



OPEN Drivers of arthropod biodiversity in an urban ecosystem

Jayne M. M. Lewthwaite^{1,10}, Teagan M. Baiotto^{1,10}, Brian V. Brown², Yan Yin Cheung¹, Austin J. Baker^{1,2}, Charles Lehnen^{1,3}, Terrence P. McGlynn^{2,4}, Vaughn Shirey^{1,5}, Lisa Gonzalez⁶, Emily Hartop⁷, Peter H. Kerr⁸, Eric Wood⁹ & Laura Melissa Guzman^{1✉}

Our world is becoming increasingly urbanized with a growing human population concentrated around cities. The expansion of urban areas has important consequences for biodiversity, yet the abiotic drivers of biodiversity in urban ecosystems have not been well characterized for the most diverse group of animals on the planet, arthropods. Given their great diversity, comparatively small home ranges, and ability to disperse, arthropods make an excellent model for studying which factors can most accurately predict urban biodiversity. We assessed the effects of (i) topography (distance to natural areas and to ocean) (ii) abiotic factors (mean annual temperature and diurnal range), and (iii) anthropogenic drivers (land value and amount of impervious surface) on the occurrence of six arthropod groups represented in Malaise trap collections run by the BioSCAN project across the Greater Los Angeles Area. We found striking heterogeneity in responses to all factors both within and between taxonomic groups. Diurnal temperature range had a consistently negative effect on occupancy but this effect was only significant in Phoridae. Anthropogenic drivers had mixed though mostly insignificant effects, as some groups and species were most diverse in highly urbanized areas, while other groups showed suppressed diversity. Only Phoridae was significantly affected by land value, where most species were more likely to occur in areas with lower land value. Los Angeles can support high regional arthropod diversity, but spatial community composition is highly dependent on the taxonomic group.

Urban ecosystems, including the developed infrastructure of towns and cities, are among the most widespread forms of land use globally¹. Cities are novel, highly manicured ecosystems composed of developed structures intermixed with vegetation and impervious surfaces where human populations are the main drivers of change^{2,3}. Thus, the traditional view of urban ecosystems is that they are generally devoid of biodiversity, especially when compared with natural areas. Nevertheless, a surge of research over the past few decades has highlighted the critical role of cities in harboring distinct flora and fauna—especially for vertebrates such as birds and mammals, which have been the predominant focus of urban wildlife study^{4–8}. Invertebrates, including arthropods, which are the most diverse group of animals on the planet, have been comparably less studied⁹. For example, despite recent efforts in describing new insect species within urban ecosystems, the number of insects in urban environments, the extent of their ranges, and other aspects of their life history are generally unknown¹⁰. One exception to this is urban pollinator research, which has rapidly grown in scope and volume in recent years. This research has shown that urban areas can act as refuges¹¹ for pollinators when there is adequate suitable habitat¹², but that these habitats disproportionately favor generalist species¹³. Given the ecological importance of arthropods in virtually all terrestrial ecosystems, and considering the widespread declines in arthropods across the globe, it is imperative to expand the taxonomic scope of urban ecology research to describe the distribution and ecology within cities in lesser-studied arthropod groups.

¹Marine and Environmental Section, Department of Biological Sciences, University of Southern California, Los Angeles 90089, USA. ²Department of Entomology, Natural History Museum of Los Angeles County, Los Angeles 90007, USA. ³Human Evolutionary Biology Section, Department of Biological Sciences, University of Southern California, Los Angeles 90089, USA. ⁴Department of Biology, California State University Dominguez Hills, Carson 90747, USA. ⁵Department of Biology, Georgetown University, Washington, DC 20057, USA. ⁶Natural History Museum of Los Angeles County, Los Angeles 90007, USA. ⁷Center for Integrative Biodiversity Discovery, Museum für Naturkunde, Berlin, Germany. ⁸California State Collection of Arthropods, CDFA Plant Pest Diagnostics Center, Sacramento, CA 95832, USA. ⁹Department of Biological Sciences, California State University Los Angeles, 5151 State University Drive, Los Angeles 90032, USA. ¹⁰These authors contributed equally: Jayne M. M. Lewthwaite and Teagan M. Baiotto. ✉email: guzmanur@usc.edu

Characterizing the drivers of arthropod distributions in cities is essential because of their critical role in urban ecosystems. They are integral components of food webs that provide a broad range of ecosystem services in cities¹⁴. When food webs in cities become simplified, the functional roles of the key players are amplified¹⁵. Considering that arthropods fill many trophic roles, maintaining arthropod biodiversity is imperative for many other organisms. For example, they are a critically important food source for urban bird species. Cities can serve as reservoirs of pollinator diversity, and green areas within them can serve as climatic “stepping stones” to connect habitats across the landscape^{16–18}. Additionally, arthropods are important for decomposition processes, nutrient cycling, and biological control of pests¹⁹. They also provide a window into other taxa that receive even less attention in urban ecology, such as saprotrophic organisms²⁰ and cryptic soil biota²¹. While there are many immediate steps that we can take to protect and foster urban arthropod biodiversity²², more effective policies and practices require a more nuanced understanding of the characteristics of urban environments that support a diverse and resilient arthropod fauna²³, as well as the taxonomic groups most susceptible to increasing urbanization²⁴.

Arthropod distributions in urban environments can be driven by a wide variety of factors, including climate-related variables, as well as income and human-related variables. Disentangling the effects of each of these becomes challenging, particularly in urban areas where micro-climates and micro-habitats depend on both large-scale climate patterns, as well as anthropogenic structures. For example, the amount of water available for arthropods depends both on the amount of precipitation, as well as the irrigation patterns in a given area. Using only observational data, ecologists have often relied on techniques like model selection in hopes of extracting cause-and-effect relationships. However, these techniques can lead to inaccurate or even biased results^{25,26}. For example, the inclusion of a third variable (Z) in a model can unintentionally remove the causal effect between X and Y if there is a causal link between X and Z that is unaccounted for²⁷. Causal inference approaches have emerged as a powerful tool to overcome this hurdle as they are concerned with predicting the consequences of intervening in a system (e.g., how X impacts Y) by considering the effect additional variables (e.g. Z) could have on causal relationships between X and Y²⁸. Essentially, the framework involves (i) identifying variables that are likely important in the system; (ii) making causal assumptions about their relationships between variables, (iii) identifying which effects to estimate; and (iv) exploring the consequences of changing the assumptions to see whether inferences hold under different sets of assumptions²⁵.

The greater Los Angeles metropolitan area (L.A.) is one of the largest urbanized areas in the United States. Due to the development patterns, in many places throughout the city, it is a short distance between the surrounding mountainous natural areas and high-density urban areas (e.g. Santa Monica Mountains to downtown and Century City), which creates a narrow urban-to-natural gradient. Along this gradient, numerous factors shift, including elevation, mean and diurnal temperature, and income distribution (via the luxury hypothesis described below), that likely affect arthropod species distributions. Although these narrow gradients may suppress diversity in many taxonomic groups, they could alternatively create high regional diversity (along with high beta diversity) and abundance levels in arthropods due to their relatively small home range sizes, high sensitivity to climatic drivers, and the variety of micro-habitats found in urbanized landscapes. For example, temperature has been shown to explain spatial insect beta diversity gradients²⁹. Furthermore, urban structures can either increase³⁰ or decrease²⁴ insect richness, depending on the interplay between densification vs. sprawl as well as the specific life history requirements of the group^{31,32}.

Beyond climatic drivers, local-scale influences like urban green areas can dramatically affect arthropod occurrence³³. However, L.A. has a long and fraught history of real estate inequity, which has left a legacy on land value across the city³⁴ where high-income neighborhoods have disproportionately higher levels of vegetation cover and green area than lower-income neighborhoods³⁵. This phenomenon is known as the “luxury effect” and hypothesizes a positive relationship between wealth and biodiversity within urban areas and has been demonstrated in groups such as birds³⁶, bats³⁷, and terrestrial mammals³⁸. However, although it may directly impact the amount of habitat and thus the likelihood of occurrence of many arthropod species, very little research to date has examined the luxury effect on arthropod biodiversity³⁹.

In this study, we characterize arthropod diversity within 6 different groups across an extensive metropolitan area that spans climatic and urban gradients. Specifically, we sought to disentangle the relative importance of (i) topography (distance to natural areas and distance to ocean) (ii) abiotic factors (mean annual temperature and diurnal range, and relative humidity), and (iii) anthropogenic drivers (land value and amount of impervious surface as proxies for luxury effect) in determining the distributions of individual species as well as comparing responses within and among several major taxonomic groups. We accomplished this with the development of a large-scale and extensive monitoring network across L.A., which sampled arthropods from 2014 to 2018 in various phases, and the application of multi-species occupancy modeling and causal inference tools.

Results

Overall, we found that arthropod diversity patterns varied dramatically across space in L.A. and diverged significantly across the 6 taxonomic groups we examined. Much of this diversity and divergence between groups can be attributed to the geographical variation in topology, environment, and anthropogenic features across this relatively large area and each group’s and species’ individual response to those predictors. This is apparent from the extensive between- and within-group variation that we observed in both the direction and magnitude of predictor importance on arthropod occupancy. Our predictors were all transformed to the same scale in order to enable meaningful comparisons between groups and drivers. Additionally, models were constructed such that they estimate the marginal contribution of each predictor; i.e. predictors are directly attributable and not confounded with other predictors (see “Methods” section).

The 60 sites included in this study were widely distributed across L.A.: as far west as the coastline, eastwards to San Bernardino in the San Gabriel Valley, northwards to the San Fernando Valley, and southwards to Long Beach

(Fig. 1). The furthest sites were 122 km apart from one another. A large proportion of sites were concentrated in central L.A. In total, 236 species of arthropods were included in our study.

When comparing predictors of occurrence across taxonomic groups, topographical elements consistently had the strongest effects on arthropod occurrence (Fig. 2a,b). Although the mean effect of distance to ocean was centered on 0 and was not statistically significant for all groups, it had the largest within-group variation. Phoridae in particular had a wide variation in responses, with a mean positive effect that overlapped with 0 (0.091, 95% CI [- 0.033, 0.209]) (i.e. many species were more likely to occur farther from the ocean, but this was not universal amongst phorid species). Similarly, distance to natural area had large within- and across-group variation, although this effect was only statistically significant in Drosophilidae (mean estimate of - 0.175, 95% CI [- 0.356, - 0.037]), which were more likely to occur in close proximity to natural areas. Syrphidae occupancy also increased near natural areas, but this was not a statistically significant effect (mean estimate of - 0.127, 95% CI [- 0.311, 0.042]). Meanwhile, there was no significant mean effect on Tipuloidea, Araneae, and Mycetophilidae (Fig. 2a,b), and interestingly, we found a positive (but not significant) effect of distance to natural areas (mean estimate of 0.050, 95% CI [- 0.042, 0.130]) on Phoridae occurrence (i.e. some species were more likely to occur farther from natural areas).

In terms of environmental predictors, the effect of mean annual temperature on species occurrence was not consistent within- or across groups on average, except we found a significant negative relationship in Syrphidae and Mycetophilidae (Fig. 2c,d), indicating that all syrphid and mycetophilid species collected in this study are more likely to occur in cooler sites. Notably, there was a large within-group variation in Phoridae. Meanwhile, mean diurnal temperature range had a consistent and negative effect across all groups, though this effect was only statistically significant in Phoridae. Phorid species were less likely to occur in sites that experienced more extreme fluctuations in temperatures. These results indicate that in L.A., mean temperature on its own does not have a directional effect on species' distributions, but the diurnal range does.

For anthropogenic predictors of arthropod distributions, there was no mean effect of the amount of impervious surface surrounding a site on occurrence in some groups (Drosophilidae and Phoridae) but a negative effect in other groups (Syrphidae and Tipuloidea) and a positive effect in others (Araneae and Mycetophilidae), though none of these effects were statistically significant (Fig. 2e,f). Similarly, land value had little predictive power in most groups, though interestingly, most Phoridae species were significantly more likely to occur in areas with lower land value (mean estimate of - 0.207, 95% CI [- 0.334, - 0.083]).

Some of the predictors do not seem to have an overall positive or negative effect on occurrence, but they do correspond to changes in community composition. Although Fig. 1a shows that overall arthropod diversity is lowest in the most urbanized areas of L.A., when this map is subdivided by taxonomic group (Fig. 1b–g), this is not the case for all groups. For example, drosophilid—and to some extent phorid and syrphid diversity—is often higher in central L.A. compared to the outskirts.

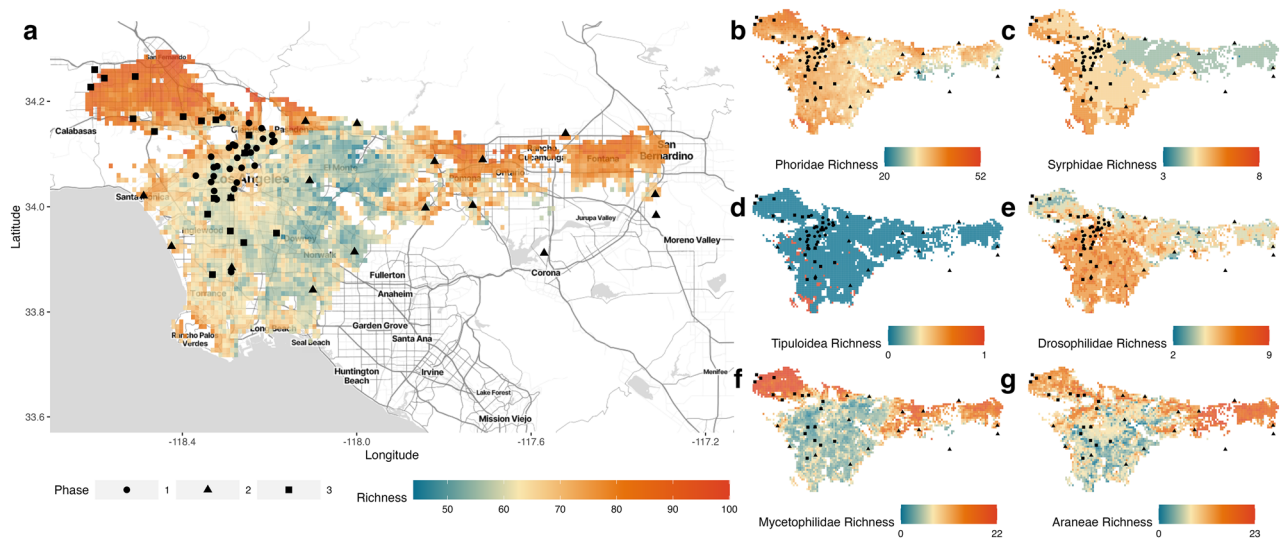


Figure 1. Spatial interpolation of diversity across the greater Los Angeles area. (a) Projected richness for all combined groups ($n = 236$ species) is shown using the continuous color map. (b–g) Projected richness is also shown for each individual taxonomic group modelled: Phoridae ($n = 108$ species), Syrphidae ($n = 34$ species), Tipuloidea ($n = 20$ species), Drosophilidae ($n = 27$ species), Mycetophilidae ($n = 23$ species), and Araneae ($n = 24$ species). BioSCAN sampling sites are overlaid with shapes displaying the collection phase (year of sampling). Across groups, total richness is not obviously driven by any single predictor but rather a result of complex group- and species-specific effects. The extent of the spatial map is based on (1) available land value data, which is confined to the Los Angeles and San Bernardino Counties and (2) the bounding of every predictor to 3 standard deviations away from the mean value present at BioSCAN sampling sites. Grid cells are at a 1 km resolution. R packages `ggplot2`, `ggmap`, and `basemaps` were used to create this map. Map tiles by Stamen Design, under CC BY 4.0. Data by OpenStreetMap, under ODbL.

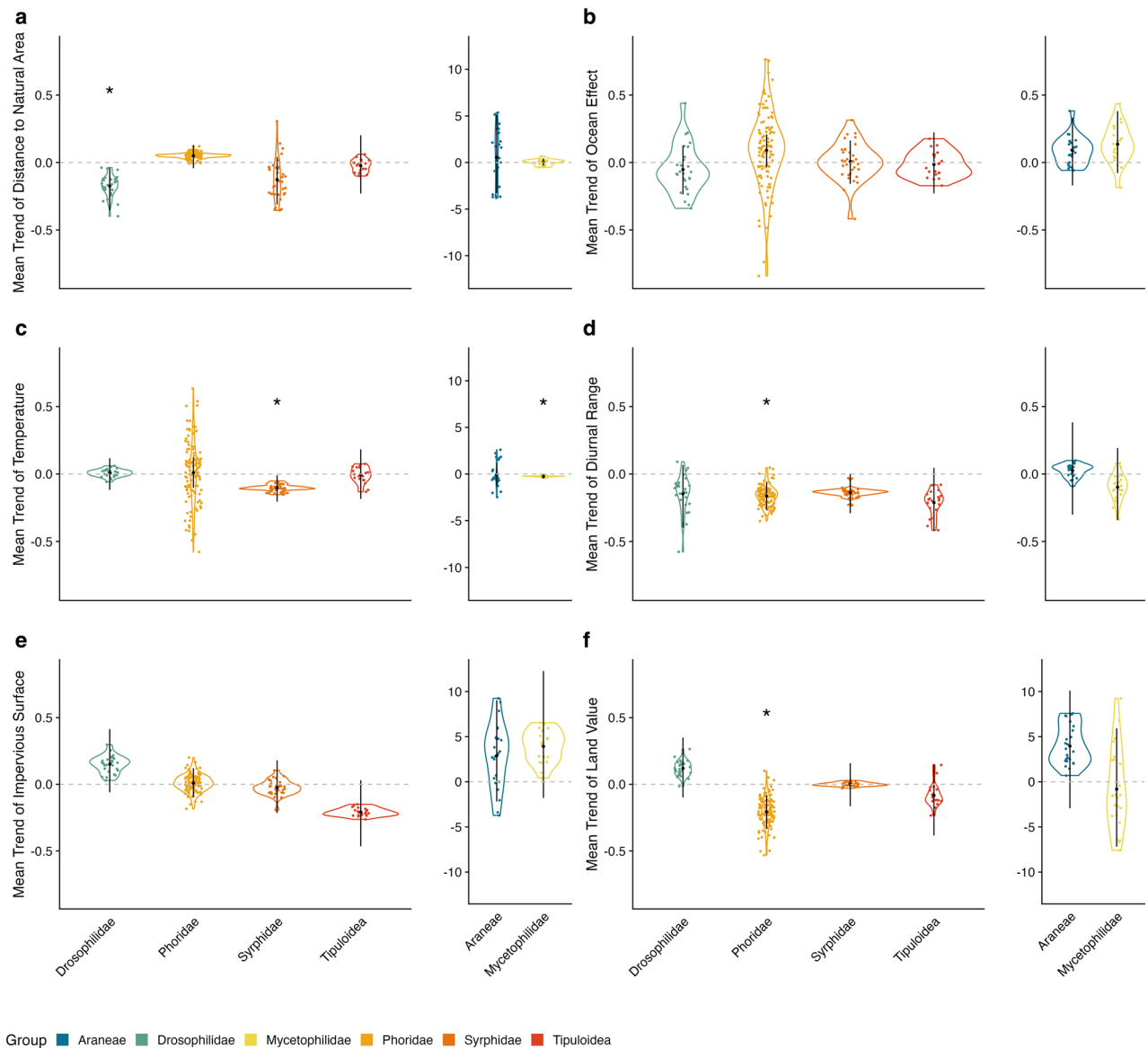


Figure 2. Across-group predictors of occupancy across environmental variables. **(a)** Distance to Natural Area. **(b)** Distance to Ocean. **(c)** Temperature. **(d)** Diurnal Range. **(e)** Impervious Surface. **(f)** Land Value. The six colors displayed in the plots represent different arthropod groups: Araneae (dark blue), Drosophilidae (dark turquoise), Mycetophilidae (yellow), Phoridae (yellow–orange), Syrphidae (orange), and Tipuloidea (red). There are 24, 27, 23, 108, 34, and 20 species in Araneae, Drosophilidae, Mycetophilidae, Phoridae, Syrphidae, Tipuloidea, respectively. Each colored point represent a species from an arthropod group with the same color. The grey horizontal line at $y = 0$ represents no significant trend. The black point represents the average trend of an arthropod group. The black vertical line represents the 95% confidence interval of the average group trend. Stars above each group represent group-level 95% confidence intervals that do not overlap with zero. Note that the effects of land value and impervious surface are measured within a 0.25 km radius for Phoridae and Araneae, and 0.5 km for Drosophilidae, Syrphidae, Tipuloidea and Mycetophilidae. Araneae and Mycetophilidae have different y-axis values for some predictors.

Group-level trends are not representative of all species contained within them, and accordingly, we also characterized within-group variation in responses. We focus on Phoridae in the main text, as they are the most diverse and best-characterized group in this dataset, but similar plots for all other taxonomic groups are located in the Supplemental Material (Supp. Figs. S7–S11). In terms of topographical variables, we found that distance to natural areas had little effect on occupancy (Fig. 3a), while individual phorid species do respond strongly to increasing distance to the ocean. Specifically, two species (*Megaselia donahuei* and *Puliciphora occidentalis*) are significantly more likely to occur at sites near the ocean, while six species (*Diplonevra setigera*, *Megaselia carthayensis*, *Megaselia losangelensis*, *Megaselia simunorum*, *Megaselia spiniclasper*, and *Trophodeinus furcatus*) are more likely to occur at sites farther from the ocean (Fig. 3b). In terms of climatic variables, we found that both temperature and diurnal range had strong effects on individual species (Fig. 3c,d), but relative humidity did not

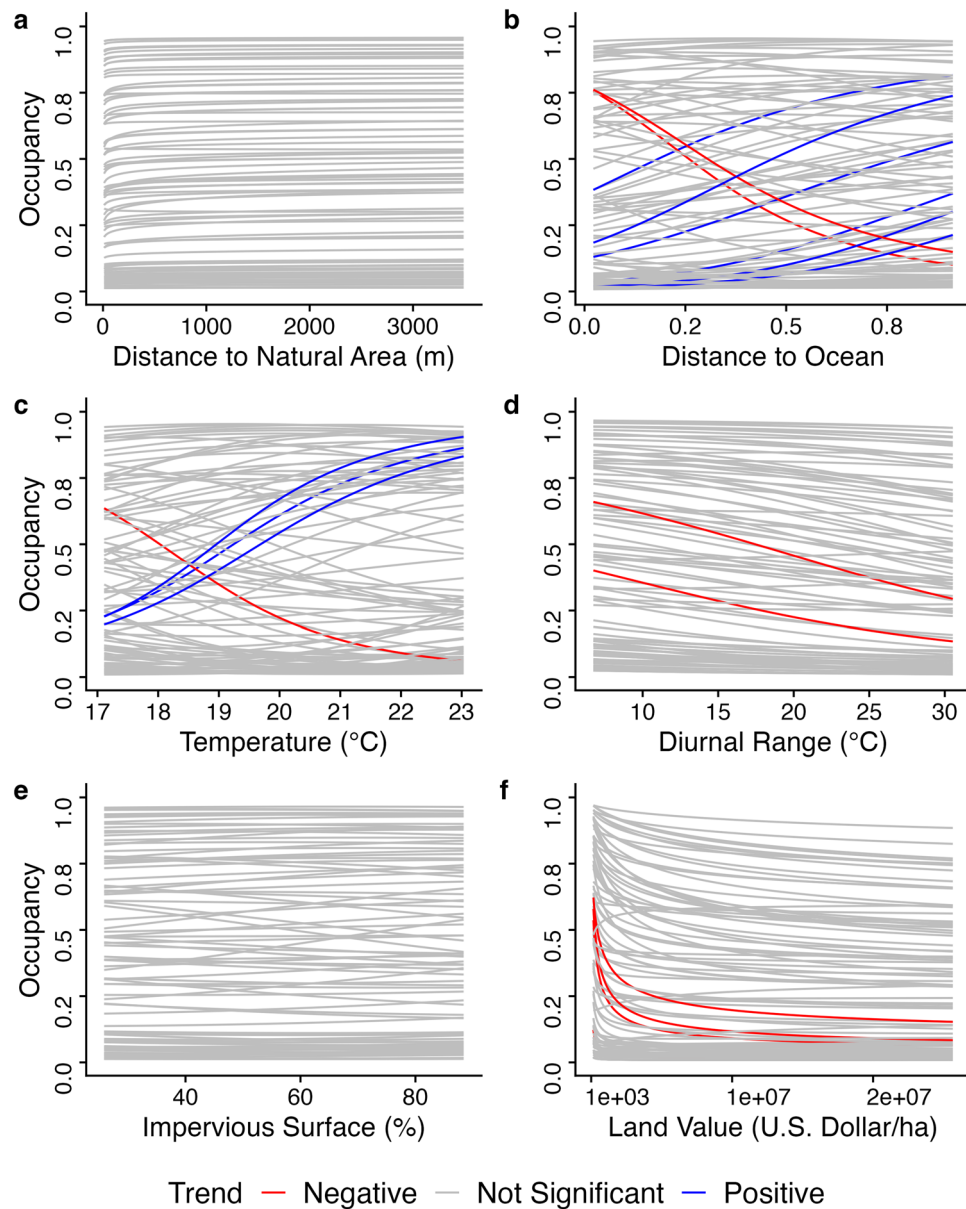


Figure 3. Within-group predictors of occupancy in Phoridae. 108 species are represented in this group. Each plot is showing the single effect of an environmental variable by keeping the other environmental variables in the same model at their average value from the sampling data. (a) Distance to Natural Area. (b) Distance to Ocean. (c) Temperature. (d) Diurnal Range. (e) Percentage of Impervious Surface. (f) Land Value. Each line represents the effect of the environmental variable on an individual species. Red lines indicate a statistically significant negative relationship between the occupancy of a species and an environmental variable, blue lines are statistically significant positive relationships, and grey lines indicate non-significant relationships.

(Supp. Fig. S6). For temperature, we found three species where occupancy increased significantly with increasing temperature (*Megaselia donahuei*, *Puliciphora occidentalis*, and *Spiniphora bergenstammii*), and one species where occupancy decreased significantly with increasing temperature (*Megaselia rodriguezorum*, Fig. 3c). The mean effect of diurnal range was negative on Phoridae as a whole, and indeed we found most species to decrease in occupancy as the diurnal range increased, though only two of those species did so significantly (*Megaselia renwickorum* and *Spiniphora bergenstammii* Fig. 3d). Finally, in terms of anthropogenic drivers, we found that impervious surface had a very small and statistically insignificant effect on all individual phorid species (Fig. 3e), but land value had a consistently negative effect. However, only four species decreased significantly as land value increased (*Anevrina variabilis*, *Megaselia hoggorum*, *Megaselia scalaris*, and *Phalacrotophora halictorum* Fig. 3f). Further, the decrease of occupancy with increasing land value was most dramatic at the lowest end of the distribution, indicating that above a certain threshold land value, phorid occurrence was no longer impacted. Some of the species listed above show similar responses to seemingly opposing predictors (e.g. *Megaselia donahuei*

is more likely to occur at sites near the ocean, and also in sites with higher mean temperatures; whereas mean temperature and distance to ocean negatively covary in space (Fig. S15 and S17). We note that these are effects marginal effects, after accounting for possible confounders.

Discussion

As cities across the world continue to grow at unprecedented rates, determining features of urban ecosystems that support or hinder biodiversity is a critical and pressing need. Here, we used a robust arthropod dataset, coupled with topographic, abiotic, and anthropogenic predictor variables, to uncover drivers of arthropod diversity and occupancy patterns across L.A. Our results highlight the differential responses of arthropods to various geographic and biological features of L.A. L.A. is surrounded by the Pacific Ocean—to the west and south—and mountains—to the north and east—that are primarily protected lands. Therefore, arthropods may be more diverse the closer they are to potential source populations in the protected areas. Abiotic factors were also key drivers of arthropod diversity patterns where the mean diurnal temperature range at a given location was a stronger predictor than average temperature values. Given that arthropods are ectothermic, species range limits may be governed by their physiological tolerances for a range of temperature conditions. Lastly, anthropogenic variables, such as land cover, were less strong in their predictive power, which is in contrast to studies of birds and other taxa³⁶, bats³⁷, and terrestrial mammals³⁸. Taken together, our study uncovers important relationships among predictor and response variables and highlights potential important paths for future research.

One of our most important findings was the large heterogeneity in responses both within taxonomic groups as well as across them. For example, land value and amount of impervious surface had opposing trends in Drosophilidae compared to Tipuloidea. Meanwhile, at the individual species level, there is large variation within phorids in their probability of occurring in close proximity to the ocean. However, some trends were broadly consistent: larger diurnal temperature ranges negatively impacted occurrence in almost all species (Fig. 2d), indicating this is a strong limiting factor in L.A.'s arthropod biodiversity.

All of the heterogeneity in species-level responses scales up to create dramatic spatial variation in arthropod biodiversity in L.A. (Fig. 1). Although some groups have high levels of diversity in the densely urbanized areas of central L.A. (e.g. Drosophilidae), other groups are most diverse in the city's outskirts. In some groups such as spiders, these patterns may be driven by invasive species (a number of which are present in our dataset), who tend to do better in low-diversity systems⁴⁰. We note that predicted species richness for Tipuloidea and Syrphidae maps is consistently low (even though our study includes 34 Syrphidae species and 20 Tipuloidea species). This could potentially be driven by low overall occupancy in species within these groups, or the inability to predict occupancy from the variables included in our models. For example, floral resources and aphids are likely very important in determining Syrphidae occurrences⁴¹, and these variables are not included in our model.

An interesting and somewhat unexpected finding from our analysis was the negative trend we observed for land value in predicting occupancy patterns in Phoridae and Tipuloidea. The luxury effect, a well-known and studied phenomenon of many cities across the globe, suggests that wealthier areas will harbor greater biodiversity as residents of these areas have the means to support lush yards and other green amenities, plus landscaping, maintenance, and irrigation that in turn supports a high diversity of wildlife. Further, wealthy areas in arid cities, such as L.A., have far greater tree canopy cover and vegetation cover⁴², potentially creating habitats that may attract arthropods and act as a refuges in an otherwise inhospitable environment⁴³. Thus, given that arthropods have particular habitat associations with various features of urban environments²³, we expected to find positive patterns of arthropods occupancy in relation to wealth patterns. Our expectations follow those of plants³⁵, birds⁴², and mammals³⁸, which were all more abundant and diverse in high-income areas of cities. This raises the question of what could be driving the contrasting patterns we uncovered for arthropods (and particularly Phoridae) when compared to other taxa. A plausible explanation could be the likely increase in landscaping in high-income areas⁴⁴, which may lead to overly manicured yards with high grass cover and high mowing intensity, each of which is negatively related to arthropod abundance and diversity. This may be particularly important in phorid flies, where some species consume and require decaying matter in order to complete their life cycle¹⁰. For example, one of the phorid species that decreased significantly in occupancy with increasing land value was *Anevrina variabilis*, a species that inhabits ground squirrel burrows and feeds on decaying squirrel carcasses. Over-watering can also negatively affect habitat for phorid flies. For example, another phorid species negatively associated with land value is *Phalacrotophora halictorum*, a parasitoid of halictid bees, a group that is known to inhabit the driest areas with little or no irrigation. Further, phorid flies have been shown to be sensitive to the invasion of Argentine ants (*Linepithema humile*)⁴⁶, as Argentine ants kill off the native ants that parasitic phorid species use as hosts. Argentine ants have been shown to be sensitive to desiccation, and require more water availability⁴⁷, whereas optimal conditions for many phorids are fairly dry natural areas with lots of leaf litter. Alternatively, lower-income areas may have more 'feral' landscaping conditions, which could lead to distinct niches that arthropods could benefit from⁴⁸. Finally, some phorid species, such as *Megaselia scalaris* (another species whose occupancy decreased with high land values) are highly-associated with human activity, and may be more likely to occur in areas of high human density and thus low land value. Whatever the driver behind the patterns we observed for arthropods, our results suggest an extension of the luxury-effect hypothesis in that it may not apply to all arthropod groups in an urban setting. Uncovering the drivers behind this pattern could be an important avenue of future research.

Another unexpected finding was that essentially all phorid species were more likely to occur in areas further away from natural areas (Fig. 2). Much of the natural area in L.A. is concentrated in the surrounding mountains and thus tends to be at higher elevations, and remnant natural areas are very different from city gardens and parks⁴⁵. However, this effect remained significantly negative even after accounting for the independent effect of elevation. It is possible that highly local factors are more important in shaping phorid diversity, who

are (presumably) less vagile than the other groups, and thus are responding on a finer spatial scale. It is also possible that there are some fundamental differences between natural areas at different elevations in L.A. that are not captured in this analysis. For example, our definition of natural areas was quite broad and included any protected areas larger than 10 hectares. This may include heavily-modified and disturbed urban parks that may not be suitable for phorid species, but act as refuges for other arthropod species to maintain relatively high levels of diversity in urbanized landscapes (Fig. 1). Further characterization of these differences in natural area types may help unravel this relationship.

We see mixed responses both within and between groups in response to impervious surface. This was unexpected, as much research has pointed towards a negative effect of impervious surface on arthropod communities^{49–51}. Often this is due to the reduction of floral resources, nesting habitat, and prey availability in these environments³¹. However, the impact of these reductions are most detrimental in specialist species who have specific life history requirements⁵², whereas generalist and highly-mobile species have the flexibility to exploit alternate resources and are less likely to be negatively impacted. Because we surveyed a broad variety of species that likely span a wide generalism-specialism spectrum, this could account for the mixed responses we see. Additionally, it is possible that impervious surface never reaches high enough levels to actually be a strongly negative effect. Recent work has hypothesized that the effects of impervious surface area don't begin to have noticeably negative impacts on insect biodiversity until they reach more than 50% of the surrounding area³³. Mean impervious surface area within a 0.25km radius from our study sites hovers near 55%, and so perhaps is only a significant deterrent in specialist species, and not broadly affecting most species.

Hover flies (Syrphidae) were one of the only groups to exhibit a negative response to temperature. One possible explanation could be that the food plants that maintain high aphid populations (the main food source of some syrphids³³) dry up rapidly in the summer, particularly in areas with high mean temperature. Another possible explanation might be tied to their role as pollinators occupying a niche similar to native bees. Urban warming generally harms the abundance of native bees^{33,54}, and honey bees forage more in higher temperatures and when it is sunny⁵⁵. Considering the ubiquity of feral and managed non-native honey bees throughout our study area, we think it is plausible that the negative effect of temperature on hover fly densities is a consequence of honey bee competition. This is compatible with recent results showing a negative effect of urban honey bees on native pollinators⁵⁶.

There are a number of caveats with our study that may influence the generalization of our findings. First, the taxonomic scope of our study is determined by the limitations imposed by our sampling method. Malaise traps were used to collect arthropods (with the exception of spiders), which trap some groups reliably, whereas some groups of arthropods are underrepresented compared to other sampling approaches^{57,58}. We included all of the taxa which we had enough representation across sites and sampling months to draw reliable inferences from but had to drop others (such as Lepidoptera, Hymenoptera, etc.; see Supp. Fig. S1). Second, we collated land value data from a number of different sources and counties. Methodologies for calculating land value potentially differ across counties, and we see a noticeable jump in values at county lines (such as San Bernardino) that may be an artifact of this (Supp. Fig. S12). Third, we chose to focus solely on spatial patterns of biodiversity, and there is no temporal dimension to this study, neither seasonal nor yearly. For the vast majority of our sites, we only have 1 year of data. Any large-scale climatic anomalies could influence the composition of the arthropod communities we sampled. Arthropod communities also tend to be very stochastic and change year-to-year⁵⁹, and we may be missing this variation. Indeed, seasonal variation in climate has been shown to be important in this system⁶⁰, and future work will aim to decouple the spatial vs. temporal components of biodiversity in this system. Finally, Los Angeles is a unique city in some respects, particularly due to its climate (dry, hot, near-desert²³), varied topography, and proximity to both the coast and to large protected areas. Additionally, it is distinctive in its combination of dense human population but sprawling urban footprint⁶¹ and the high levels of wealth disparity among its residents. However, there are many cities in the US arid west that have emulated the model of urban sprawl that Los Angeles pioneered. Therefore, we believe our results may be representative of other urban areas in western North America, though it is quite possible that we may see different trends in cities that are cooler, wetter, and have a different urbanization footprint (e.g. London or Paris).

In summary, we found that high heterogeneity in responses within and across 6 arthropod groups to (i) topographical, (ii) abiotic, and (iii) anthropogenic factors that can affect occupancy across L.A. This translates into high regional diversity and highly variable composition across space. This high variability in composition across space can have large consequences on the ecosystem services that arthropods provide across the city. Our findings are of general interest with respect to urban arthropod distributions and spatial factors that drive heterogeneity in community composition. These findings also highlight that arthropod diversity is not limited to affluent neighborhoods, and that with proper land management (such as prioritizing green areas, reducing pesticide application, and creating microhabitats in our built environments⁶²), landowners can ensure that the ecosystem benefits derived from arthropods are distributed amongst all urban residents⁶³. Finally, this study illustrates the unique value of community science in urban ecosystems. Not only has this project engaged Angelenos in their own backyard biodiversity and contributed to broader educational initiatives within the community⁶⁴, it has yielded numerous novel scientific findings such as the description of species previously unknown to science^{65,66}, seasonal trends in abundance¹⁰, and has documented major range extensions, potentially due to the introduction of non-native species⁶⁷, all of which are relevant to site- and species-based management strategies. As such, museum-led community science initiatives like this one have an important role to play in the conservation arena⁶⁸.

Methods

Community data collection

BioSCAN survey

BioSCAN is a community science project conducted by the Natural History Museum of Los Angeles County. The BioSCAN team recruited project participants to become BioSCAN site hosts and agreed to host arthropod sampling and environmental data collection on their properties over the course of one continuous year. All sampling took place in residential backyards and gardens in L.A. (Fig. 1). None of the sites were placed in natural areas.

Insects were sampled using a Malaise trap, a tent-like passive collection device that collects insects entering the space of the tent. Insects that are attracted to light and that move upwards when encountering obstacles will move up the mesh and fall into a sampling bottle at the top of the trap that contains ethanol to kill and preserve the specimens. Participants conducted sampling for 1 week each month for one continuous calendar year. Three distinct 1-year sampling campaigns were conducted, accumulating a total of 60 sites that were sampled for a minimum of 1 year. Each sampling campaign (phase) occurred in different years, where phase 1 occurred during 2014, phase 2 during 2015 and 2016, and phase 3 during 2017 and 2018. Staff of the museum visited each location periodically to collect accumulated insect samples and the paired environmental data (see below). For further details, see McGlynn et al.⁶⁰.

Spiders were hand collected by the museum staff via generalized hand-collecting on the property and the general vicinity where the traps were placed.

Arthropod identification

BioSCAN staff and cooperating expert taxonomists identified arthropod specimens using morphology. The following groups were primarily identified to species or morphospecies level and as such are included in the analyses: Mycetophilidae, Phoridae, Drosophilidae, Tipuloidea, Syrphidae, and Araneae (Supp. Fig. S1).

Environmental data and site characteristics

Arthropods are small-bodied ectotherms with specialized diets, so their occurrence and sampling of arthropods is a function of microhabitat characteristics. For this reason, fine-scale measurement of habitat characteristics is essential to create a predictive framework for arthropod biodiversity⁶⁹. We used microclimate variables collected directly at the site of arthropod collections. This approach resulted in insights into the assembly of arthropod communities that would not have been possible with remote sensing or other more coarse approaches to microhabitat assessment^{23,60}.

HOBO data loggers

Environmental data trackers were placed alongside every trap. These trackers recorded data every five minutes on air temperature, soil temperature, relative humidity, and insolation (solar radiation).

We cleaned the HOBO collected data following McGlynn et al.⁶⁰. Specifically, we used the following steps: (i) We aggregated all site-level data into three dataframes, each corresponding to a phase of the BioSCAN data collection, removing duplicate entries and standardizing variable units across phases. For example, temperature was collected in Celsius in some phases, while in others in Fahrenheit. (ii) We filtered the variables to air temperature and relative humidity and excluded all other variables. (iii) We removed potential errors in the data. For temperature, we removed values less than -5°C and more than 50°C , and for relative humidity, we removed values less than 0% and greater than 100%. (iv) Site 07 had a sequence of months containing temperatures higher than 60°C , which is an unreasonably high temperature, so we removed all of this temperature data. (v) We calculated the minimum and maximum temperature (T_{min} and T_{max}) at the hourly scale to smooth any unreasonable temperature values captured at the 5-min period. (vi) For any hourly or daily data, we treated these data as NAN (not a number) if more than 10% of 5-min data used to aggregate was itself NAN. (vii) We then further removed observations from all sites if more than 25% of sites had NAN for that date-time combination for the variable in question. (viii) For the remaining NAN, we used inverse-distance weighting (IDW) interpolation with three nearest neighbors and a power of 1, where the neighbors are other BioSCAN HOBO sensors. IDW was performed at an hourly scale. (ix) Our annual summary statistics were then computed using this normalized, semi-gap-filled data.

Using the site-level micro-meteorological data from the HOBO data loggers, we calculated several bioclimatic variables at an annual scale⁷⁰ to obtain environmental metrics for each site that encompassed temporal variation. This way, landscape and site environmental characteristics were comparable (i.e., accounting spatial variation) and occupancy estimates could be interpolated across L.A. From the available suite of bioclimatic variables we used only mean annual temperature (BIO1) and mean diurnal range (BIO2) in our occupancy models.

Meteorological data

We downloaded Bioclimatic variables from WorldClim⁷⁰ to create the interpolated map of species richness for L.A. (see Spatial Interpolation).

Land value

Land parcel data, including parcel delineation, land value, use type, and year assessed, was available and downloaded for Los Angeles and San Bernardino counties^{71,72}. We filtered for residential parcels only and scaled parcel values to the estimated current value using historic Consumer Price Index estimates. Parcel values were then normalized by area and unreasonably low values were filtered from the dataset.

Using this standardized parcel dataset, we calculated neighborhood-level land value estimates for each of the BioSCAN sites by taking the mean land value of residential parcels within two buffer sizes around each sampling site: 0.25 km and 0.5 km. We chose to calculate these buffers at different scales because of the differing dispersal abilities between the taxonomic groups included in our study. While some groups (e.g. Syrphids) can disperse relatively far⁷³, and will likely be influenced by the habitat within a radius of 0.5–1 km⁷³, other groups (e.g. Phoridae) are known to not disperse very far at all⁷⁴ (and sometimes less than 30 m⁷⁵), and so habitat beyond 0.25 km is likely not relevant for these groups. Land value buffers were set at 0.25 km for Phoridae, Araneae; and 0.5 km for Syrphidae, Tipuloidea, Drosophilidae, and Mycetophilidae.

Impervious surface

We calculated neighborhood-level impervious surface percentage estimates in a similar way to land value. We downloaded 2019 National Land Cover Database (NLCD) Impervious Surface cover data⁷⁶ at the 30 m resolution for L.A. and the mean impervious cover percentage value was calculated within two buffer sizes around each sampling site: 0.25 km, and 0.5 km. The same buffer sizes as from land value were assigned to each taxonomic group.

Elevation

We downloaded and extracted elevation data at the 30 m resolution using the void-filled NASADEM product⁷⁷.

Distance to ocean

Sites in our study vary in their proximity to the Pacific Ocean and topological features. Proximity to the ocean mediates many of the meteorological effects that differentiate coastal and inland areas, such as solar radiation, weather systems, and the distribution of moisture. Additionally, meteorological conditions do not change linearly with geographic distance; they are mediated by topological features, such as mountains (e.g. rain shadows). As such, we developed a cost distance to ocean map for L.A. We created a cost distance raster using the cost distance tool in ArcGIS⁷⁸ where a polygon of the Pacific Ocean was the source layer and slope was the cost (derived from the NASADEM elevation data). Here, slope acts as a surrogate for topological features that disrupt meteorological conditions, namely mountains.

Distance to natural areas

Some parks and protected areas contain native habitat and complex vegetation structures not commonly found across urbanized L.A., potentially resulting in different community structures in neighboring areas. Additionally, natural areas are thought to be important reservoirs of source populations for insects⁷⁹, and so sites closer to natural areas may experience more frequent colonization events. We used the `sf` package⁸⁰ in R⁸¹ to calculate a raster with cell values corresponding to the distance to nearest protected area. We used the California Protected Area Database⁸² as a surrogate for natural areas, removing protected areas less than 10 hectares in size to remove small, manicured parks that may not provide the habitat characteristics that larger, more natural areas do.

Causal analysis

To estimate the effect sizes of different variables of interest (mean annual air temperature, mean diurnal range, land value, percentage impervious cover, distance to natural areas, and distance to ocean) on occupancy, we use a structural causal modeling framework⁸³. For each group studied, we developed a directed acyclic graph (DAG) based on shared life history characteristics of the group and probable influential variables related to our research question (Supp. Figs. S3–S5). Visualizations of each DAG were produced using the R packages `daggity`⁸⁴ and `ggplot2`⁸⁵. The relationships we established in the DAGs were dependent on the climatic and social relationships we find in the greater L.A. metro area. For example, in L.A., we observe that higher elevation areas or areas closer to the ocean have higher living desirability.

After developing a DAG for each taxonomic group studied, we followed the structural causal modelling framework described by Arif and MacNeil 2022⁸³ to validate our DAGs with observed variable data and apply backdoor criterion in order to select the appropriate controlling variables for each regression model. The R package `daggity`⁸⁴ was used once more to test DAG and data consistency among observed variables that could not be explained by the relationships specified. After correcting the DAGs for unexplained correlations, we applied the backdoor criterion to select a single model for each variable of interest to estimate the variable's effect on occupancy.

Occupancy models

We used an occupancy modeling approach to evaluate the effect of the different predictors for each arthropod group. We ran a multi-species occupancy model for each group (Mycetophilidae, Phoridae, Drosophilidae, Tipuloidea, Syrphidae, Araneae). Following Stewart et al.²⁵, we predicted occupancy probability using an information-theoretic approach using all observed variables but used a structural causal modeling framework approach to estimate and compare the effects of each observed variable on arthropod occupancy using separate models informed by our DAGs. Because of this, we ran a total of 37 occupancy models (6 arthropod groups × 6 environmental predictors of interest, plus an additional environmental predictor for Phoridae only). Observed variables include site averaged annual mean temperature, site averaged mean diurnal range, site averaged relative humidity, site elevation, site distance to ocean, site distance to natural areas, neighbourhood level impervious surface, and neighbourhood level land value. We ran the following models.

Model all: is the occupancy model that we developed to predict arthropod occupancy across L.A.

$$\begin{aligned}
 \text{logit}(\psi_{ij}) = & \psi_0 + \\
 & \psi_{\text{Species}[i]} + \\
 & \psi_{T_{\text{Mean}}[i]} \times T_{\text{Mean}}[j] + \\
 & \psi_{T_{\text{DiurnalRange}}[i]} \times T_{\text{DiurnalRange}}[j] + \\
 & \psi_{\text{Elevation}[i]} \times \text{Elevation}[j] + \\
 & \psi_{D_{\text{Ocean}}[i]} \times D_{\text{Ocean}}[j] + \\
 & \psi_{D_{\text{NaturalAreas}}[i]} \times D_{\text{NaturalAreas}}[j] + \\
 & \psi_{\text{ImperviousSurface}[i]} \times \text{ImperviousSurface}[j] + \\
 & \psi_{\text{LandValue}[i]} \times \text{LandValue}[j].
 \end{aligned} \tag{1}$$

We estimated occupancy for each species i in each site j , where ψ_0 and, $\psi_{\text{Species}[i]}$ are defined as a global intercept and a species-specific intercept, respectively. $\psi_{T_{\text{Mean}}[i]}$, $\psi_{T_{\text{DiurnalRange}}[i]}$, $\psi_{\text{Elevation}[i]}$, $\psi_{D_{\text{Ocean}}[i]}$, $\psi_{D_{\text{NaturalAreas}}[i]}$, $\psi_{\text{ImperviousSurface}[i]}$, and $\psi_{\text{LandValue}[i]}$ are species-specific linear effects of mean annual temperature, mean annual diurnal range, elevation, site distance to the ocean, site distance to natural areas, neighborhood impervious surface, and neighborhood land value, respectively.

Model distance to ocean is the occupancy model that we used to estimate the effect of distance to the ocean on occupancy.

$$\begin{aligned}
 \text{logit}(\psi_{ij}) = & \psi_0 + \\
 & \psi_{\text{Species}[i]} + \\
 & \psi_{D_{\text{Ocean}}[i]} \times D_{\text{Ocean}}[j] + .
 \end{aligned} \tag{2}$$

Model distance to natural areas is the occupancy model that we used to estimate the effect of distance to natural areas on occupancy.

$$\begin{aligned}
 \text{logit}(\psi_{ij}) = & \psi_0 + \\
 & \psi_{\text{Species}[i]} + \\
 & \psi_{D_{\text{NaturalAreas}}[i]} \times D_{\text{NaturalAreas}}[j] + \\
 & \psi_{\text{Elevation}[i]} \times \text{Elevation}[j].
 \end{aligned} \tag{3}$$

Model mean annual temperature is the occupancy model that we used to estimate the effect of mean annual temperature on occupancy.

$$\begin{aligned}
 \text{logit}(\psi_{ij}) = & \psi_0 + \\
 & \psi_{\text{Species}[i]} + \\
 & \psi_{T_{\text{Mean}}[i]} \times T_{\text{Mean}}[j] + .
 \end{aligned} \tag{4}$$

Model Diurnal range is the occupancy model that we use to estimate the effect of the Diurnal range on occupancy.

$$\begin{aligned}
 \text{logit}(\psi_{ij}) = & \psi_0 + \\
 & \psi_{\text{Species}[i]} + \\
 & \psi_{T_{\text{DiurnalRange}}[i]} \times T_{\text{DiurnalRange}}[j] + \\
 & \psi_{D_{\text{Ocean}}[i]} \times D_{\text{Ocean}}[j].
 \end{aligned} \tag{5}$$

Model relative humidity is the occupancy model that we use to estimate the effect of the Relative humidity on occupancy. This model was run on Phoridae only.

$$\begin{aligned}
 \text{logit}(\psi_{ij}) = & \psi_0 + \\
 & \psi_{\text{Species}[i]} + \\
 & \psi_{T_{\text{Mean}}[i]} \times T_{\text{Mean}}[j] + \\
 & \psi_{RH_{\text{Mean}}[i]} \times RH_{\text{Mean}}[j] + \\
 & \psi_{\text{ImperviousSurface}[i]} \times \text{ImperviousSurface}[j] + \\
 & \psi_{D_{\text{Ocean}}[i]} \times D_{\text{Ocean}}[j] + \\
 & \psi_{\text{Elevation}[i]} \times \text{Elevation}[j].
 \end{aligned} \tag{6}$$

Model impervious surface and land value is the occupancy model that we used to estimate the effect of impervious surface and land value on occupancy.

$$\begin{aligned} \text{logit}(\psi_{ij}) = & \psi_0 + \\ & \psi_{\text{Species}[i]} + \\ & \psi_{\text{Species}[i]} + \\ & \psi_{\text{Elevation}[i]} \times \text{Elevation}[j] + \\ & \psi_{\text{DOcean}[i]} \times \text{DOcean}[j] + \\ & \psi_{\text{DNaturalAreas}[i]} \times \text{DNaturalAreas}[j] + \\ & \psi_{\text{ImperviousSurface}[i]} \times \text{ImperviousSurface}[j] + \\ & \psi_{\text{LandValue}[i]} \times \text{LandValue}[j]. \end{aligned} \quad (7)$$

For all models, we assume that species-specific intercepts and slopes in both of the above models are normally distributed about some mean. Specifically,

$$\begin{aligned} \psi_{\text{Species}[i]} & \sim \mathcal{N}(0, \sigma_{\psi_{\text{Species}}}) \\ \psi_{\text{TMean}[i]} & \sim \mathcal{N}(\mu_{\psi_{\text{TMean}}}, \sigma_{\psi_{\text{TMean}}}) \\ \psi_{\text{TDiurnalRange}[i]} & \sim \mathcal{N}(\mu_{\psi_{\text{TDiurnalRange}}}, \sigma_{\psi_{\text{TDiurnalRange}}}) \\ \psi_{\text{RHM}[i]} & \sim \mathcal{N}(\mu_{\psi_{\text{RHM}}}, \sigma_{\psi_{\text{RHM}}}) \\ \psi_{\text{Elevation}[i]} & \sim \mathcal{N}(\mu_{\psi_{\text{Elevation}}}, \sigma_{\psi_{\text{Elevation}}}) \\ \psi_{\text{DOcean}[i]} & \sim \mathcal{N}(\mu_{\psi_{\text{DOcean}}}, \sigma_{\psi_{\text{DOcean}}}) \\ \psi_{\text{DNaturalAreas}[i]} & \sim \mathcal{N}(\mu_{\psi_{\text{DNaturalAreas}}}, \sigma_{\psi_{\text{DNaturalAreas}}}) \\ \psi_{\text{ImperviousSurface}[i]} & \sim \mathcal{N}(\mu_{\psi_{\text{ImperviousSurface}}}, \sigma_{\psi_{\text{ImperviousSurface}}}) \\ \psi_{\text{LandValue}[i]} & \sim \mathcal{N}(\mu_{\psi_{\text{LandValue}}}, \sigma_{\psi_{\text{LandValue}}}), \end{aligned} \quad (8)$$

where $\mu_{\psi_{\text{TMean}}}$, $\mu_{\psi_{\text{TDiurnalRange}}}$, $\mu_{\psi_{\text{Elevation}}}$, $\mu_{\psi_{\text{DOcean}}}$, $\mu_{\psi_{\text{DNaturalAreas}}}$, $\mu_{\psi_{\text{ImperviousSurface}}}$, $\mu_{\psi_{\text{LandValue}}}$ denote the mean effect of each corresponding predictor, across species, and σ terms denote the variances about these means.

In both of the above models, we model detection probability as

$$\begin{aligned} \text{logit}(p_{ijk}) = & p_0 + \\ & p_{\text{Month}^2} \times \text{Month}^2[i] + \\ & p_{\text{Species}[i]}, \end{aligned} \quad (9)$$

where p_0 denotes the mean detection probability and $p_{\text{Species}[i]}$ denotes a species-specific random effect. While p_{Month^2} are fixed effects that change as a function of the month. The fixed effect of month assumes detection is highest during the middle of the year and decreases towards the beginning and end of the year.

Additionally, we assume

$$p_{\text{Species}[i]} \sim \mathcal{N}(0, \sigma_{p_{\text{Species}}}). \quad (10)$$

We ran occupancy models in JAGS⁸⁶ and assessed model convergence both by visually inspecting chains and checking The Gelman–Rubin statistic (we ensured that \hat{R} was < 1.1 for all parameters). We used flat, uninformative priors for all parameters and ran models for 100,000 iterations, with a burn-in of 1000 iterations and thinning every 100 iterations, across 3 chains, resulting in a total of 300 posterior samples.

Spatial interpolation of results

We estimated species richness for our study groups across L.A. by spatially interpolating our linear occupancy model, fit with species-specific mean posterior effect sizes for each predictor and intercept. Input spatial data was consistent with the data used for occupancy models with the exception of the two bioclimatic variables: mean annual air temperature (BIO1) and mean diurnal range (BIO2). The HOBO temperature data is extremely fine-scale in its temporal and spatial resolution and so will be the most accurate data in predicting species-specific occupancy at each site. However, the HOBO temperature data is very limited in spatial extent (we only had data loggers at the 60 sites) so interpolating site-level temperature values across the remainder of the LA Basin would be unfeasible. For this reason, we used WorldClim⁷⁰ data to interpolate site-level climate across the basin.

Data processing to produce these maps included the following. (i) We aligned and projected all spatial predictors to the WGS 84 projection system, masked and snapped to a common configuration, and assigned the same cell size. (ii) We scaled and log-transformed input data, as was done to prepare the data for the occupancy model. We also re-centered mean annual air temperature and mean diurnal range estimates due to varying sampling-site mean estimates (as a result of the different data sources). Here, we assume that the BioCLIM variables scale linearly with site-level difference observations. (iii) We filtered values of each predictor if it was beyond three standard deviations about the mean of our model inputs (zero after scaling); values beyond this were set to Not-a-number (NaN). (iv) For each 1-km grid cell across L.A., if any predictor had a value of NaN, we set to NaN to normalize the extent, thus generating the extent seen in Fig. 1.

After processing the spatial data, we calculated species-specific occupancy posterior probabilities for each 1-km cell using the *Model All*. We then summed the median posterior occupancy for each species at each cell for all species, first according to group, then collectively for all groups, to obtain richness estimates (Fig. 1).

For most analyses we used R V4.2.1⁸¹. For spatial manipulations we used the packages `sf`⁸⁰ and `raster`⁸⁷; for data manipulation and visualization we used `tidyverse`⁸⁸; for running models, we used `rjags`⁸⁹, `R2jags`⁹⁰, and `runjags`⁹¹. For manipulation of the environmental data we used Python V3.8.9⁹², and for some spatial analyses we used ArcGIS Pro V3.0⁷⁸. To access and download the spatial data used, we used Google Earth Engine⁹³. To generate maps for Fig. 1, Figs. S2, S12–S17, we used the packages `ggplot2`⁸⁵ and `ggmap`⁹⁴ for visualization, and `basemaps`⁹⁵ for downloading the basemap, which used the toner-lite map tiles by Stamen Design, under CC BY 4.0. Data by OpenStreetMap, under ODbL.

Code availability

All code used to produce these results is freely available in a public GitHub repository (https://github.com/EcolD ataSciUSC/BioSCAN_2023.git). The repository includes most external data—as well as the steps to access and download it—and BioSCAN collection data can be obtained by contacting the corresponding author.

Received: 10 August 2023; Accepted: 22 December 2023

Published online: 03 January 2024

References

- Foley, J. A. *et al.* Global consequences of land use. *Science* **309**, 570–574 (2005).
- Dobbs, C., Escobedo, F. J. & Zipperer, W. C. A framework for developing urban forest ecosystem services and goods indicators. *Landscape Urban Plan.* **99**, 196–206 (2011).
- Escobedo, F. J., Kroeger, T. & Wagner, J. E. Urban forests and pollution mitigation: Analyzing ecosystem services and disservices. *Environ. Pollut.* **159**, 2078–2087 (2011).
- Collins, J. P. *et al.* A new urban ecology: Modeling human communities as integral parts of ecosystems poses special problems for the development and testing of ecological theory. *Am. Sci.* **88**, 416–425 (2000).
- Grimm, N. B. *et al.* Global change and the ecology of cities. *Science* **319**, 756–760 (2008).
- McPhearson, T. *et al.* Advancing urban ecology toward a science of cities. *Bioscience* **66**, 198–212 (2016).
- Rebele, F. Urban ecology and special features of urban ecosystems. *Glob. Ecol. Biogeogr. Lett.* **4**, 173–187 (1994).
- Wu, J. Urban ecology and sustainability: The state-of-the-science and future directions. *Landscape Urban Plan.* **125**, 209–221 (2014).
- McIntyre, N. E. Ecology of urban arthropods: A review and a call to action. *Ann. Entomol. Soc. Am.* **93**, 825–835 (2000).
- Brown, B. V. & Hartop, E. A. Big data from tiny flies: Patterns revealed from over 42,000 phorid flies (Insecta: Diptera: Phoridae) collected over one year in Los Angeles, California, USA. *Urban Ecosyst.* **20**, 521–534 (2017).
- Hall, D. M. *et al.* The city as a refuge for insect pollinators. *Conserv. Biol.* **31**, 24–29 (2017).
- Baldock, K. C. *et al.* A systems approach reveals urban pollinator hotspots and conservation opportunities. *Nat. Ecol. Evol.* **3**, 363–373 (2019).
- Silva, V. H. *et al.* Diverse urban pollinators and where to find them. *Biol. Conserv.* **281**, 110036 (2023).
- Narango, D. L., Tallamy, D. W. & Marra, P. P. Nonnative plants reduce population growth of an insectivorous bird. *Proc. Natl. Acad. Sci.* **115**, 11549–11554 (2018).
- Felson, A. J. & Ellison, A. M. Designing (for) urban food webs. *Front. Ecol. Evol.* **9**, 582041 (2021).
- Baldock, K. C. *et al.* Where is the UK's pollinator biodiversity? The importance of urban areas for flower-visiting insects. *Proc. R. Soc. B* **282**, 20142849 (2015).
- Braaker, S., Obrist, M. K., Ghazoul, J. & Moretti, M. Habitat connectivity and local conditions shape taxonomic and functional diversity of arthropods on green roofs. *J. Anim. Ecol.* **86**, 521–531 (2017).
- Gathof, A. K., Grossmann, A. J., Herrmann, J. & Buchholz, S. Who can pass the urban filter? A multi-taxon approach to disentangle pollinator trait-environmental relationships. *Oecologia* **199**, 165–179 (2022).
- Ossola, A., Hahs, A. K., Nash, M. A. & Livesley, S. J. Habitat complexity enhances comminution and decomposition processes in urban ecosystems. *Ecosystems* **19**, 927–941 (2016).
- Newbound, M., McCarthy, M. A. & Lebel, T. Fungi and the urban environment: A review. *Landscape Urban Plan.* **96**, 138–145 (2010).
- Tresch, S. *et al.* Litter decomposition driven by soil fauna, plant diversity and soil management in urban gardens. *Sci. Total Environ.* **658**, 1614–1629 (2019).
- Forister, M. L., Pelton, E. M. & Black, S. H. Declines in insect abundance and diversity: We know enough to act now. *Conserv. Sci. Pract.* **1**, e80 (2019).
- Adams, B. J. *et al.* Local-and landscape-scale variables shape insect diversity in an urban biodiversity hot spot. *Ecol. Appl.* **30**, e02089 (2020).
- Fenoglio, M. S., Rossetti, M. R. & Videla, M. Negative effects of urbanization on terrestrial arthropod communities: A meta-analysis. *Glob. Ecol. Biogeogr.* **29**, 1412–1429 (2020).
- Stewart, P. S., Stephens, P. A., Hill, R. A., Whittingham, M. J. & Dawson, W. Model selection in occupancy models: Inference versus prediction. *Ecology* **104**, e3942 (2023).
- Arif, S. & MacNeil, M. A. Predictive models aren't for causal inference. *Ecol. Lett.* **25**, 1741–1745 (2022).
- Cinelli, C., Forney, A. & Pearl, J. A crash course in good and bad controls. *Sociol. Methods Res.* <https://doi.org/10.1177/00491241221099552> (2022).
- Laubach, Z. M., Murray, E. J., Hoke, K. L., Safran, R. J. & Perng, W. A biologist's guide to model selection and causal inference. *Proc. R. Soc. B* **288**, 20202815 (2021).
- Heino, J., Alahuhta, J., Fattorini, S. & Schmera, D. Predicting beta diversity of terrestrial and aquatic beetles using ecogeographical variables: Insights from the replacement and richness difference components. *J. Biogeogr.* **46**, 304–315 (2019).
- Theodorou, P. *et al.* Urban areas as hotspots for bees and pollination but not a panacea for all insects. *Nat. Commun.* **11**, 576 (2020).
- Fenoglio, M. S., Calviño, A., González, E., Salvo, A. & Videla, M. Urbanisation drivers and underlying mechanisms of terrestrial insect diversity loss in cities. *Ecol. Entomol.* **46**, 757–771 (2021).
- Corcos, D. *et al.* Impact of urbanization on predator and parasitoid insects at multiple spatial scales. *PLoS ONE* **14**, e0214068 (2019).
- Wenzel, A., Grass, I., Belavadi, V. V. & Tschardt, T. How urbanization is driving pollinator diversity and pollination—A systematic review. *Biol. Conserv.* **241**, 108321 (2020).
- Wood, E. M. *et al.* Historical racial redlining and contemporary patterns of income inequality negatively affect birds, their habitat, and people in Los Angeles, California. *Ornithol. Appl.* **126**, 044 (2023).
- Clarke, L. W., Jenerette, G. D. & Davila, A. The luxury of vegetation and the legacy of tree biodiversity in Los Angeles, California. *Landscape Urban Plan.* **116**, 48–59 (2013).
- Chamberlain, D. E., Henry, D. A., Reynolds, C., Caprio, E. & Amar, A. The relationship between wealth and biodiversity: A test of the luxury effect on bird species richness in the developing world. *Glob. Change Biol.* **25**, 3045–3055 (2019).

37. Li, H., Parker, K. A. & Kalcounis-Rueppell, M. C. The luxury effect beyond cities: Bats respond to socioeconomic variation across landscapes. *BMC Ecol.* **19**, 1–13 (2019).
38. Magle, S. B. *et al.* Wealth and urbanization shape medium and large terrestrial mammal communities. *Glob. Change Biol.* **27**, 5446–5459 (2021).
39. Kaiser, A. & Resasco, J. The impact of the urban matrix on arthropod biodiversity and ecosystem services in community gardens. <https://doi.org/10.21203/rs.3.rs-2949625/v1> (2023).
40. Kempf, J. K., Adams, B. J. & Brown, B. V. Urban spider diversity in Los Angeles assessed using a community science approach. *Urban Nat.* **40**, 1–10 (2021).
41. van Rijn, P. C., Kooijman, J. & Wäckers, F. L. The contribution of floral resources and honeydew to the performance of predatory hoverflies (Diptera: Syrphidae). *Biol. Control* **67**, 32–38 (2013).
42. Wood, E. M. & Esaian, S. The importance of street trees to urban avifauna. *Ecol. Appl.* **30**, e02149 (2020).
43. Vasquez, A. V. & Wood, E. M. Urban parks are a refuge for birds in park-poor areas. *Front. Ecol. Evol.* **10**, 1048 (2022).
44. Jenerette, G. D. *et al.* Urban vegetation and income segregation in drylands: A synthesis of seven metropolitan regions in the southwestern United States. *Environ. Res. Lett.* **8**, 044001 (2013).
45. Avolio, M. *et al.* Urban plant diversity in Los Angeles, California: Species and functional type turnover in cultivated landscapes. *Plants People Planet* **2**, 144–156 (2020).
46. Human, K. G. & Gordon, D. M. Effects of Argentine ants on invertebrate biodiversity in northern California. *Conserv. Biol.* **11**, 1242–1248 (1997).
47. Holway, D. A., Suarez, A. V. & Case, T. J. Role of abiotic factors in governing susceptibility to invasion: A test with argentine ants. *Ecology* **83**, 1610–1619 (2002).
48. Chalker-Scott, L. Nonnative, noninvasive woody species can enhance urban landscape biodiversity. *Arboric. Urban For.* **41**, 173–186 (2015).
49. Sattler, T., Duelli, P., Obrist, M., Arlettaz, R. & Moretti, M. Response of arthropod species richness and functional groups to urban habitat structure and management. *Landsc. Ecol.* **25**, 941–954 (2010).
50. Herrmann, J., Buchholz, S. & Theodorou, P. The degree of urbanisation reduces wild bee and butterfly diversity and alters the patterns of flower-visitation in urban dry grasslands. *Sci. Rep.* **13**, 2702 (2023).
51. Lagucki, E., Burdine, J. D. & McCluney, K. E. Urbanization alters communities of flying arthropods in parks and gardens of a medium-sized city. *PeerJ* **5**, e3620 (2017).
52. Rocha, E. A. & Fellowes, M. D. Urbanisation alters ecological interactions: Ant mutualists increase and specialist insect predators decrease on an urban gradient. *Sci. Rep.* **10**, 6406 (2020).
53. Nelson, E. H., Hogg, B. N., Mills, N. J. & Daane, K. M. Syrphid flies suppress lettuce aphids. *Biocontrol* **57**, 819–826 (2012).
54. Hamblin, A. L., Youngsteadt, E. & Frank, S. D. Wild bee abundance declines with urban warming, regardless of floral density. *Urban Ecosyst.* **21**, 419–428 (2018).
55. Clarke, D. & Robert, D. Predictive modelling of honey bee foraging activity using local weather conditions. *Apidologie* **49**, 386–396 (2018).
56. MacInnis, G., Normandin, E. & Ziter, C. D. Decline in wild bee species richness associated with honey bee (*Apis mellifera* L.) abundance in an urban ecosystem. *PeerJ* **11**, e14699 (2023).
57. Campbell, J. W. & Hanula, J. Efficiency of malaise traps and colored pan traps for collecting flower visiting insects from three forested ecosystems. *J. Insect Conserv.* **11**, 399–408 (2007).
58. Uhler, J. *et al.* A comparison of different malaise trap types. *Insect Conserv. Divers.* **15**, 666–672 (2022).
59. Hawkins, B. A. & Holyoak, M. Transcontinental crashes of insect populations? *Am. Nat.* **152**, 480–484 (1998).
60. McGlynn, T. P. *et al.* Temperature accounts for the biodiversity of a hyperdiverse group of insects in urban Los Angeles. *Proc. R. Soc. B* **286**, 20191818 (2019).
61. Eidlin, E. The worst of all worlds: Los Angeles, California, and the emerging reality of dense sprawl. *Transp. Res. Rec.* **1902**, 1–9 (2005).
62. Hunter, M. R. & Hunter, M. D. Designing for conservation of insects in the built environment. *Insect Conserv. Divers.* **1**, 189–196 (2008).
63. Lerman, S. B., Larson, K. L., Narango, D. L., Goddard, M. A. & Marra, P. P. Humanity for habitat: Residential yards as an opportunity for biodiversity conservation. *BioScience* **73**, 671 (2023).
64. Pauly, G. B., Brown, B. V. & Bettison-Varga, L. Fostering community engagement with nature at the natural history museums of Los Angeles County. In *Proc. Conference on Science and Actions For Species Protection* 46 (2020).
65. Hartop, E. A., Brown, B. V. & Disney, R. H. L. Flies from LA, the Sequel: A further twelve new species of *Megaselia* (Diptera: Phoridae) from the BioSCAN project in Los Angeles (California, USA). *Biodivers. Data J.* **4**, e7756 (2016).
66. Hartop, E. A., Brown, B. V. & Disney, R. H. L. Opportunity in our ignorance: Urban biodiversity study reveals 30 new species and one new Nearctic record for *Megaselia* (Diptera: Phoridae) in Los Angeles (California, USA). *Zootaxa* **3941**, 451–484 (2015).
67. Grimaldi, D. *et al.* Strange little flies in the big city: Exotic flower-breeding Drosophilidae (Diptera) in urban Los Angeles. *PLoS ONE* **10**, e0122575 (2015).
68. Ballard, H. L. *et al.* Contributions to conservation outcomes by natural history museum-led citizen science: Examining evidence and next steps. *Biol. Conserv.* **208**, 87–97 (2017).
69. Murdock, C. C., Evans, M. V., McClanahan, T. D., Miazgowiec, K. L. & Tesla, B. Fine-scale variation in microclimate across an urban landscape shapes variation in mosquito population dynamics and the potential of *Aedes albopictus* to transmit arboviral disease. *PLoS Negl. Trop. Dis.* **11**, e0005640 (2017).
70. Fick, S. E. & Hijmans, R. J. Worldclim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
71. County of Los Angeles, Department of Regional Planning. *Parcels* (2023).
72. County of San Bernardino California. *SBCo Parcel Polygons* (2023).
73. Moquet, L., Laurent, E., Bacchetta, R. & Jacquemart, A.-L. Conservation of hoverflies (Diptera, Syrphidae) requires complementary resources at the landscape and local scales. *Insect Conserv. Divers.* **11**, 72–87 (2018).
74. Morrison, L. W., Kawazoe, E. A., Guerra, R. & Gilbert, L. E. Phenology and dispersal in Pseudacteon flies (Diptera: Phoridae), parasitoids of Solenopsis fire ants (Hymenoptera: Formicidae). *Ann. Entomol. Soc. Am.* **92**, 198–207 (1999).
75. Henne, D. C. & Johnson, S. J. Short-term population redistribution of *Pseudacteon tricuspis* (Diptera: Phoridae) from point source releases. *Environ. Entomol.* **40**, 73–81 (2011).
76. Dewitz, J. & Survey, U. G. *National Land Cover Database (NLCD) 2019 Products (ver. 2.0, June 2021)*. Google Earth Engine. <https://doi.org/10.5066/P9KZCM54> (2021).
77. JPL, N. *NASADEM Merged DEM Global 1 Arc Second v001 [Data Set]*. Google Earth Engine. https://doi.org/10.5067/MEASURES/NASADEM/NASADEM_HGT.001 (2020).
78. Environmental Systems Research Institute, Inc. ArcGIS [GIS software].
79. Samways, M. J. *et al.* Solutions for humanity on how to conserve insects. *Biol. Conserv.* **242**, 108427 (2020).
80. Pebesma, E. Simple features for R: Standardized support for spatial vector data. *R J.* **10**, 439–446 (2018).
81. R Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, 2022).
82. CPAD 2016a. <https://www.calands.org> (2023).

83. Arif, S. & MacNeil, M. A. Applying the structural causal model framework for observational causal inference in ecology. *Ecol. Monogr.* **93**, e1554 (2023).
84. Textor, J., van der Zander, B., Gilthorpe, M. S., Liškiewicz, M. & Ellison, G. T. Robust causal inference using directed acyclic graphs: The R package 'dagitty'. *Int. J. Epidemiol.* **45**, 1887–1894 (2016).
85. Wickham, H., Chang, W. & Wickham, M. H. Package 'ggplot2'. *Create Elegant Data Visualisations Using the Grammar of Graphics. Version 2* 1–189 (2016).
86. Plummer, M. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proc. 3rd International Workshop on Distributed Statistical Computing (DSC 2003)* 1–8 (2003).
87. Hijmans, R. J. *et al. raster: Geographic Data Analysis and Modeling* (2023).
88. Wickham, H. *et al.* Welcome to the tidyverse. *J. Open Source Softw.* **4**, 1686 (2019).
89. Plummer, M. *rjags: Bayesian Graphical Models Using MCMC. R Package Version 4-10* (2019).
90. Su, Y.-S. & Yajima, M. *R2jags: Using R to Run 'JAGS'. R Package Version 0.6-1* (2021).
91. Denwood, M. J. runjags: An R package providing interface utilities, model templates, parallel computing methods and additional distributions for MCMC models in JAGS. *J. Stat. Softw.* **71**, 1–25. <https://doi.org/10.18637/jss.v071.i09> (2016).
92. Van Rossum, G. & Drake, F. L. Jr. *Python Tutorial* (Centrum voor Wiskunde en Informatica Amsterdam, 1995).
93. Gorelick, N. *et al.* Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **202**, 18–27 (2017).
94. Kahle, D. & Wickham, H. ggmap: Spatial visualization with ggplot2. *R J.* **5**, 144–161 (2013).
95. Schwalb-Willmann, J. *basemaps: Accessing Spatial Basemaps in R. R Package Version 0.0.5* (2022).

Acknowledgements

The authors thank all of the collaborators in the community who graciously operated a trap in their gardens for an entire year with us. They also thank Estella Hernandez, Regina Wetzter, Dean Pentcheff, Maria Wong, Kelsey Bailey and James Hogue. L.M.G. acknowledges funding from USC Wise and the Wrigley Institute. T.G. acknowledges funding from the Wrigley Institute. A.B. acknowledges funding from the CIBI project via the California Institute for Biodiversity. BioSCAN was funded by the Natural History Museum of Los Angeles County.

Author contributions

B.V.B. and T.P.M. conceived the sampling, B.V.B. coordinated the data collection, and did some of the phorid sample identifications. E.H. and M.W. also did phorid sample identifications. L.G. identified much of the Drosophilidae samples, P.H.K. the Mycetophilidae, and James Hogue the Syrphidae and Tipulomorpha samples. All authors contributed to the conceptualization of the analysis. J.M.M.L. and T.M.B. performed the analyses, with support of Y.Y.J., C.L., V.S., and A.B. L.M.G. supervised the data analysis. J.M.M.L. and T.M.B. drafted the manuscript, and all authors contributed to the writing and editing.

Competing interests

The authors declare no competing interests.


Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-50675-3>.

Correspondence and requests for materials should be addressed to L.M.G.

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.

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024