needed to meet the pollination threshold (50% of mean observed pollination) in all years at eastern watermelon farms (**a**), western watermelon farms (**b**) and blueberry farms (**c**). Each point represents one site. The solid line represents the slope and intercept estimated from the GLMM with random slopes and intercepts for site. The shaded area represents the 95% CI around the GLMM estimates.

Across multiple years, we predicted that changes in species relative abundances would occur due to asynchronous fluctuations in the populations of different species across years. These asynchronous changes should allow species experiencing high abundance in a given year to functionally compensate for species experiencing low abundance in that year^{17,37,38,63}, resulting in a greater number of species needed to maintain a given level of function over time. Our results provide support for this prediction of the insurance hypothesis: in the 6-year time series, we found a 47–219% increase in the number of species needed as we increased the number of years (Fig. 2). The null model again confirms that this increase was largely due to interannual turnover of bee species, not merely to sampling effects (Extended Data Fig. 2). Together, these findings suggest that for crop pollination, pollinator diversity provides insurance against fluctuations in species abundance over time as predicted by the insurance hypothesis^{36,38,61}.

A third mechanism through which the need for biodiversity may increase with the scale examined, and one that has largely been overlooked in the biodiversity–ecosystem functioning literature (but see refs. ^{10,21,22,64}), is variation in the total abundance of individuals across space or time. In our results, variation in total bee abundance across time was a major source of accumulation of bee species needed. Dates with high bee abundance typically only needed a few species to meet the target pollination threshold (Fig. 3). In contrast, many dates with lower abundance needed all of their species to reach the threshold (Supplementary Table 8). This result suggests that species diversity may result in a different type of insurance effect, in which low abundance years or weeks lead to an ‘all hands on deck’ situation where even rare species are needed for a target level of ecosystem function to be achieved^{34,48}. The conservation literature has tended to emphasize the role of diversity while de-emphasizing the role of total abundance in ecosystem function²².

However, both diversity and abundance can decrease as a result of anthropogenic impacts, for bees^{3,5} and most other taxa that have been studied^{65–68}. Our results suggest that losses of both abundance and diversity can have important effects on ecosystem function.

One limitation of our analytical approach is that the threshold is based on the amount of pollination currently received, rather than the amount that is actually needed to maximize crop yield. While the number of pollen grains needed to maximize fruit set on a single flower has been estimated empirically (that is, the biological pollination threshold^{56,69,70}), it would not be meaningful to compare our observed function data to this per-flower threshold. This is because our data are in units of pollen deposition per transect per unit time, rather than per flower over its entire lifetime, which are the units of the biological pollination threshold. Net collecting all bees in a transect is necessary to get species identifications on the entire bee community since many species cannot be identified in the field. These species-level identifications were essential for our study’s goal of measuring the pollination function contributed by each bee species. Because we could not observe flowers over their entire lifetime, and do not know for how long we looked at each individual flower in the transect, we cannot convert our transect-level function measurements to the units of pollen deposition per flower.

How conservative our threshold based on 50% of observed function is regarding the number of bee species needed for consistent pollination will depend to some degree on whether the amount of pollination a crop currently receives is close to the amount needed to maximize yield, which can vary across cultivars and environmental contexts^{57,71}. Studies have found that blueberry is often pollen-limited, suggesting that our 50% threshold is likely conservative for this crop and that in fact more bee species are needed than we report in this study⁵⁶. While watermelon has often been found to currently receive adequate pollination in our eastern study

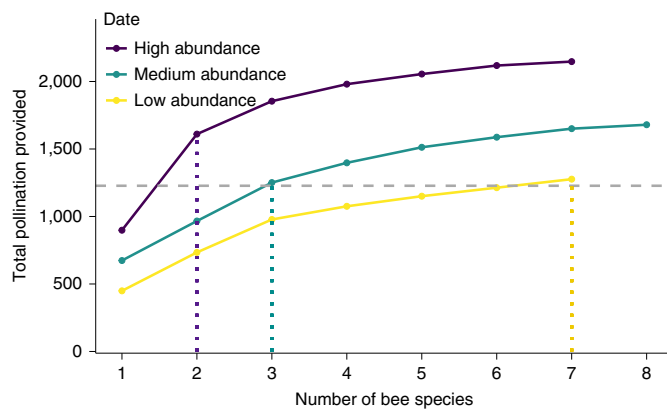


Fig. 3 | Accumulation curves of pollination provided at a single site on three different days, across which aggregate bee abundance varied from low (yellow) to medium (blue) and high (purple). The solid lines show the cumulative pollination received versus cumulative bee species richness. Bee species are added on the x axis in descending order of function. Each vertical dotted line shows the minimum number of bee species needed to meet the function threshold (dashed grey line) for that date. For the date with the highest overall bee abundance at the site (27 bees, solid purple line), only 2 of the 7 species were needed to meet the threshold, while for the date with the lowest abundance at the site (18 bees, solid yellow line), all species present were needed.

system^{70,72}, our sensitivity analysis found that our results for watermelon were relatively robust to our choice of threshold (Extended Data Figs. 5 and 6). Indeed, given that our objective was to determine the relative number of bee species needed over different timescales, rather than the absolute number of species needed at a particular place or time, we believe that our more abstracted measure provides a consistent framework that can be broadly applied and can reveal important information about how the need for species diversity changes with timescale.

There are several other limitations of our study, which could be useful topics for future work. First, given that our research focused on ecosystem services, our data and analysis are based on function provided by wild bees; indeed, we did not have permission to net-collect honeybees, which are a domesticated (and privately owned) species in our study system. But we acknowledge that there were domestic honeybees present at most of our study farms and that they are a dominant crop pollinator. Nonetheless, wild bees often make important contributions to crop yield even in the presence of high honeybee abundance^{18,73–76} and there is evidence that the demand for crop pollination is growing faster than the supply of managed honeybees⁷⁷, suggesting a vital need to understand the magnitude of ecosystem services that would be lost if wild bee species were lost from the system.

Second, we acknowledge that additional factors beyond bee phenology can produce changes in species abundances on crop flowers over time. For example, bees may change their visitation to flowers over time due to competitive interactions or switching to more rewarding flower species. Such rewiring of interactions can occur over time even without turnover of species present⁶². If other bee species would increase their visitation to crop flowers following the loss of a competitor, then the results reported in this study might overestimate the number of bee species needed. However, previous work suggested that such compensatory effects are generally more important in influencing abundance in experiments than in natural systems^{78,79}. In addition, the interpretation that phenology plays a role in the species turnover we observed is supported by the fact that interspecific differences in phenology occur for specimens sampled across a wide variety of habitats, rather than just on a single flower

species (Extended Data Figs. 3 and 4). Lastly, regardless of the exact cause of the temporal turnover we observed in the bee species visiting crop flowers, the consequence for crop pollination is the same: more bee species are needed to achieve the pollination threshold over a longer timescale.

Although we only had 1 system with data for more than 3 years (eastern watermelon) and 1 with data spanning a wide range of dates within a year (western watermelon), the magnitude of increase in species needed with timescale was remarkably consistent among crop systems when considered over the same interval of time. In addition, the fact that the relationship between timescale and the number of species needed did not level off suggests that even longer time series, spanning multiple seasons, may further bolster the need for biodiversity to ensure reliable ecosystem service.

Conclusions

Given the current biodiversity crisis, it is vital to understand how species diversity affects ecosystem function at the spatio-temporal scales at which biodiversity loss occurs in real-world systems. Our results suggest that for the important real-world ecosystem service of crop pollination, a greater number of bee species are needed at longer timescales due to turnover in species composition and variation in abundance of the most functionally important bee species. At any given time, most pollination is performed by a few highly abundant species. However, the identity of these abundant species changes within and across years; even numerically dominant species are not always present in sufficient numbers to meet function thresholds without additional contributions by less abundant species. These results support the idea that biodiversity provides insurance against fluctuations in ecosystem function across time.

Methods

Study sites. We collected three datasets in which we measured crop pollination by wild bees: one from highbush blueberry (*Vaccinium corymbosum* L., cultivars Duke and Bluecrop) in the eastern United States, one from watermelon (*Citrullus lanatus*, various cultivars) in the eastern United States and one from watermelon (cultivars Crimson Sweet and Yellow Baby) in the western United States. The blueberry study sites consisted of 16 farms within a 35 × 55 km² area in southern New Jersey. Sites were sampled 3 times, over a 17–23 d period each year, for 3 years (Supplementary Tables 1 and 2). The eastern watermelon study sites consisted of 25 farms within a 90 × 60 km² area of central New Jersey and east central Pennsylvania. Farms were sampled 1–3 times, over a 24–52 d period each year, for a total of 6 years (Supplementary Tables 1 and 2). The western watermelon study sites consisted of 36 farms within a 64 × 49 km² area of the Northern Central Valley of California. Farms were sampled between 3 and 9 times, over an 8–12-week period each year, for 3 years (Supplementary Tables 1 and 2).

One of the common challenges to measuring seasonal changes in pollination is that as the plant community changes, bees must be collected off different plants; therefore, it can be difficult to determine whether phenological turnover in the pollinator community is due to plant identity or bee phenology. A useful feature of the western watermelon farms is that the planting times within a field are staggered so that the same fields experience multiple peak blooms at several times across the growing season. This gives us the rare opportunity to measure changes in bee diversity and abundance over a long seasonal timescale on a single plant species and in the same locations. To ensure that each farm represented an independent bee community, all sites were separated by a minimum of 1.0 km, which is outside the flight range of most bee species in these communities⁸⁰. Further details on these field studies can be found in refs. ^{22,81}.

Flower visitation measurements. We used parallel study designs and methods to measure crop pollination by wild bees in our three systems. We chose sampling dates to correspond with the peak bloom of each crop (Supplementary Table 1). On each sampling day, we netted all bees observed visiting crop flowers within a predefined 50–200 m² transect. Netting was repeated in four 10 min periods each sampling day for the western watermelon sites and in three 20 min periods each sampling day for the eastern watermelon and blueberry sites. All collected bee specimens were curated in the laboratory and identified to species where possible with the remainder lumped into groups of indistinguishable species. Overall, 97.8% of the specimens were identified to the species level, with most of the remaining 2.2% consisting of unidentified morphospecies of the diverse subgenus *Lasioglossum* (subgenus *Dialictus*). Specimens were identified using a combination of published taxonomic keys and assistance by expert taxonomists, including R. Thorp, J. Gibbs, J. Ascher and Z. Portman.

Pollination function measurements. To measure the per-visit pollination provided by each type of bee, we experimentally measured the number of conspecific pollen grains deposited on stigmas by individual bees in single visits to flowers. Before data collection, we placed pollinator enclosures around unopened flowers. Once flowers were open, we removed the enclosures, cut the flowers and offered them individually to foraging bees. One individual bee was allowed to visit each flower, after which the flower was protected from further visitation. Bees visiting experimental flowers were visually identified to morphogroups (Supplementary Tables 3–5) since it is not possible to identify many wild bees to species without catching them. In the western watermelon study, bees were also identified to sex, when possible, for pollen deposition measurements.

After being visited, experimental flowers were allowed to sit in water at room temperature for 24 h (watermelon) or 48 h (blueberry) to permit pollen tube growth. Stigmas were then softened in 10% KOH and stained with 1% fuchsin (watermelon) or softened in 1 M NaOH and stained with 0.01% aniline blue (blueberry). We then counted the number of conspecific pollen grains on each stigma using a compound microscope. We assumed that most ungerminated grains fall off the stigma during processing and that this count therefore represents germinated pollen grains⁷⁰. Blueberry packages pollen in tetrads (groups of four pollen grains) but it is rare that all four grains develop pollen tubes. Therefore, we counted only tetrads with at least one germinating pollen tube as in Cariveau et al.⁸¹. Single-visit pollen deposition data were collected at all farms for blueberry and a subset of farms that had sufficient bee abundance and diversity to get adequate sample sizes for the two watermelon systems. Pollen deposition data were pooled across sites and a mean pollen deposition value was calculated for each morphospecies; therefore, we assumed that for bees foraging in monoculture crop fields, the same bee species had similar per-visit pollen deposition at all farms.

Minimum set analysis. To examine how the number of bee species required for pollination changes with timescale, we first calculated the total pollination that each bee species provides at each site-date in a given crop system. Site-date is a logical unit of analysis for measuring pollination because watermelon flowers are only open for 1 d. Although blueberry flowers are open for 5–8 d, we used the same site-date unit to allow comparison among study systems. In addition, sampling effort was standardized across site-dates, allowing us to directly compare total species abundance. We calculated the pollination provided by each bee species at a site-date as the number of visits by that species at that site-date $\times$ the mean per-visit pollen deposition for the morphogroup to which that species belonged^{34,48,81}. We summed the pollination provided by all species at the site-date to estimate the total observed pollination for each site-date.

For each crop system, we then defined a single threshold level of pollination to be met at all site-dates in that system (pollination threshold). We defined this threshold as 50% of the mean observed pollination per site-date, with the mean taken across all site-dates in the crop system. Finding the level of biodiversity required to retain 50% of the observed function is a standard approach used in the biodiversity–ecosystem function literature^{34,49,82,83} and is analogous to a half maximal effective concentration in ecotoxicology. While a threshold of 50% of the maximum function has been used in previous work^{49,82}, we chose 50% of the mean function as our threshold because the mean is less sensitive to sampling effort and outliers (that is, the maximum function across samples increases as more samples are taken but we do not expect the same pattern for mean function across samples). Setting the pollination threshold based on mean function makes our results more conservative than previous work, which used the higher threshold of 50% of maximum function. Lastly, because the function threshold is necessarily arbitrary, we also performed a sensitivity analysis to determine whether our choice of threshold influenced the relationship between temporal scale and the number of bee species needed to meet the threshold. Specifically, we repeated the minimum set analysis using thresholds of 10–90% of the mean function.

We first examined how the number of species needed to meet the pollination threshold at a site changes with the number of sampling days over the course of a single year for which the threshold must be met. Across the 3 crop systems, the mean time interval between successive rounds of sampling was 6 d with a range of 1–41 d. We defined the minimum set for m dates as the minimum number of species needed to meet the pollination threshold for all dates in the set $1 \dots m$, starting with the earliest and adding dates 1 at a time in chronological order. We then plotted the cumulative number of bee species needed to meet the function threshold against the number of dates in the set. We repeated this calculation, using the same mean function threshold for each site-year combination in the dataset. Site-years therefore act as replicates in our within-year analysis.

We next examined how the number of species needed to meet the pollination threshold at a site changes with the number of years over which the threshold must be met. We defined the minimum set for n years as the minimum number of species needed to meet the 50% function threshold for all years in the set $1 \dots n$, starting with the earliest and adding years 1 at a time in chronological order. For the threshold to be met for a given year, we required that the threshold be met on each of three sampling days for the blueberry and western watermelon crops. Because the eastern watermelon sites were only sampled 1 d in 3 years of the study, we required the threshold be met on 1 sampling day each year for the eastern watermelon crop. We then plotted the cumulative number of bee species needed

to meet the function threshold against the number of years in the set. We repeated this calculation, using the same mean function threshold, for each site in the dataset. Sites therefore act as replicates in our across-years analysis.

For both the across-season and across-years analyses, we determined the minimum number of species needed to meet the function threshold for each set of sampling dates (as in Winfree et al.³⁴). We calculated the minimum number of species as follows: (1) starting with the earliest date in the set, first add the species that deposits the most pollen at that site-date; (2) continue adding species in descending order of pollen deposition until the function threshold is met or until all species present at that date have been added; (3) if the running species list for the first date also satisfies the threshold for the second date in the set, move on to the third date, and so on; (4) if not, add species present at the next date to the running list until the threshold is met for that date or all the species present at that date have been added; (5) repeat this process for each additional date until a running list is created that meets the function threshold for all dates in the set. Because the species that provide the most function on any single date may not be the best species for meeting the function threshold on other dates, the minimum set may depend on the order in which species are added. To ensure that we found the actual minimum set of species, we used a genetic algorithm optimizer to shuffle the order in which species are added to the list. We performed the optimization using the GAPerm function in the R package gaoptim⁸⁴. To achieve a stable minimum set, we ran the optimizer until the resulting number of species did not change for 30 generations.

Statistical analysis. To assess how the number of species in the minimum set changed with the number of dates within a year over which the threshold must be met, we fitted generalized linear mixed models (GLMMs) on the results of the minimum set analysis using the R package lme4 v.1.1.26 (ref. ⁸⁵). For each model, we also performed tests for overdispersion and extracted R^2 values using the R package performance v.0.7.0⁸⁶. For each crop system, we fitted a full model with the number of dates as a fixed effect and site as a random effect. Because the data were right-skewed, as is typical for count data, we used a Poisson distribution with a log link for the blueberry and western watermelon datasets. For the eastern watermelon dataset, we detected overdispersion (dispersion ratio = 2.126, $\chi^2 = 501.67$, $P < 0.001$). Therefore, we used a negative binomial distribution with a log link function instead of a Poisson distribution, using the R package MASS v.7.3.53.1⁸⁷. For each crop, we fitted a full mixed model with random slopes and intercepts, as well as a mixed model with only the fixed effect and random intercepts. For all three crops, we retained the mixed model with random intercepts because it had the lowest Akaike information criterion (AIC) (Supplementary Table 6).

To assess how the number of species in the minimum set changed with the number of years over which the threshold must be met, we fitted GLMMs using a Poisson distribution with a log link function. For each crop system, we fitted a full model with the number of years as a fixed effect and site as a random slope and intercept. For the western watermelon, we fitted the GLMM using the R package glmmTMB v.1.1.3 instead of lme4 due to small random effects⁸⁸. We also fitted a simpler GLMM with only the fixed effect and random intercepts. For each crop, we selected the model with the lowest AIC. For the eastern watermelon, we retained the full mixed model with random slopes and intercepts. For the western watermelon and blueberry, we retained the simpler mixed model with random intercepts. No overdispersion was detected.

Because our statistical analyses were performed on a derived quantity (the size of the minimum species set), we did not feel it was appropriate to use P values to perform significance tests on our results. We instead report the effect sizes, along with confidence intervals (CIs) and R^2 values, estimated from the GLMMs (Tables 1 and 2). We performed all statistical analyses in R v.4.0.4 (ref. ⁸⁹).

Null model. Even in the absence of real turnover in the community, repeated samples taken from the same species pool may differ in species composition and abundances due to chance differences in which individuals are sampled (sampling effects). Therefore, we developed a null model to determine how much of the increase in the number of bee species needed to meet the pollination threshold was due to biological effects (species turnover over time) versus sampling effects. We estimated the magnitude of sampling effects by resampling from our pool of collected bees. For the within-years analysis, we compared a series of resamples taken from the same site-date (that is, the null) to a series of resamples taken from different dates at the same site (that is, the observed). For the across-years analysis, we compared a series of resamples taken from the same site-year (the null) to a series of resamples taken from different site-years (the observed). Specifically, for each resample, we randomly drew 30 individuals with replacement.

For each series of resamples, we recalculated the pollination threshold (as 50% of the mean pollination across all resamples) and then reran the minimum set analysis on the resampled data (as described above). For the across-season analysis, we repeated the resampling procedure 100 times for each site-year. For the across-years analysis, we repeated the resampling procedure 100 times for each site. We then compared the observed results, in which resamples were drawn from different species pools (that is, different dates) to the null results, in which resamples were drawn from the same species pool (that is, a single site-date). Any difference between the null and observed in the increase in the bee biodiversity

needed over time can be attributed to biological turnover in the bee community rather than sampling effects. We calculated the factor increase in the number of species needed with timescale as: the number of species needed for the maximum number of days minus the number of species needed for 1 d, divided by the number of species needed for 1 d. We also calculated the proportion of the increase that was due to turnover, rather than sampling effects, as the change in the number of species needed in the observed minus the change in number of species needed in the null, divided by the change in the number of species needed in the observed.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data used to generate the results of this study have been deposited in Figshare (<https://doi.org/10.6084/m9.figshare.20083916>, <https://doi.org/10.6084/m9.figshare.20010191>, <https://doi.org/10.6084/m9.figshare.20010179>). The bee specimens on which the data are based are permanently housed at Rutgers University and University of California, Davis.

Code availability

The R code used to generate the results of this study is available on GitHub (https://github.com/nlemanski/Bee_diversity_ecosystem_function).

Received: 13 December 2021; Accepted: 13 July 2022;

Published online: 22 August 2022

References

- Cardinale, B. J. et al. Biodiversity loss and its impact on humanity. *Nature* **486**, 59–67 (2012).
- Hooper, D. U. et al. A global synthesis reveals biodiversity loss as a major driver of ecosystem change. *Nature* **486**, 105–108 (2012).
- Potts, S. G. et al. Global pollinator declines: trends, impacts and drivers. *Trends Ecol. Evol.* **25**, 345–353 (2010).
- Kennedy, C. M. et al. A global quantitative synthesis of local and landscape effects on wild bee pollinators in agroecosystems. *Ecol. Lett.* **16**, 584–599 (2013).
- Cameron, S. A. et al. Patterns of widespread decline in North American bumble bees. *Proc. Natl Acad. Sci. USA* **108**, 662–667 (2011).
- Xu, S. et al. Species richness promotes ecosystem carbon storage: evidence from biodiversity–ecosystem functioning experiments. *Proc. Biol. Sci.* **287**, 20202063 (2020).
- Jochum, M. et al. The results of biodiversity–ecosystem functioning experiments are realistic. *Nat. Ecol. Evol.* **4**, 1485–1494 (2020).
- Isbell, F. et al. High plant diversity is needed to maintain ecosystem services. *Nature* **477**, 199–202 (2011).
- Barnes, A. D. et al. Species richness and biomass explain spatial turnover in ecosystem functioning across tropical and temperate ecosystems. *Philos. Trans. R. Soc. B Biol. Sci.* **371**, 20150279 (2016).
- Manning, P. & Cutler, G. C. Ecosystem functioning is more strongly impaired by reducing dung beetle abundance than by reducing species richness. *Agric. Ecosyst. Environ.* **264**, 9–14 (2018).
- van der Plas, F. Biodiversity and ecosystem functioning in naturally assembled communities. *Biol. Rev. Camb. Philos. Soc.* **94**, 1220–1245 (2019).
- Blüthgen, N. & Klein, A.-M. Functional complementarity and specialisation: the role of biodiversity in plant–pollinator interactions. *Basic Appl. Ecol.* **12**, 282–291 (2011).
- Loreau, M. Biodiversity and ecosystem functioning: a mechanistic model. *Proc. Natl Acad. Sci. USA* **95**, 5632–5636 (1998).
- Loreau, M. et al. Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science* **294**, 804–808 (2001).
- Tilman, D. The ecological consequences of changes in biodiversity: a search for general principles. *Ecology* **80**, 1455–1474 (1999).
- Duffy, J. E., Godwin, C. M. & Cardinale, B. J. Biodiversity effects in the wild are common and as strong as key drivers of productivity. *Nature* **549**, 261–264 (2017).
- Gonzalez, A. et al. Scaling-up biodiversity–ecosystem functioning research. *Ecol. Lett.* **23**, 757–776 (2020).
- Garibaldi, L. A. et al. Wild pollinators enhance fruit set of crops regardless of honey bee abundance. *Science* **340**, 1608–1611 (2013).
- Greenop, A., Woodcock, B. A., Wilby, A., Cook, S. M. & Pywell, R. F. Functional diversity positively affects prey suppression by invertebrate predators: a meta-analysis. *Ecology* **99**, 1771–1782 (2018).
- McGill, B. J. et al. Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecol. Lett.* **10**, 995–1015 (2007).
- Genung, M. A. et al. The relative importance of pollinator abundance and species richness for the temporal variance of pollination services. *Ecology* **98**, 1807–1816 (2017).
- Winfree, R., Fox, J. W., Williams, N. M., Reilly, J. R. & Cariveau, D. P. Abundance of common species, not species richness, drives delivery of a real-world ecosystem service. *Ecol. Lett.* **18**, 626–635 (2015).
- Kleijn, D. et al. Delivery of crop pollination services is an insufficient argument for wild pollinator conservation. *Nat. Commun.* **6**, 7414 (2015).
- Smith, M. D. & Knapp, A. K. Dominant species maintain ecosystem function with non-random species loss. *Ecol. Lett.* **6**, 509–517 (2003).
- Lohbeck, M., Bongers, F., Martinez-Ramos, M. & Poorter, L. The importance of biodiversity and dominance for multiple ecosystem functions in a human-modified tropical landscape. *Ecology* **97**, 2772–2779 (2016).
- Balvanera, P., Kremen, C. & Martinez-Ramos, M. Applying community structure analysis to ecosystem function: examples from pollination and carbon storage. *Ecol. Appl.* **15**, 360–375 (2005).
- Maureaud, A. et al. Biodiversity–ecosystem functioning relationships in fish communities: biomass is related to evenness and the environment, not to species richness. *Proc. Biol. Sci.* **286**, 20191189 (2019).
- Genung, M. A., Fox, J. & Winfree, R. Species loss drives ecosystem function in experiments, but in nature the importance of species loss depends on dominance. *Glob. Ecol. Biogeogr.* **29**, 1531–1541 (2020).
- Potts, S. G., Vulliamy, B., Dafni, A., Neeman, G. & Willmer, P. Linking bees and flowers: how do floral communities structure pollinator communities? *Ecology* **84**, 2628–2642 (2003).
- Tilman, D., Isbell, F. & Cowles, J. M. Biodiversity and ecosystem functioning. *Annu. Rev. Ecol. Syst.* **45**, 471–493 (2014).
- Craven, D. et al. A cross-scale assessment of productivity–diversity relationships. *Glob. Ecol. Biogeogr.* **29**, 1940–1955 (2020).
- Thompson, P. L., Isbell, F., Loreau, M., O'Connor, M. I. & Gonzalez, A. The strength of the biodiversity–ecosystem function relationship depends on spatial scale. *Proc. Biol. Sci.* **285**, 20180038 (2018).
- Qiu, J. & Cardinale, B. J. Scaling up biodiversity–ecosystem function relationships across space and over time. *Ecology* **101**, e03166 (2020).
- Winfree, R. et al. Species turnover promotes the importance of bee diversity for crop pollination at regional scales. *Science* **359**, 791–793 (2018).
- Albrecht, J. et al. Species richness is more important for ecosystem functioning than species turnover along an elevational gradient. *Nat. Ecol. Evol.* **5**, 1582–1593 (2021).
- Yachi, S. & Loreau, M. Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. *Proc. Natl Acad. Sci. USA* **96**, 1463–1468 (1999).
- Shanafelt, D. W. et al. Biodiversity, productivity, and the spatial insurance hypothesis revisited. *J. Theor. Biol.* **380**, 426–435 (2015).
- Naeem, S. & Li, S. Biodiversity enhances ecosystem reliability. *Nature* **390**, 507–509 (1997).
- Tilman, D. Biodiversity: population versus ecosystem stability. *Ecology* **77**, 350–363 (1996).
- Herrera, C. M. Variation in mutualisms: the spatiotemporal mosaic of a pollinator assemblage. *Biol. J. Linn. Soc. Lond.* **35**, 95–125 (1988).
- McCormack, M. L., Adams, T. S., Smithwick, E. A. H. & Eissenstat, D. M. Variability in root production, phenology, and turnover rate among 12 temperate tree species. *Ecology* **95**, 2224–2235 (2014).
- Wright, K. W., Vanderbilt, K. L., Inouye, D. W., Bertelsen, C. D. & Crimmins, T. M. Turnover and reliability of flower communities in extreme environments: insights from long-term phenology data sets. *J. Arid Environ.* **115**, 27–34 (2015).
- Tylianakis, J. M. et al. Resource heterogeneity moderates the biodiversity–function relationship in real world ecosystems. *PLoS Biol.* **6**, e122 (2008).
- Kremen, C. Managing ecosystem services: what do we need to know about their ecology? *Ecol. Lett.* **8**, 468–479 (2005).
- Iserbyt, S. & Rasmont, P. The effect of climatic variation on abundance and diversity of bumblebees: a ten years survey in a mountain hotspot. *Ann. Soc. Entomol. Fr.* **48**, 261–273 (2012).
- Houlahan, J. E. et al. Compensatory dynamics are rare in natural ecological communities. *Proc. Natl Acad. Sci. USA* **104**, 3273–3277 (2007).
- Ernest, S. K. M. & Brown, J. H. Homeostasis and compensation: the role of species and resources in ecosystem stability. *Ecology* **82**, 2118–2132 (2001).
- Kremen, C., Williams, N. M. & Thorp, R. W. Crop pollination from native bees at risk from agricultural intensification. *Proc. Natl Acad. Sci. USA* **99**, 16812–16816 (2002).
- Allan, E. et al. More diverse plant communities have higher functioning over time due to turnover in complementary dominant species. *Proc. Natl Acad. Sci. USA* **108**, 17034–17039 (2011).
- Tilman, D., Reich, P. B. & Knops, J. M. H. Biodiversity and ecosystem stability in a decade-long grassland experiment. *Nature* **441**, 629–632 (2006).
- Awasthi, A., Singh, M., Soni, S. K., Singh, R. & Kalra, A. Biodiversity acts as insurance of productivity of bacterial communities under abiotic perturbations. *ISME J.* **8**, 2445–2452 (2014).
- Tuck, S. L. et al. The value of biodiversity for the functioning of tropical forests: Insurance effects during the first decade of the Sabah biodiversity experiment. *Proc. Biol. Sci.* **283**, 20161451 (2016).

53. Isbell, F. et al. Quantifying effects of biodiversity on ecosystem functioning across times and places. *Ecol. Lett.* **21**, 763–778 (2018).
54. Reich, P. B. et al. Impacts of biodiversity loss escalate through time as redundancy fades. *Science* **336**, 589–592 (2012).
55. Perry, C. J., Sovik, E., Myerscough, M. R. & Barron, A. B. Rapid behavioral maturation accelerates failure of stressed honey bee colonies. *Proc. Natl Acad. Sci. USA* **112**, 3427–3432 (2015).
56. Benjamin, F. E. & Winfree, R. Lack of pollinators limits fruit production in commercial blueberry (*Vaccinium corymbosum*). *Environ. Entomol.* **43**, 1574–1583 (2014).
57. Isaacs, R. & Kirk, A. K. Pollination services provided to small and large highbush blueberry fields by wild and managed bees. *J. Appl. Ecol.* **47**, 841–849 (2010).
58. Zhang, Y., Chen, H. Y. H. & Reich, P. B. Forest productivity increases with evenness, species richness and trait variation: a global meta-analysis. *J. Ecol.* **100**, 742–749 (2012).
59. Baumgärtner, S. The insurance value of biodiversity in the provision of ecosystem services. *Nat. Resour. Model.* **20**, 87–127 (2007).
60. Manning, P. et al. in *Advances in Ecological Research* (eds Eisenhauer, N., Bohan, D. A. & Dumbrell, A. J.) 323–356 (Academic Press, 2019).
61. Naem, S. Species redundancy and ecosystem reliability. *Conserv. Biol.* **12**, 39–45 (1998).
62. CaraDonna, P. J. et al. Interaction rewiring and the rapid turnover of plant–pollinator networks. *Ecol. Lett.* **20**, 385–394 (2017).
63. Gonzalez, A. & Loreau, M. The causes and consequences of compensatory dynamics in ecological communities. *Annu. Rev. Ecol. Evol. Syst.* **40**, 393–414 (2009).
64. Liu, D., Chang, P.-H. S., Power, S. A., Bell, J. N. B. & Manning, P. Changes in plant species abundance alter the multifunctionality and functional space of heathland ecosystems. *New Phytol.* **232**, 1238–1249 (2021).
65. Buschke, F. T., Hagan, J. G., Santini, L. & Coetzee, B. W. T. Random population fluctuations bias the Living Planet Index. *Nat. Ecol. Evol.* **5**, 1145–1152 (2021).
66. Almond, R. E. A., Grooten, M. & Peterson, T. *Living Planet Report 2020: Bending the Curve of Biodiversity Loss* (World Wildlife Fund, 2020).
67. Collen, B. et al. Monitoring change in vertebrate abundance: the Living Planet Index. *Conserv. Biol.* **23**, 317–327 (2009).
68. Wagner, D. L. Insect declines in the Anthropocene. *Annu. Rev. Entomol.* **65**, 457–480 (2020).
69. Stanghellini, M. S., Ambrose, J. T. & Schultheis, J. R. The effects of honey bee and bumble bee pollination on fruit set and abortion of cucumber and watermelon. *Am. Bee. J.* **137**, 386–391 (1997).
70. Winfree, R., Williams, N. M., Dushoff, J. & Kremen, C. Native bees provide insurance against ongoing honey bee losses. *Ecol. Lett.* **10**, 1105–1113 (2007).
71. Tamburini, G., Bommarco, R., Kleijn, D., van der Putten, W. H. & Marini, L. Pollination contribution to crop yield is often context-dependent: a review of experimental evidence. *Agric. Ecosyst. Environ.* **280**, 16–23 (2019).
72. Stanghellini, M. S., Ambrose, J. T. & Schultheis, J. R. Seed production in watermelon: a comparison between two commercially available pollinators. *HortScience* **33**, 28–30 (1998).
73. Reilly, J. R. et al. Crop production in the USA is frequently limited by a lack of pollinators. *Proc. Biol. Sci.* **287**, 20200922 (2020).
74. Greenleaf, S. S. & Kremen, C. Wild bees enhance honey bees' pollination of hybrid sunflower. *Proc. Natl Acad. Sci. USA* **103**, 13890–13895 (2006).
75. Sáez, A. Managed honeybees decrease pollination limitation in self-compatible but not in self-incompatible crops. *Proc. Biol. Sci.* **289**, 20220086 (2022).
76. Brittain, C., Williams, N., Kremen, C. & Klein, A. M. Synergistic effects of non-*Apis* bees and honey bees for pollination services. *Proc. Biol. Sci.* **280**, 20122767 (2013).
77. Aizen, M. A. & Harder, L. D. The global stock of domesticated honey bees is growing slower than agricultural demand for pollination. *Curr. Biol.* **19**, 915–918 (2009).
78. Houlahan, J. E. et al. Negative relationships between species richness and temporal variability are common but weak in natural systems. *Ecology* **99**, 2592–2604 (2018).
79. Winfree, R. Global change, biodiversity, and ecosystem services: what can we learn from studies of pollination? *Basic Appl. Ecol.* **14**, 453–460 (2013).
80. Greenleaf, S. S., Williams, N. M., Winfree, R. & Kremen, C. Bee foraging ranges and their relationship to body size. *Oecologia* **153**, 589–596 (2007).
81. Cariveau, D. P., Williams, N. M., Benjamin, F. E. & Winfree, R. Response diversity to land use occurs but does not consistently stabilise ecosystem services provided by native pollinators. *Ecol. Lett.* **16**, 903–911 (2013).
82. Gamfeldt, L., Hillebrand, H. & Jonsson, P. R. Multiple functions increase the importance of biodiversity for overall ecosystem functioning. *Ecology* **89**, 1223–1231 (2008).
83. Zavaleta, E. S., Pasari, J. R., Hulvey, K. B. & Tilman, G. D. Sustaining multiple ecosystem functions in grassland communities requires higher biodiversity. *Proc. Natl Acad. Sci. USA* **107**, 1443–1446 (2010).
84. Haupt, R. L. & Haupt, S. E. *Practical Genetic Algorithms* (Wiley, 2004).
85. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
86. Lüdtke, D., Makowski, D., Waggoner, P. & Patil, I. performance: Assessment of regression models performance. R package version 0.7.0 <https://doi.org/10.5281/zenodo.3952174> (2020).
87. Venables, W. N. & Ripley, B. D. *Modern Applied Statistics with S* (Springer, 2002).
88. Brooks, M. et al. glmmTMB: Generalized linear mixed models using template model builder. R package version 1.1.3 <https://glmmTMB.github.io/glmmTMB/> (2022).
89. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2021).

Acknowledgements

Funding for this work was provided by the National Science Foundation (NSF) DEB no. 2019863 to R.W., NSF DEB no. 1556885 to N.M.W. and U.S. Department of Agriculture, National Institute of Food and Agriculture, Agriculture and Food Research Initiative no. 65104-05782 to R.W. (principal investigator) and N.M.W. (co-principal investigator).

Author contributions

N.J.L., N.M.W. and R.W. conceived the research question and study design. N.M.W. and R.W. oversaw data collection. N.J.L. performed the analyses and wrote the original manuscript draft. All authors reviewed and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41559-022-01847-3>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41559-022-01847-3>.

Correspondence and requests for materials should be addressed to Natalie J. Lemanski.

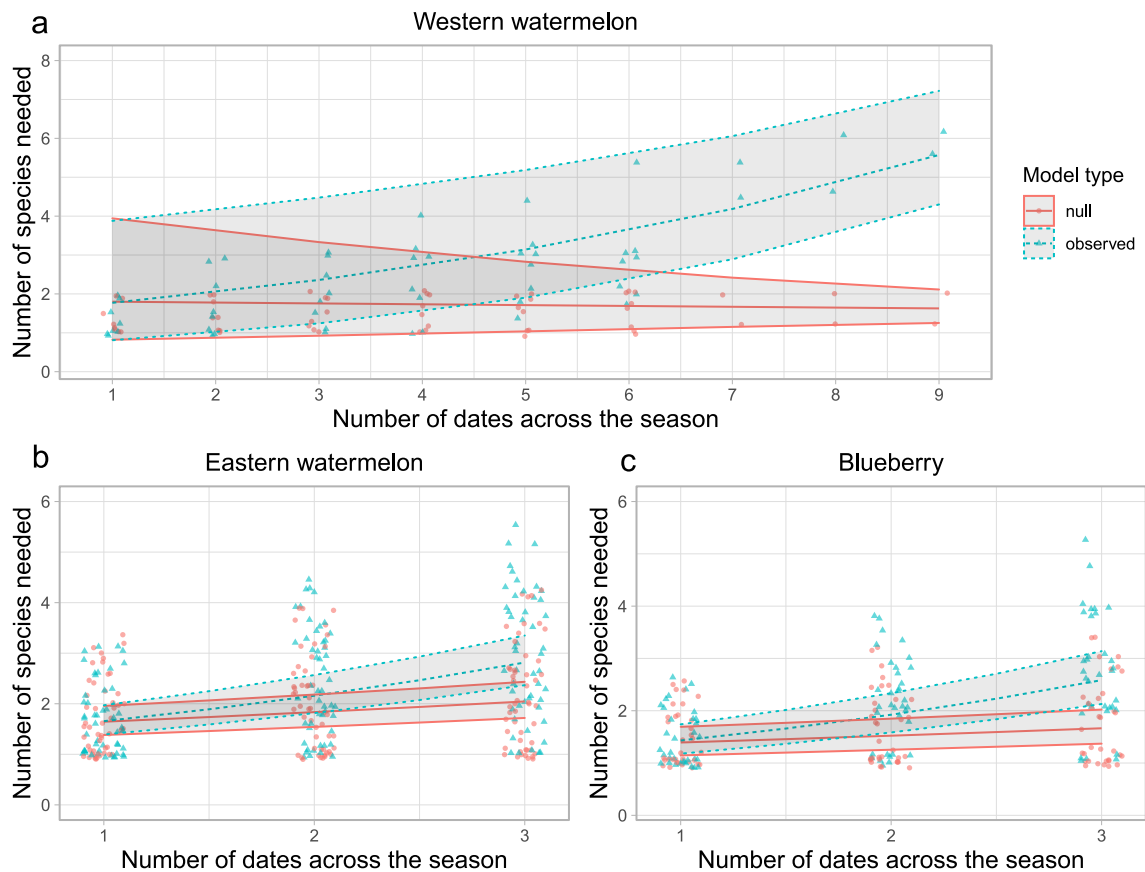
Peer review information *Nature Ecology & Evolution* thanks Marcelo Aizen, Amy Iler and Joanne Bennett for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.

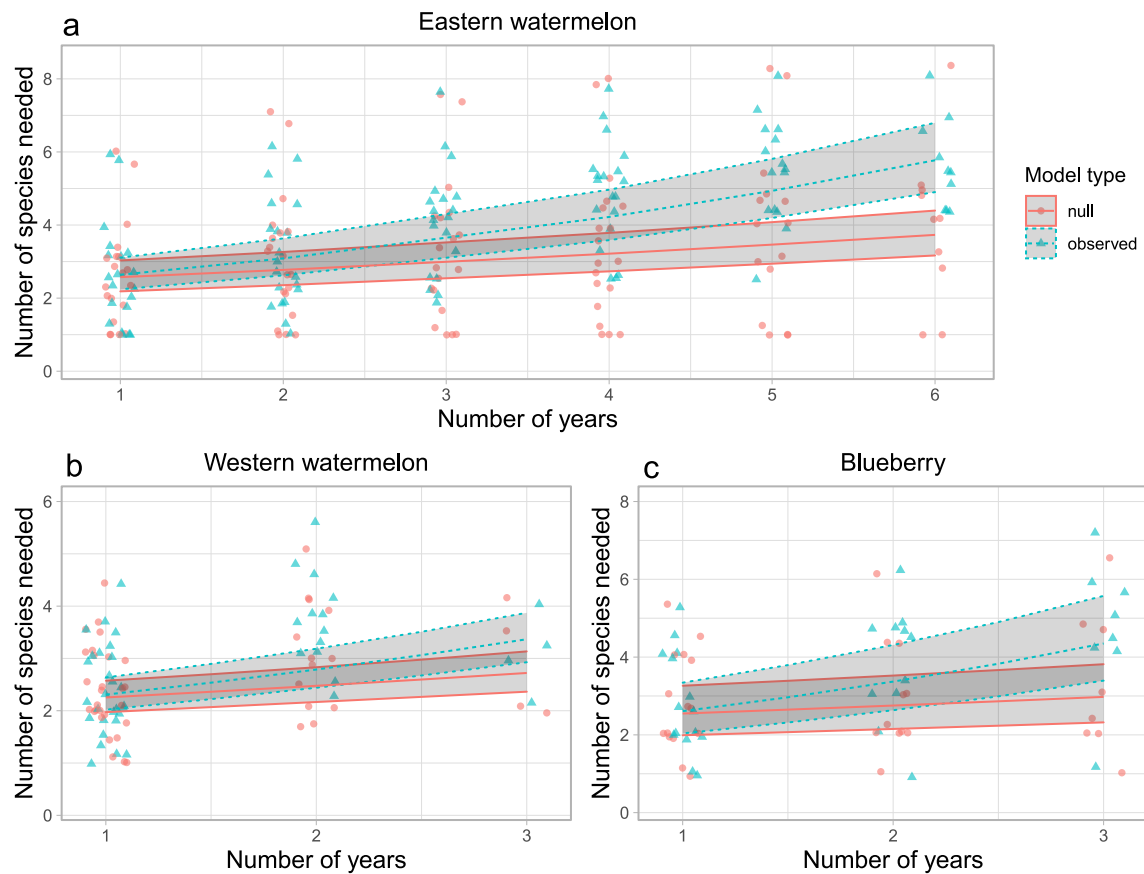
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

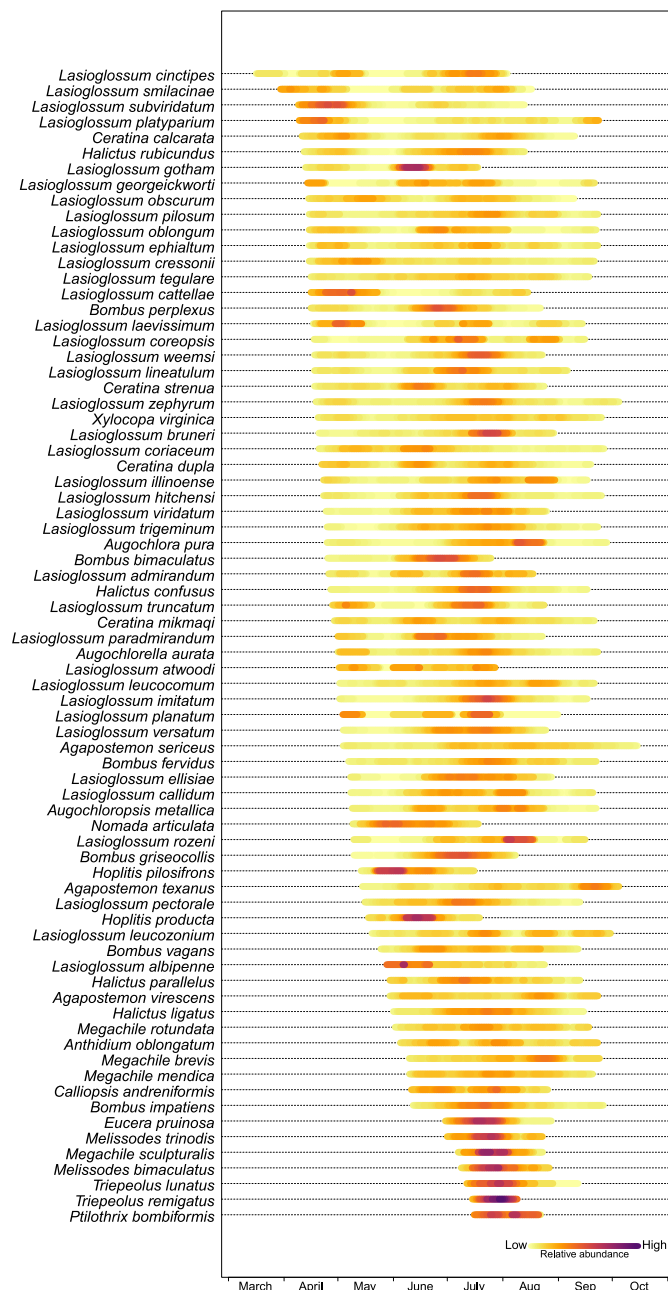
© The Author(s), under exclusive licence to Springer Nature Limited 2022



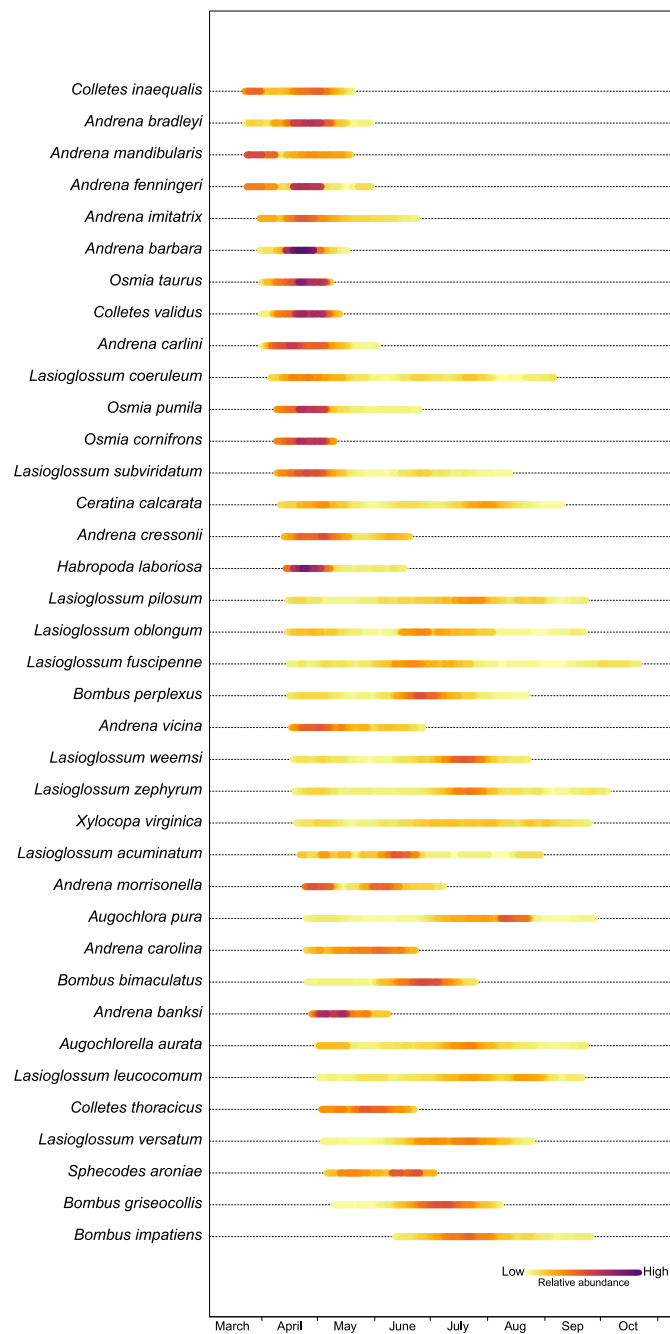
Extended Data Fig. 1 | Minimum set analysis on a subsample of the data in which 30 individuals were randomly drawn each different date within a site-year. We compared the results (dashed blue line) to a null model (solid red line) in which subsamples of 30 individuals were all drawn from the same date. Shaded areas represent the 95% confidence intervals across 100 replicates of the sampling process. Each point represents the mean across all replicates for a single site-year. The difference between the endpoints of the observed and null accumulation curves represents the increase in the minimum set that is due to turnover across days within a year.



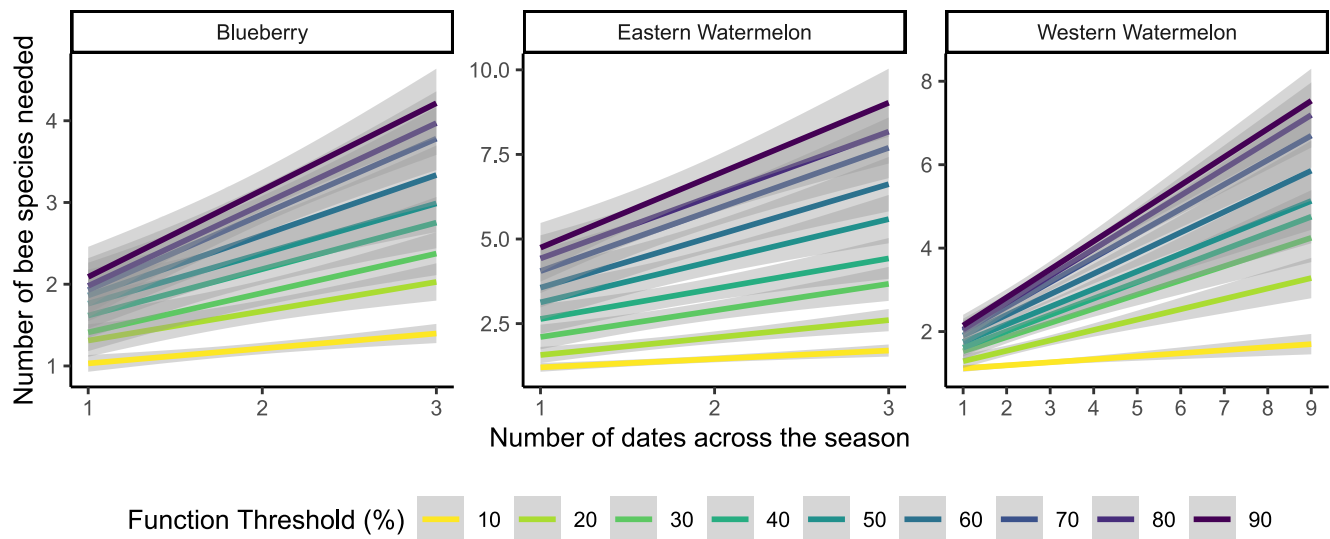
Extended Data Fig. 2 | Minimum set analysis on a subsample of the data in which 30 individuals were randomly drawn from each different year within a site. We compared the results (dashed blue line) to a null model (solid red line) in which subsamples of 30 individuals were all drawn from the same year. Thus, confidence intervals include both variation across sites in the number of bee species needed, and uncertainty from the sampling process. Specifically, shaded areas represent the 95% confidence intervals across 100 replicates of the sampling process. Each point represents the mean across all replicates for a single site. The difference between the endpoints of the observed and null accumulation curves represents the increase in number of species needed that is due to turnover in species composition across years.



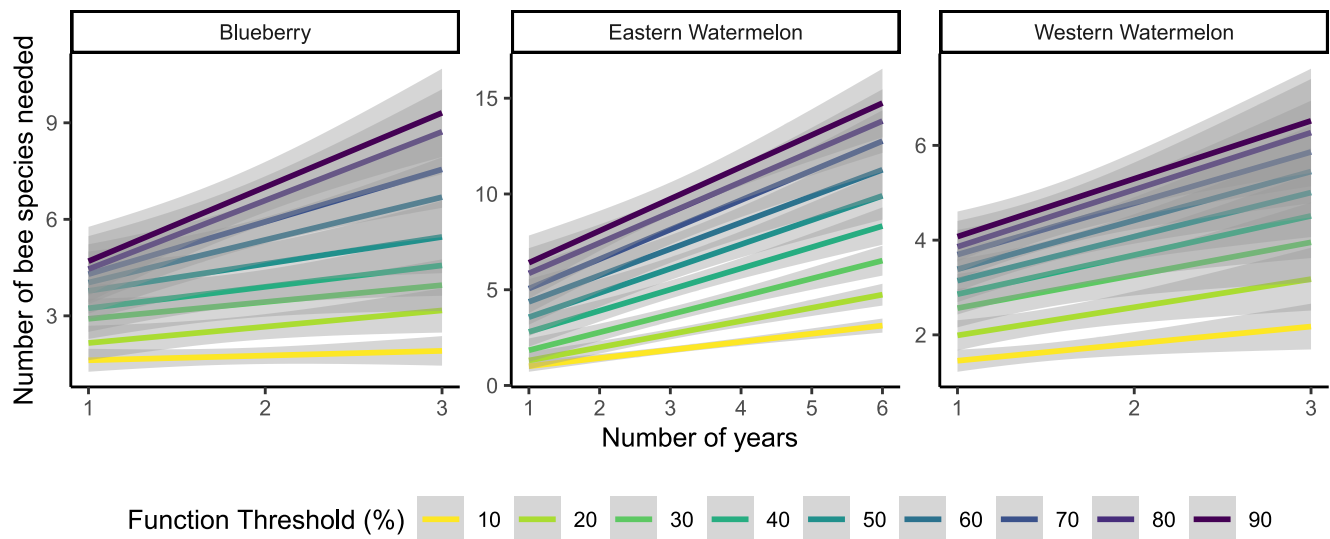
Extended Data Fig. 3 | Phenological windows of bee species observed at eastern watermelon farms. Horizontal lines show the timing of peak abundance for each bee species observed at the eastern watermelon farms. Phenological data is based on all observations of that species across all datasets collected by the lab. Comparable data for species visiting western watermelon farms was not available.



Extended Data Fig. 4 | Phenological windows of bee species observed at blueberry farms. Horizontal lines show the timing of peak abundance for each bee species observed at the blueberry farms. Phenological data is based on all observations of that species across all datasets collected by the lab.



Extended Data Fig. 5 | Sensitivity analysis of the effect of function threshold on the number of bee species needed for analyses done across the growing season within one year. The function threshold is a percentage of the mean observed pollination per site-date, averaged across all site-dates.



Extended Data Fig. 6 | Sensitivity analysis of the effect of function threshold on the number of bee species needed for analyses done across years. The function threshold is a percentage of the mean observed pollination per site-date, averaged across all site-dates.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- | | | |
|-------------------------------------|-------------------------------------|--|
| ☐ | ☒ | The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☐ | ☒ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☐ | ☒ | Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection No software was used to collect data for this study.

Data analysis All analyses were performed using R (version 4.0.4). All code for the analyses, null model, and figures are deposited on GitHub (https://github.com/nlemanski/Bee_diversity_ecosystem_function).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The data used to generate the results of this study have been deposited in Figshare (doi.org/10.6084/m9.figshare.20083916, doi.org/10.6084/m9.figshare.20010191, doi.org/10.6084/m9.figshare.20010179). The bee specimens upon which the data are based are permanently housed at Rutgers University and University of California Davis.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	The study used three existing datasets of bees visiting crop flowers. The bee community at a single farm field was sampled on different days throughout the year and for multiple years. In this study we calculated the minimum number of bee species needed to provide a threshold level of pollination services at a single site across different periods of time (number of dates within a year and number of years). Sites acted as replicates for the across-years analysis and site-years acted as replicates for the within-years analysis. The number of sites sampled was 16 for blueberry, 25 for eastern watermelon, and 36 for western watermelon.
Research sample	Samples consisted of all bees (Hymenoptera, clade Anthophila) observed on crop flowers within a predefined 50- to 200-m ² transect at a single farm field on a single date. The samples collected are meant to represent the community of bees pollinating the crop at that site-date. Farm fields were far enough apart to represent independent bee communities. In addition, farm field is a scale at which pollination services are relevant to individual producers.
Sampling strategy	Samples consisted of all bees observed visiting crop flowers within a predefined 50- to 200-m ² transect. Sampling was repeated 1-9 times per year and for 3-6 years at each site. The number of sites sampled was 16 for blueberry, 25 for eastern watermelon, and 36 for western watermelon.
Data collection	The observers walked the transect, net collecting all bees observed visiting crop flowers. Netting was repeated in four ten-minute periods for western watermelon sites and in three 20-minute periods for eastern watermelon and blueberry sites. Sampled bees were then preserved and identified to species in the lab.
Timing and spatial scale	The blueberry study sites consisted of 16 farms within a 35 by 55 km area. The eastern watermelon study sites consisted of 25 farms within a 90 x 60 km area. The western watermelon study sites consisted of 36 farms within a 64 x 49 km area. Western watermelon data was collected in 2010, 2011, and 2012, between May 29 and Aug. 21, with 2-41 days (mean 6 days) between consecutive sampling rounds. Eastern watermelon data was collected in 2004, 2005, 2007, 2008, 2010, 2011, and 2012, between June 26 and Aug. 20, with 3-12 days (mean 5 days) between consecutive sampling rounds. Blueberry data was collected in 2010, 2011, and 2012, between Apr. 13 and May 12, with 1-18 days (mean 7 days) between consecutive sampling rounds.
Data exclusions	Data from 2004 was excluded for the eastern watermelon data. This year was excluded because it was a pilot study with lower sampling effort and slightly different sampling methods than the other years of the study.
Reproducibility	No attempts have been made to reproduce the analysis presented in the current manuscript.
Randomization	Randomization is not relevant to the current study because no experimental manipulation of samples was done. Only observational data (the number and species identity of bees visiting crop flowers) was collected.
Blinding	Blinding was not relevant to this study. Observers collected all bees observed on flowers within the transects.
Did the study involve field work?	☒ Yes ☐ No

Field work, collection and transport

Field conditions	Samples were collected during clear (sunny to overcast) weather to ensure that bees would be active.
Location	The datasets used in the study were collected in agricultural fields located in southern New Jersey (blueberry data), central New Jersey and east-central Pennsylvania (eastern watermelon data), and the Northern Central Valley of California (western watermelon data).
Access & import/export	All bee specimens were collected from private farms with the permission of the landowners. No samples were imported/exported across international borders.
Disturbance	Bee specimens were collected from the field in order to sample the community of bees visiting crop plants. Previous work suggests that the number of individuals collected in a study of this type is not large enough to have a noticeable impact on the bee populations at the study sites.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	The study did not involve laboratory animals.
Wild animals	Wild bees were collected by net in the field and killed on site using standard procedures. Collected specimens were then pinned, preserved, and identified in the lab.
Field-collected samples	The study did not involve samples collected from the field.
Ethics oversight	No ethical approval was required because the study only involved invertebrate animals (insects).

Note that full information on the approval of the study protocol must also be provided in the manuscript.