From Concrete to Canopy: Illuminating Moth Biodiversity in New York City's Urban Jungle

Shira Linsk[1], Anna Thonis[1], & Kristin M. Winchell[1,2]

[1]New York University, Department of Biology
[2]Correspondence: kristin.winchell@nyu.edu

Author contact information: sl8825@nyu.edu, at6689@nyu.edu, kristin.winchell@nyu.edu

Data Availability: All data and associated code are archived on Zenodo with doi: 10.5281/zenodo.14756963

Funding statement: This work was supported College of Arts and Science, New York University, Dean's Undergraduate Research Fund

Conflict of interest disclosure: The authors declare no conflicts of interest.

RRH: NOCTURNAL URBAN MOTHS

Words: 3,989

0 tables, 5 figures

Abstract

Moths (Lepidoptera) are sensitive to anthropogenic threats and serve as valuable bioindicators. Despite the remarkable diversity and abundance of Lepidoptera globally, there is a lack of information on how moth species are impacted by urbanization. Notably, very little is known about moths in the largest city of the United States, New York City, where pervasive urban pollutants, artificial light at night, land cover change, and habitat fragmentation are severe. We examined the effects of urbanization on moth biodiversity in New York City, with a focus on green spaces. We sampled twelve locations across six parks at night and complemented these surveys with citizen science records of moths from *iNaturalist*. Overall, we found a negative relationship between greater urbanization and moth biodiversity, with community similarity correlated with environmental similarity. Our results support greater biodiversity with less light at night, urban development, and more deciduous tree cover and open land. Our structural equation model reveals additional insight: although we detected a strong negative direct effect of developed land, urbanization also impacts diversity via indirect effects of reducing open space and deciduous tree cover. Developed open space alone does not directly affect diversity, but may positively impact diversity through its covariance with vegetation cover. These findings support the importance of mitigating artificial light at night in urban green spaces and maintaining urban vegetation to ensure nocturnal Lepidoptera are able to persist in rapidly urbanizing landscapes.

Keywords: Lepidoptera, moths, urbanization, ALAN, urban greenspace, invertebrates, urban ecology, biodiversity, path analysis

Introduction

Urbanization is linked to declines in biodiversity through biotic homogenization and species extirpation (McKinney *et al.*, 2006; Piano *et al.*, 2020). Although these impacts are seen across the tree of life, the effects of urbanization on invertebrates have been particularly drastic, with steep declines across insect orders (Hallman *et al.,* 2017; Vaz *et al.*, 2023). Studies have demonstrated local extinctions of insect populations and reduced insect diversity with increasing urbanization (Bergerot *et al.*, 2010; Fattorini, 2011). Much of this work has focused on bees and butterflies, revealing decreased diversity and abundance with increased urbanization (Giuliano, 2004; Bergerot *et al.*, 2010; Lizee *et al.*, 2015; Ramírez-Restrepo & MacGregor-Fors, 2016).

Urbanization presents many challenges for insects, including the urban heat island effect, habitat fragmentation, pollution, impervious surfaces, and exotic plants (Fenoglio *et al.*, 2021). Habitat fragmentation and landscape maintenance have been major contributors to pollinator decline, including Lepidoptera (Sanetra *et al.*, 2024). However, alongside these challenges come opportunities in terms of anthropogenic resources

that may be exploited, as seen in common urban species like cockroaches, flies, and carpet beetles (Frankie & Ehler, 1978). Many species may thrive in cities despite the threats of pest control and habitat modification (Frankie & Ehler, 1978). For example, urban plantings may generate habitat and resources that support invertebrate diversity (Sharmin *et al.*, 2024; Fenoglio *et al.*, 2024), including Hymenoptera (Theodorou *et al.*, 2020) and spiders (Argañaraz *et al.*, 2017).

Relatively little research on urban invertebrates has focused on moths, and there is a notable bias toward diurnal Lepidoptera (Fox, 2012; Anderson *et al.*, 2023; Giuliano, 2004). A few studies on moths have found decreasing populations with increased artificial light at night (Belitz *et al.*, 2024; Boyes *et al.*, 2020; Liang *et al.,* 2023). Additionally, anthropogenic environments generate intense selection pressures and filtering processes that may favor the survival of only a few species, such as heat-tolerant moths (Franzen *et al.*, 2020), or produce phenotypic shifts, such as increased body size in urban moths (Merckx *et al.*, 2018). Indeed, studies of industrial melanism in moths represent some of the first examples of rapid evolutionary change in cities (Kettlewell, 1955; Kettlewell, 1956; Hagen, 1999).

Understanding the effects of urbanization on moths is important because of the crucial roles they play in ecosystem services (e.g., pollination, food, and nutrient cycling). Moths are fundamental to ecosystem health and stability and serve as bioindicators (Dar & Jamal, 2021; Liang *et*

al., 2023; New, 2023) since they respond quickly to environmental change (Belitz *et al.*, 2024; Choi, 2008; Gaona *et al.*, 2021). There are approximately 170,000 known species of moths globally (Carter, 1992), and estimates range as high as 500,000 (Kristensen *et al.,* 2007). Yet, we know very little about the species that persist in urban landscapes. For conservation initiatives to be successful in these spaces, we must first take stock of the species that reside there and relate patterns of diversity to urban stressors.

We investigate patterns of moth diversity in New York City, where artificial light at night, pollution, and habitat fragmentation are extreme. We hypothesize that habitat-scale features such as pollution, light at night, and microclimate correlate with moth diversity. In addition, we hypothesize that landscape-scale features influence species diversity and predict developed land use negatively impacts moth biodiversity. By connecting moth biodiversity to specific urban features, our findings contribute to our understanding of the ecological effects of urbanization and can aid the development of urban conservation strategies.

Methods

Statistical analyses were conducted using R studio (Posit, 2023, ver. 2023.12.0.369) and R (v.4.3.2; R Core Team, 2023). For all linear models, we evaluated multicollinearity with the ‘VIF’ function in the package *car* (Fox & Weisberg, 2019), assessed statistical significance using t-tests for

individual coefficients, and evaluated model fit based on Akaike Information Criterion (AIC) and residual deviance.

Site Selection: We identified 12 on-the-ground sampling sites in six green spaces in Brooklyn and Queens, New York, that reflect variation in urbanization (Fig. 1). We visually assessed satellite imagery to identify green spaces, then chose general locations within each based on their position along a continuum of urbanization quantified from remote-sensed landcover (see below). We chose specific sampling sites spanning a gradient of light pollution based on on-the-ground measurement of light (rather than remote-sensed) to capture variation in this variable at an organismally relevant spatial scale.

Field Sampling: We surveyed each location once shortly after dusk (8-10 pm) between March 3-April 29, 2024. We restricted the temporal sampling period to minimize seasonal variation in species detected across sites. We constructed a light trap using a suspended 5 m^2 white sheet directly illuminated with two flood lights (6000 Kelvin, visible light spectrum) and one ultraviolet LED light (385-395 nm) (as in Merckx & Slade, 2014). Moths are easily sampled with light traps, allowing for robust estimates of species richness and abundance (Choi, 2008). Each survey consisted of two people photographing all moths observed landing on or in the immediate vicinity (on illuminated ground within 2m) of the light trap. We identified species from the photographs and counted the total number of individuals per species encountered during each sampling event. In

addition, for each sampling event, we measured the following site-level variables: air pollution (PM 2.5; Temtop M2000C), light at night (lux; EXTECH Light Meter LT300), temperature, humidity, and wind speed (Kestrel 3000).

iNaturalist data: We downloaded 31,793 Research-Grade observations of the Order Lepidoptera using a polygon spatial filter around all five boroughs of New York City. We manually removed butterfly Families (Papilionidae, Nymphalidae, Pieridae, Lycaenidae, Hesperiidae), resulting in 12,956 observations of moths. We imported records into QGIS (QGIS, 2024, v.3.34.3-Prizren) and used the New York City administrative boundary layer (GIS, 2025) to reduce this dataset to 10,740 records. We used the plugin *Density Analysis* (Hamilton, 2024) to visualize observation density using a grid of 1 km hexagonal cells across the region (Fig. 1). We excluded hexagonal zones with undefined landscape types or >70% water cover to help ensure the zones have relevant habitat data for terrestrial species, resulting in 9,394 records. Three additional observations were excluded because they were from our on-the-ground sampling, resulting in a final dataset of 9,391 records.

Landscape analysis: We mapped on-the-ground sampling locations and *iNaturalist* records using QGIS. We extracted zonal statistics of land cover classification (National Land Cover Database “NLCD”; Dewitz, 2023) within a 150-meter buffer around each on-the-ground sampling location and within each hexagonal zone, and calculated the percentage of each land

cover class within each. We summarized land cover classifications (Deciduous, Mixed Forest, Grassland, Pasture/Hay, Developed Open Space, Developed Low-Intensity, Developed Medium-Intensity, Developed High-Intensity) with principal component analysis (PCA) using the function 'prcomp' in base R *stats* (R Core Team, 2023). We visualized results with packages *factoextra (*Kassambara & Mundt, 2020) and *corrplot* (Wei & Simko, 2021).

Statistical analyses: We summarized species diversity with Simpson's Diversity Index (Simpson's D; Simpson, 1949), which considers both species richness and evenness. We calculated Simpson's D (1-D) using species-level identification for each on-the-ground location and each *iNaturalist* hex bin in R using the function 'diversity' in the package *vegan* (Oksanen *et al.*, 2022).

For on-the-ground sampling, we conducted two analyses. We first analyzed the similarity of community composition with respect to environmental similarity across sites using multiple matrix regression. This entailed generating a distance matrix of species composition similarity across sites, as well as for the environment (percentages of all NLCD land cover classes, PM2.5, lux, temperature, humidity, and wind speed). We used the function 'dist' in the R base *stats* for all distance matrix analyses (R Core Team, 2023) and performed matrix regression using the function 'MRM' in R package *ecodist* (Goslee & Urban, 2007). Second, we analyzed diversity using a linear regression of Simpson's Diversity and site-level

habitat features. We did not include remote-sensed land cover in this model because the resolution of these data and our sample size resulted in low statistical power (overparameterization) and multicollinearity when these variables were included. Explanatory variables included temperature, pollution (PM 2.5), and light intensity (lux) measured during on-the-ground sampling. We simplified this model using stepwise model simplification (backwards and forwards) based on AIC using the function 'step' in R base *stats*.

For the *iNaturalist* data, we conducted three analyses. First, we analyzed the similarity of community composition with respect to environmental similarity across sites using multiple matrix regression. As with the on-the-ground MRM analysis, this entailed generating a distance matrix of species composition similarity across hex-bins, and for the environment (percentages of all NLCD land cover classes only). Second, we analyzed diversity using a linear regression of Simpson's Diversity and landscape-level (remote-sensed) habitat features. Explanatory variables in the model included NLCD categories (as a proportion of hex-bin area): deciduous tree cover, developed land cover (high-, medium-, low-intensity), pasture, developed open land, shrub, and grassland. We simplified this model using stepwise model simplification (backwards and forwards) based on AIC using the function 'step' in R base *stats*.

Third, we developed a structural equation model (SEM) using *iNaturalist* observations to examine relationships between species diversity,

vegetation (deciduous forest, grassland, pasture, shrub), developed land (high-, medium-, low-intensity), and developed open land cover. SEMs analyze multiple relationships at the same time while considering latent factors (i.e., unobserved relationships, such as between deciduous land cover and grassland). We fit our SEM using maximum likelihood with the 'sem' function in the *lavaan* R package (v.0.6-19, Rosseel 2012). Model fit was assessed using Chi-squared (χ^2), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Root Mean Square Error of Approximation (RMSEA). To account for shared variance between variables, we specified covariances between vegetation cover and development intensity, development intensity and developed open space, and between several vegetation variables, including between deciduous and grassland, and between grassland and shrub. We used the z-statistic to assess the significance of each path coefficient in the SEM and visualized observed relationships with the function 'semCors' in the R package *semPlot* (Epskamp, 2022*)*.

In addition, we evaluated relationships between diversity and urbanization more generally across all on-the-ground sample sites combined with *iNaturalist* data using a linear model of Simpson's Diversity Index and the first four principal components of our PCA. We retained four principal components in this analysis based on the percentage of variance captured and visual inspection of the scree plot.

Results

Urban environmental variation: Land cover analyses using remote-sensed data verified that our on-the-ground sampling locations spanned a broad range of urbanization (Fig. 2a). Larger parks (Forest Park, Highland Park, Prospect Park) exhibited greater habitat heterogeneity and had more deciduous forest cover. Small parks (Maria Hernandez, Grover Cleveland, Irving Square), by contrast, had little to no deciduous tree cover and were composed almost entirely of developed land. Grover Cleveland Park stood out with only a small percentage of high-intensity developed land as it borders a private cemetery with green space. These differences are clearly visible in satellite imagery: Forest Park is dominated by forest, Prospect Park has less tree cover, and Maria Hernandez is substantially smaller and consists mostly of developed land (Fig. 2c).

Our PCA incorporated land cover for both on-the-ground sites and *iNaturalist* records (12 on-the-ground locations, 1115 *iNaturalist* hexagonal zones). The PCA captured 95.22% of the variance in the first four principal components, with PC1 capturing 44.65% of the variance. Visual inspection of the scree plot suggests that 4 principal components should be retained to describe variation in the dataset. The first and second principal components (PC1 and PC2) represent urbanization, with strong positive loadings for developed high-intensity land cover on PC1 (Fig. 2b) and for medium-intensity developed land cover on PC2. The third component (PC3) was negatively correlated with water and positively correlated with developed open space and deciduous tree cover. The fourth principal component (PC4)

was positively correlated with developed open space and developed low-intensity land cover and negatively associated with deciduous tree cover. Grassland, pasture, and shrub land cover classifications did not load strongly on any of the first four principal components.

Diversity of Lepidoptera: In our on-the-ground surveys, we observed a total of 69 individuals, representing 11 families, 31 genera, and 33 species (Fig. 3a). Observations were identified to species in 100% of samples. From *iNaturalist*, we extracted data from 9391 "Research Grade" observations, representing 955 species, 537 genera, and 50 families (Fig. 3b). The number of species documented on *iNaturalist* from the locations where we sampled on the ground were much higher for Prospect Park (110 species on *iNaturalist* versus 12 in our sampling) and similar for Grover Cleveland (2 on *iNaturalist* versus 3). By contrast, in two parks, we detected more species in our on-the-ground sampling: Forest Park (29 on *iNaturalist* versus 35) and Highland Park (9 on *iNaturalist* versus 19). In the two parks where no species were detected in our on-the-ground sampling (Maria Hernandez and Irving Square), there were also no observations in *iNaturalist*. Notably, we detected individuals from the family Eriocraniidae in our on-the-ground sampling, but this family was not present in the *iNaturalist* dataset. Conversely, *iNaturalist* observations documented 40 additional families not detected in our on-the-ground sampling.

Based on our on-the-ground sampling, large parks (Forest Park, Highland Park, Prospect Park) displayed the highest levels of biodiversity

(1-D: 0.44–0.92, mean±sd: 0.72±0.15). Among them, Forest Park had a mean diversity of 0.80±0.1, Highland Park had a mean diversity of 0.71±0.19, and Prospect Park had a mean diversity of 0.64±0.17. In contrast, the small parks (Maria Hernandez, Grover Cleveland, Irving Square) exhibited lower biodiversity (1-D: 0–0.66, mean±sd: 0.22±0.38). No species were detected at Maria Hernandez and Irving Square in our on-the-ground sampling; since Simpson's Diversity Index is undefined in the absence of observations, we recorded 1-D as 0 for these sites.

Community Composition: In both our on-the-ground sampling and the *iNaturalist* data, sites with similar environments had similar community composition at the species level (multiple-matrix regression; on-the-ground: $F=43.127$, $R^2=0.501$, $p=0.031$; *iNaturalist*: $F=3943.21$, $R^2=0.014$, $p<0.001$). Species diversity is strongly negatively correlated with PC1 and PC2 (representing increased high- and medium-intensity development, respectively) across both datasets combined (PC1 estimate = -0.198±0.041, $t=-4.778$, $p<0.001$; PC2 estimate= -0.446±0.052, $t=-8.637$, $p<0.001$) (Fig. 4). Diversity was also positively correlated with PC3, representing less water and more developed open space and deciduous tree cover (estimate= 0.415±0.073, $t=5.653$, $p<0.00$). Diversity was not significantly correlated with PC4 ($t=0.095$, $p=0.924$).

Environmental Correlations: Our two linear models—one using on-the-ground sampling with field-measured environmental variation and the other using *iNaturalist* sampling with remote-sensed environmental variation—

revealed correlational relationships between environmental features and biodiversity across both spatial scales. Our simplified regression of Simpson's Diversity for on-the-ground sampling retained two variables: air pollution and light at night. Simpson's diversity was not significantly correlated with air pollution ($t=1.280$, $p=0.233$) but was negatively correlated with light at night (estimate=-0.512±0.062 lux, $t=-8.214$, $p<0.001$). In comparison, our simplified linear regression of Simpson's Diversity by land cover using *iNaturalist* and remote-sensed data retained five significant variables: medium-intensity developed land, low-intensity developed land, developed open land, shrubland, and deciduous tree cover. Biodiversity was negatively correlated with medium-intensity developed land cover (estimate=-0.348±0.064, $t=-5.465$, $p<0.001$) and positively correlated with deciduous tree cover (estimate=0.468±0.107, $t=4.364$, $p<0.001$) and developed open space (estimate=0.374±0.119, $t=3.159$, $p=0.002$). There was also a near-significant positive trend between diversity and low-intensity developed space (estimate=0.280±0.153, $t=1.832$, $p=0.068$), and the correlation with shrubland was non-significant ($t=1.457$, $p=0.145$).

Structural Equation Model: Our structural equation model (SEM) exhibited an overall good fit to the data (CFI: 0.955, TLI: 0.915, RMSEA: 0.067). The chi-square statistic was significant ($\chi^2=83.670$, df=19, $p<0.001$), although this is expected with larger sample sizes, and the overall fit indices suggest the model effectively captures the relationships in

the data. Our SEM detected a significant positive effect of vegetation cover on diversity (estimate=1.856±0.394, z=4.715, p<0.001; Fig. 5) and a significant negative effect of developed land on diversity (estimate=-0.301±0.100, z=-3.011, p=0.003). We did not find a significant direct effect of developed open space (z=0.435, p=0.664) on diversity (Fig. 5). Significant negative covariance was observed between vegetation cover and development intensity (covariance=-0.012, z=-10.725, p<0.001), and significant positive covariance between vegetation cover and open space (covariance=0.002, z=4.15, p<0.001). Similarly, we found a significant negative covariance between development intensity and developed open space
(covariance=-0.014, z=-11.702, p<0.001). We also observed weak but significant positive covariance between developed low-intensity and developed open space (covariance=0.005, z=9.396, p<0.001) and grassland and shrub (covariance<0.001, z=11.033, p<0.001). We observed weak but significant negative covariance between developed high-intensity and developed low-intensity (covariance=0.015, z=-11.270, p<0.001), and deciduous and grassland (covariance<-0.001, z=-2.655, p = 0.008). Further inspection of relationships among variables indicates that developed medium-intensity land cover is driving the direct effect of development on diversity, whereas high and low-intensity development impact diversity more indirectly via their relationships with open land and deciduous tree cover. Further, among the vegetation variables, only deciduous tree cover

has a strong direct effect on diversity, whereas grassland, shrubland, and pasture all have indirect effects related to their covariance with developed land.

Discussion

Urbanization drastically alters landscapes, replacing naturally occurring habitats and vegetation with impervious surfaces and buildings. This process not only impacts land cover but also increases artificial light at night (ALAN), which can both have significant consequences for biodiversity. Lepidoptera (moths and butterflies) are highly sensitive to land cover changes (Gaona et al., 2021; Beck et al., 2002), and are also influenced by habitat patch size, quality, and shape (Usher & Keiller, 1998; Lintott et al., 2014). In this study, we examined how urbanization relates to biodiversity of nocturnal Lepidoptera (moths), which perform vital ecosystem services and act as bioindicators (Dar & Jamal, 2021). Our findings suggest that urbanization is associated with lower moth diversity, with specific elements of the urban landscape shaping patterns of biodiversity in complex ways.

We found comparable moth community composition between locations with similar environments in both our on-the-ground and *iNaturalist* datasets, and a general trend of decreasing diversity with increasing urbanization. This result suggests that moth community composition across New York City is strongly influenced by the environmental features that

define different locations. Our linear models provide some insight into those features: we found strong negative correlations between diversity medium-intensity urbanization and strong positive correlations between diversity and deciduous tree cover and open land. In addition, our on-the-ground sampling found a strong negative relationship between moth diversity and light at night, with a high explained variance (R^2), suggesting light at night accounts for a substantial portion of the observed variation in species diversity across sampled locations. Although several studies have found no consistent effects of light at night on biodiversity (White, 2018; Sanetra *et al.*, 2024), others have identified significant impacts on moth community composition (Langevelde *et al.*, 2017; Wilson *et al.*, 2018; Grenis *et al.*, 2023).

Our PCA analysis across both on-the-ground and *iNaturalist* data reinforces these findings: we found strong negative correlations between diversity and principal components primarily defined by high- and medium-intensity development, and strong positive correlations with the principal component representing deciduous tree cover and open space. The identification of deciduous tree cover, open space, and high- and medium-intensity developed land across multiple analyses highlights the importance of these features. Our results are consistent with those of similar studies performed in metropolitan areas where highly disturbed habitats have been shown to correlate with reduced moth species richness (Franzen *et al.*, 2020; Rice, 2015; Belitz *et al.,* 2024).

We used structural equation modeling (SEM) to gain deeper insight into the correlations and interactions among environmental variables shaping biodiversity. SEMs more holistically represent the complex effects of urbanization and can be insightful for informing conservation (Yang *et al.*, 2022; Wang *et al.*, 2019; Santibañez-Andrade *et al.*, 2015). While we did not perform a SEM using our on-the-ground data due to the larger sample sizes required, our SEM using *iNaturalist* data and remote-sensed land cover found, similar to our other analyses, that higher diversity is associated with greater vegetation cover and less developed land. Further inspection of these relationships revealed that although vegetation and development have direct impacts on diversity, multiple indirect effects exist. We found a strong negative covariance between vegetation and development and a strong positive covariance between vegetation and developed open land, suggesting that characteristics of more intense urbanization may drive reduced moth diversity. The many indirect effects detected in our SEM provide additional insight: both high- and medium-intensity development negatively impact moth diversity through their negative impact on tree cover and open space. Together, these results suggest that developed land conversion itself does not necessarily drive species declines but may lead to declines through strong indirect effects. This finding points to potential avenues for moth conservation through the maintenance of tree cover, vegetation, and open space in more developed parts of the city.

Surprisingly, we did not find a relationship between air pollution and moth diversity at our study sites. This finding contradicts our expectations, as researchers have documented declines across various taxa in response to higher levels of air pollution (Lovett *et al.,* 2009; Barker & Tingey, 2012). However, Zvereva and Kozlov (2009) suggest that the negative effects of air pollution on insects may have been overestimated. It is also possible that our findings were influenced by confounding variables such as temperature and wind speed, as higher levels of particulate matter (PM2.5) coincided with warmer temperatures and lower wind speeds during sampling. A longer temporal scale may be required to detect a correlation between air pollution and moth biodiversity, given the high variation and covarying nature of this variable. These factors highlight the complexity of isolating individual environmental drivers and underscore some limitations of our study.

Incorporating *iNaturalist* observations enabled us to expand beyond the spatial and temporal scope of our on-the-ground sampling. This resulted in detecting an additional 40 families not observed in person, most of which were not active during the sampling timeframe. Moreover, *iNaturalist* has many records of larval stage moths, which would not be recorded during a light trap survey, and adult stages of some species are difficult to record via light trapping due to variation in phototaxis (Merckx & Slade, 2014). We also acknowledge the inherent limitations of using *iNaturalist* data, which can be influenced by biases in socioeconomic factors and uneven sampling

efforts. Research suggests that areas with higher socioeconomic status often experience greater sampling efforts by community scientists (Carlen *et al.* 2024). Taxonomic biases also exist; charismatic fauna, such as butterflies, are reported more frequently despite the greater overall diversity of moth species (Shirey *et al.*, 2021). Although our sampling documented high diversity in Forest and Highland Park, these parks have few moth records on *iNaturalist*. In fact, our on-the-ground sampling documented more moth species over a single evening at three locations in each of these parks than have ever been documented on *iNaturalist* for the entire park. Forest Park had 29 records on *iNaturalist* in over 20 years of observations in New York City, versus our 35 observations from our on-the-ground surveys; Highland Park is similar with 9 *iNaturalist* observations versus our 19. Indeed, we detected members of a family (Eriocraniidae) that had never been observed in New York City on *iNaturalist*, perhaps because of their particularly small size (~10mm wingspan) and early spring emergence. This discrepancy between datasets highlights the importance of integrating *iNaturalist* data with on-the-ground field surveys to develop a more comprehensive understanding of biodiversity in a given area, and particularly in cities where habitat is highly heterogeneous and strong sampling biases can arise.

In conclusion, our analyses revealed significant relationships between moth biodiversity and urbanization at both local and landscape spatial scales. Our on-the-ground sampling revealed negative effects of light at

night, while our landscape-level analyses found higher vegetation cover and developed open space support greater moth diversity. Further, our findings highlight the importance of using advanced modeling approaches, such as SEM, alongside linear regression to reveal hidden or unobserved relationships in environmental data. Understanding these relationships is essential for designing targeted conservation strategies, as singular actions—such as increasing vegetation alone—may be insufficient. We hope this study sparks further research into urban moth biodiversity, as these remarkable and essential creatures serve vital functions in our ecosystems.

References

Anderson, M., Rotheray, E. L., & Mathews, F. (2023). Marvellous moths! pollen deposition rate of bramble (Rubus futicosus L. agg.) is greater at night than day. *PLOS ONE*, *18*(3), e0281810. https://doi.org/10.1371/journal.pone.0281810

Argañaraz, C. I., Rubio, G. D., & Gleiser, R. M. (2017). Spider communities in urban green patches and their relation to local and landscape traits. *Biodiversity and Conservation*, *27*(4), 981–1009. https://doi.org/10.1007/s10531-017-1476-8

Barker, J. R., & Tingey, D. T. (2012). *Air Pollution Effects on Biodiversity*. Springer Science & Business Media.

Beck, J., Schulze, C. H., Linsenmair, K. E., & Fielder, K. (2002). From forest to farmland: diversity of geometrid moths along two habitat gradients on Borneo. *Journal of Tropical Ecology*, *18*(1), 33–51. https://doi.org/10.1017/s026646740200202x

Belitz, M. W., Sawyer, A., Hendrick, L. K., Kawahara, A. Y., & Guralnick, R. P. (2024). Substantial urbanization-driven declines of larval and adult moths in a subtropical environment. *Global Change Biology*, *30*(3). https://doi.org/10.1111/gcb.17241

Bergerot, B., Fontaine, B., Julliard, R., & Baguette, M. (2010). Landscape variables impact the structure and composition of butterfly assemblages along an urbanization gradient. *Landscape Ecology*, *26*(1), 83–94. https://doi.org/10.1007/s10980-010-9537-3

Boyes, D. H., Evans, D. M., Fox, R., Parsons, M. S., & Pocock, M. J. O. (2020). Is light pollution driving moth population declines? A review of causal mechanisms across the life cycle. *Insect Conservation and Diversity*, *14*(2). https://doi.org/10.1111/icad.12447

Carlen, E. J., Estien, C. O., Caspi, T., Perkins, D., Goldstein, B. R., Kreling, S., Hentati, Y., Williams, T. D., Stanton, L. A., Des Roches, S., Johnson, R. F., Young, A. N., Cooper, C. B., & Schell, C. J. (2024). A framework for contextualizing social-ecological biases in contributory science data. *People and Nature*, *6*(2). https://doi.org/10.1002/pan3.10592

Carter, D. J. (1992). *Butterflies and Moths*. Dorling Kindersley.

Choi, S.-W. (2007). Diversity and composition of larger moths in three different forest types of Southern Korea. *Ecological Research*, *23*(3), 503–509. https://doi.org/10.1007/s11284-007-0406-8

Dar, A. A., & Jamal, K. (2021). Moths As Ecological Indicators: A REVIEW. Munis Entomology & Zoology. https://www.munisentzool.org/yayin/Vol_16/Issue_2/20210520-BX9WJO4X.pdf

Dewitz, J., 2023, National Land Cover Database (NLCD) 2021 Products: U.S. Geological Survey data release, https://doi.org/10.5066/P9JZ7AO3.

Epskamp S (2022). _semPlot: Path Diagrams and Visual Analysis of Various SEM Packages' Output_. R package version 1.1.6, https://CRAN.R-project.org/package=semPlot

Fattorini, S. (2011). Insect extinction by urbanization: A long term study in Rome. *Biological Conservation*, *144*(1), 370–375. https://doi.org/10.1016/j.biocon.2010.09.014

Fenoglio, M. S., Calviño, A., González, E., Salvo, A., & Videla, M. (2021). Urbanisation drivers and underlying mechanisms of terrestrial insect diversity loss in cities. *Ecological Entomology*, *46*(4), 757–771. https://doi.org/10.1111/een.13041

Fenoglio, M. S., González, E., Calviño, A., & Videla, M. (2024). Urban Insect Communities in the Neotropics: A Systematic Literature Review and a Green Path to Promote Biodiversity Conservation. *Springer EBooks*, 163–188. https://doi.org/10.1007/978-3-031-49255-6_8

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression, Third Ed.* Sage Publications. https://www.john-fox.ca/Companion/

Fox, R. (2012). The decline of moths in Great Britain: a review of possible causes. *Insect Conservation and Diversity*, *6*(1), 5–19. https://doi.org/10.1111/j.1752-4598.2012.00186.x

Frankie, G. W., & Ehler, L. E. (1978). Ecology of Insects in Urban Environments. *Annual Review of Entomology*, *23*(1), 367–387. https://doi.org/10.1146/annurev.en.23.010178.002055

Franzén, M., Betzholtz, P.-E., Pettersson, L. B., & Forsman, A. (2020). Urban moth communities suggest that life in the city favours thermophilic multi-dimensional generalists. *Proceedings of the Royal*

Society B: Biological Sciences, *287*(1928), 20193014. https://doi.org/10.1098/rspb.2019.3014

Gaona, F. P., Iñiguez-Armijos, C., Brehm, G., Fiedler, K., & Espinosa, C. I. (2021). Drastic loss of insects (Lepidoptera: Geometridae) in urban landscapes in a tropical biodiversity hotspot. *Journal of Insect Conservation*, *25*(3), 395–405. https://doi.org/10.1007/s10841-021-00308-9

GIS DCP. (2025). *NYC Borough Boundary* [shapefile]. Arcgis.com. https://hub.arcgis.com/datasets/DCP::nyc-borough-boundary/about

Giuliano, W. M., Accamando, A. K., & Mcadams, E. J. (2004). Lepidoptera-habitat relationships in urban parks. *Urban Ecosystems*, *7*(4), 361–370. https://doi.org/10.1007/s11252-005-6835-7

Golsee, S. C., & Urban, D. (2007). The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, *22*(7). 10.18637/jss.v022.i07

Grenis, K., Nufio, C. R., Wimp, G. M., & Murphy, J. M. (2023). Does artificial light at night alter moth community composition? *Philosophical Transactions of the Royal Society B*, *378*(1892). https://doi.org/10.1098/rstb.2022.0365

Hagen, J. B. (1999). Retelling Experiments: H.B.D. Kettlewell's Studies of Industrial Melanism in Peppered Moths. *Biology & Philosophy*, *14*(1), 39–54. https://doi.org/10.1023/a:1006576518561

Hallmann, C. A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Schwan, H., Stenmans, W., Müller, A., Sumser, H., Hörren, T., Goulson, D., & de Kroon, H. (2017). More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PLOS ONE*, *12*(10), e0185809. https://doi.org/10.1371/journal.pone.0185809

Hamilton, C. (2024). *QGIS Density Analysis Plugin* (Version 2024.1.23) [Computer software]. GitHub. https://github.com/hamiltoncj/qgis-densityanalysis-plugin

Kassambara, A., & Mundt, F. (2020). _factoextra: Extract and Visualize the Results of Multivariate Data Analyses_. R package version 1.0.7,

Kettlewell, H. (1955). Selection experiments on industrial melanism in the Lepidoptera. *Heredity*, *9*(3), 323–342. https://doi.org/10.1038/hdy.1955.36

Kettlewell, H. (1956). Further selection experiments on industrial melanism in the Lepidoptera. *Heredity*, *10*(3), 287–301. https://doi.org/10.1038/hdy.1956.28

Kristensen, N. P., Scoble, M. J., & Karsholt, O. (2007). “Lepidoptera phylogeny and systematics: the state of inventorying moth and butterfly diversity.” Arthropod Systematics & Phylogeny, 65(1), 1-50.

Langevelde, F., Braamburg-Annegarn, M., Huigens, M. E., Groendijk, R., Poitevin, O., Deijk, J. R., Ellis, W. N., Grunsven, R. H. A., Vos, R., Vos, R. A., Franzén, M., & WallisDeVries, M. F. (2018). Declines in moth

populations stress the need for conserving dark nights. *Global Change Biology*, *24*(3), 925–932. https://doi.org/10.1111/gcb.14008

Liang, H., He, Y.-D., Theodorou, P., & Yang, C.-F. (2023). The effects of urbanization on pollinators and pollination: A meta-analysis. *Ecology Letters*, *26*(9). https://doi.org/10.1111/ele.14277

Lintott, P. R., Bunnefeld, N., Fuentes-Montemayor, E., Minderman, J., Blackmore, L. M., Goulson, D., & Park, K. J. (2014). Moth species richness, abundance and diversity in fragmented urban woodlands: implications for conservation and management strategies. *Biodiversity and Conservation*, *23*(11), 2875–2901. https://doi.org/10.1007/s10531-014-0753-z

Lizee, M.-H., Tatoni, T., & Deschamps-Cottin, M. (2015). Nested patterns in urban butterfly species assemblages: respective roles of plot management, park layout and landscape features. *Urban Ecosystems*, *19*(1), 205–224. https://doi.org/10.1007/s11252-015-0501-5

Lovett, G. M., Tear, T. H., Evers, D. C., Findlay, S. E. G., Cosby, B. J., Dunscomb, J. K., Driscoll, C. T., & Weathers, K. C. (2009). Effects of Air Pollution on Ecosystems and Biological Diversity in the Eastern United States. *Annals of the New York Academy of Sciences*, *1162*(1), 99–135. https://doi.org/10.1111/j.1749-6632.2009.04153.x

McKinney, M. L. (2006). Urbanization as a major cause of biotic homogenization. *Biological Conservation*, *127*(3), 247–260. https://doi.org/10.1016/j.biocon.2005.09.005

Merckx, T., Kaiser, A., & Van Dyck, H. (2018). Increased body size along urbanization gradients at both community and intraspecific level in macro-moths. *Global Change Biology*, *24*(8), 3837–3848. https://doi.org/10.1111/gcb.14151

Merckx, T., & Slade, E. M. (2014). Macro-moth families differ in their attraction to light: implications for light-trap monitoring programmes. *Insect Conservation and Diversity*, *7*(5), 453–461. https://doi.org/10.1111/icad.12068

New, T. R. (2023). Moth Ecology and Conservation Importance. *Springer EBooks*, 21–37. https://doi.org/10.1007/978-3-031-32103-0_2

Oksanen, J., Simpson, G., Blanchet, F., Kindt, R., Legendre, P., Minchin, P., O'Hara, R., Solymos, P., Stevens, M., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S., Evangelista, H., FitzJohn, R., Friendly, M., Furneaux, B., Hannigan, G., Hill, M., Lahti, L., McGlinn, D., Ouellette, M., Ribeiro, Cunha, E., Smith, T., Stier, A., Ter, Braak C., Weedon, J. (2022). _vegan: Community Ecology Package_. R package version 2.6-4,

Piano, E., Souffreau, C., Merckx, T., Baardsen, L. F., Backeljau, T., Bonte, D., Brans, K. I., Cours, M., Dahirel, M., Debortoli, N., Decaestecker, E., De Wolf, K., Engelen, J. M. T., Fontaneto, D., Gianuca, A. T., Govaert, L., Hanashiro, F. T. T., Higuti, J., Lens, L., & Martens, K. (2020). Urbanization drives cross-taxon declines in abundance and

diversity at multiple spatial scales. *Global Change Biology*, *26*(3). https://doi.org/10.1111/gcb.14934

Posit team (2023). RStudio: Integrated Development Environment for R (ver. 2023.12.0.369). Posit Software, PBC, Boston, MA. http://www.posit.co/

QGIS Development Team. (2024). QGIS Geographic Information System (ver. 3.34.3-Prizren). QGIS Association. https://www.qgis.org/

Ramírez-Restrepo, L., & MacGregor-Fors, I. (2016). Butterflies in the city: a review of urban diurnal Lepidoptera. *Urban Ecosystems*, *20*(1), 171–182. https://doi.org/10.1007/s11252-016-0579-4

R Core Team (2023). _R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria.

Rice, A. J. (2015). Community Patterns in Urban Moth Assemblages. *Journal of the Lepidopterists' Society*, *69*(3), 149–156. https://doi.org/10.18473/lepi.69i3.a2

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. Journal of Statistical Software, 48(2), 1-36. https://doi.org/10.18637/jss.v048.i02

Sanetra, D., Berger, J., Hartlieb, M., Simons, N. K., Walther, G., Nico Blüthgen, & Staab, M. (2024). Disentangling how urbanisation influences moth diversity in grasslands. *Insect Conservation and Diversity*, *17*(2), 229–242. https://doi.org/10.1111/icad.12713

Santibañez-Andrade, G., Castillo-Argüero, S., Vega-Peña, E. V., Lindig-Cisneros, R., & José Alejandro Zavala-Hurtado. (2015). Structural equation modeling as a tool to develop conservation strategies using environmental indicators: The case of the forests of the Magdalena river basin in Mexico City. *Ecological Indicators*, *54*, 124–136. https://doi.org/10.1016/j.ecolind.2015.02.022

Sharmin, M., Tjoelker, M. G., Esperon-Rodriguez, M., Katlav, A., Gilpin, A.-M., Rymer, P. D., & Power, S. A. (2024). Urban greening with shrubs can supercharge invertebrate abundance and diversity. *Scientific Reports*, *14*(1), 8735. https://doi.org/10.1038/s41598-024-58909-8

Shirey, V., Belitz, M. W., Barve, V., & Guralnick, R. (2021). A complete inventory of North American butterfly occurrence data: narrowing data gaps, but increasing bias. *Ecography*, *44*(4), 537–547. https://doi.org/10.1111/ecog.05396

Simpson, E. H. (1949). Measurement of Diversity. *Nature*, *163*(4148), 688–688. https://doi.org/10.1038/163688a0

Theodorou, P., Radzevičiūtė, R., Lentendu, G., Kahnt, B., Husemann, M., Bleidorn, C., Settele, J., Schweiger, O., Grosse, I., Wubet, T., Murray, T. E., & Paxton, R. J. (2020). Urban areas as hotspots for bees and pollination but not a panacea for all insects. *Nature Communications*, *11*(1), 1–13. https://doi.org/10.1038/s41467-020-14496-6

Usher, M. B., & Keiller, S. W. (1998). The macrolepidoptera of farm woodlands: determinants of diversity and community

structure. *Biodiversity & Conservation*, *7*, 725-748. https://doi.org/10.1023/a:1008836302193

Vaz, S., Manes, S., Khattar, G., Mendes, M., Lima, F., Teles, E., de, E., Gama-Maia, D., Maria Lucia Lorini, Valverde, M., & Paulo Cesar Paiva. (2023). Global meta-analysis of urbanization stressors on insect abundance, richness, and traits. *Science of the Total Environment*, 165967–165967. https://doi.org/10.1016/j.scitotenv.2023.165967

Wang, Y., Kotze, D. J., Vierikko, K., & Niemelä, J. (2019). What makes urban greenspace unique – Relationships between citizens' perceptions on unique urban nature, biodiversity and environmental factors. *Urban Forestry & Urban Greening*, *42*, 1–9. https://doi.org/10.1016/j.ufug.2019.04.005

Wei, T., & Simko, V. (2021). R package 'corrplot': Visualization of a Correlation Matrix (Version 0.92). https://github.com/taiyun/corrplot

White, P. J. T. (2018). An Aerial Approach to Investigating the Relationship between Macromoths and Artificial Nighttime Lights Across an Urban Landscape. *Journal of Agricultural and Urban Entomology*, *34*(1), 1–14. https://doi.org/10.3954/1523-5475-34.1.1

Wilson, J. F., Baker, D., Cheney, J., Cook, M., Ellis, M., Freestone, R., Gardner, D., Geen, G., Hemming, R., Hodgers, D., Howarth, S., Jupp, A., Lowe, N., Orridge, S., Shaw, M., Smith, B., Turner, A., & Young, H. (2018). A role for artificial night-time lighting in long-term changes in populations of 100 widespread macro-moths in UK and Ireland: a

citizen-science study. *Journal of Insect Conservation*, *22*(2), 189–196. https://doi.org/10.1007/s10841-018-0052-1

Yang, L., Zhao, S., & Liu, S. (2022). A global analysis of urbanization effects on amphibian richness: Patterns and drivers. *Global Environmental Change*, *73*, 102476. https://doi.org/10.1016/j.gloenvcha.2022.102476

Figure Captions

Figure 1. (a) On-the-ground sampling locations for 12 locations in Brooklyn and Queens, New York City, USA, where we sampled Lepidoptera, and National Land Cover (NLCD) data for New York City. (b) Species diversity in 1-km hexagonal bins derived from iNaturalist data. Base maps by Open Street Map.

Figure 2. (a) Stacked bar plot of NLCD land cover composition at on-the-ground sampling sites. Colors represent the proportion of each land cover type in 150m buffers around each sampling location. (b) Principal components analysis of NLCD land cover for our on-the-ground sampling locations and all hex-bins used with iNaturalist data. PC1 represents high-intensity developed land cover. (c) Satellite images of sampling locations in Brooklyn and Queens, New York.

Figure 3. (a) Stacked bar plot of Lepidoptera Family diversity per site for on-the-ground sampling. (b) Number of observations of each Lepidoptera Family across all iNaturalist sampling sites. Colored bars are Families observed in iNaturalist and on-the-ground sampling. (c) Images of moths taken during on-the-ground sampling.

Figure 4. Simpson’s diversity (D-1) as it relates to land cover summarized by principal components analysis: (a) PC1 representing high-intensity

developed land, (b) PC2 representing medium-intensity developed land, and (c) PC3 representing open developed land and deciduous tree cover.

Figure 5. Relationships among predictor variables and Simpson's diversity generated by the structural equation model. Nodes represent variables included in the model, while edges indicate associations between them. Thicker lines represent stronger relationships, with positive correlations shown in green and negative correlations in red. Edge weights are proportional to the strength of standardized regression coefficients.

Figures

Figure 1

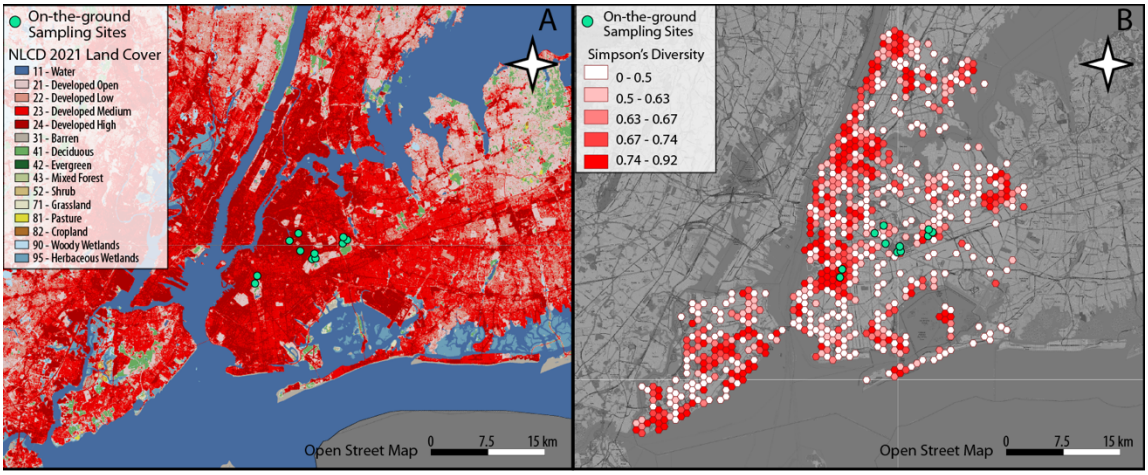


Figure 2

Figure 3

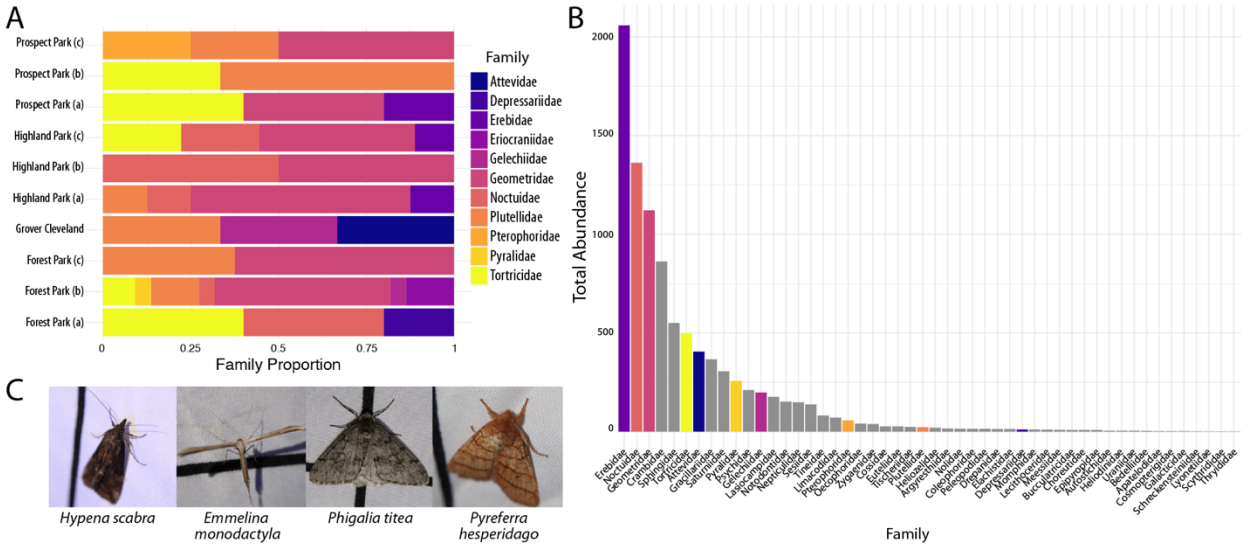


Figure 4

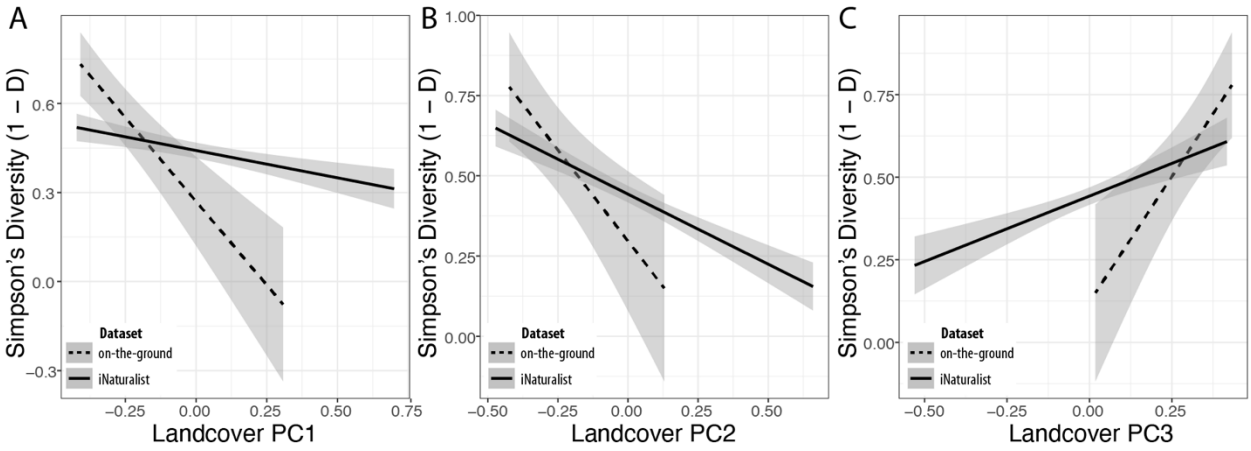
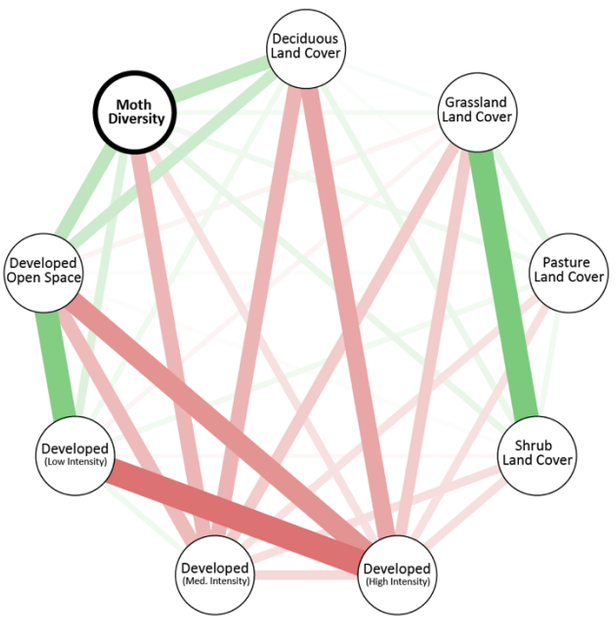
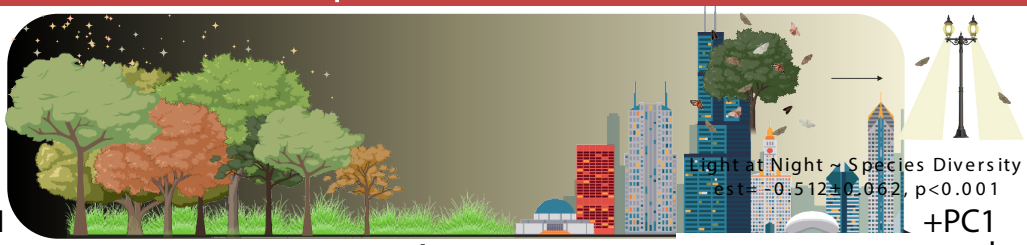


Figure 5


Deciduous Land Cover
Grassland Land Cover
Pasture Land Cover
Shrub Land Cover
Developed (High Intensity)
Developed (Med. Intensity)
Developed (Low Intensity)
Developed Open Space
Moth Diversity

MOTH DIVERSITY ACROSS NYC URBAN MOSAIC

Methods: Sample Moths Across Urban Gradient

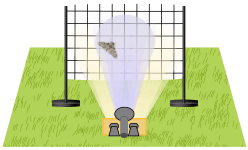


-PC1
less urban

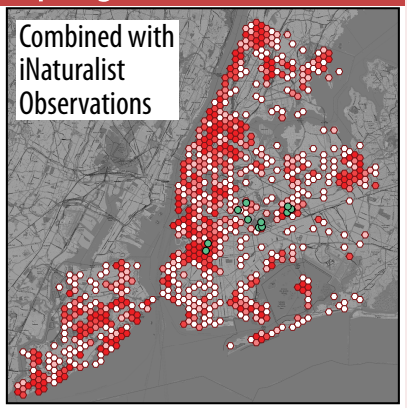
Land Cover PC1

+PC1
more urban

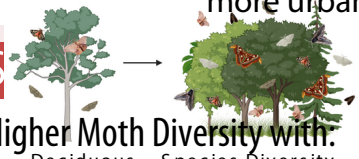
Sampling



Light-trapping on the ground

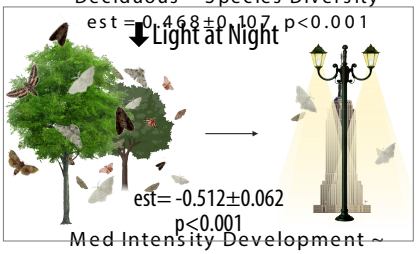


S

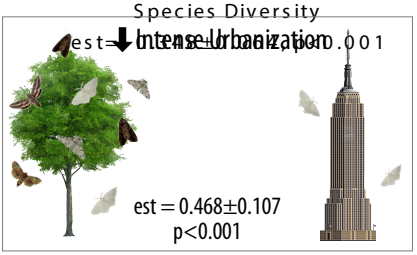


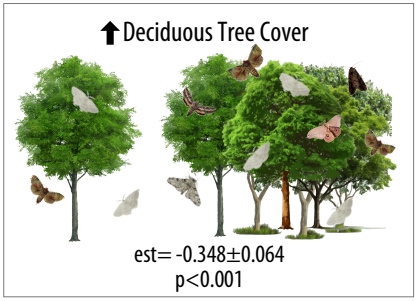
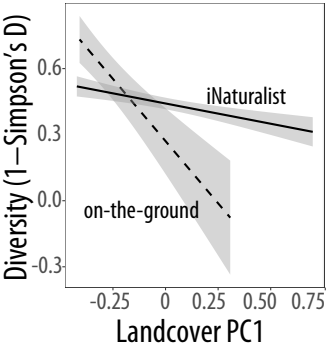
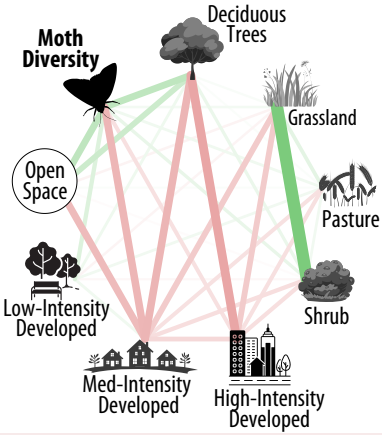
Higher Moth Diversity with:

Deciduous ~ Species Diversity
est = 0.468±0.107, p<0.001



Med Intensity Development ~ Species Diversity
est = 0.468±0.107, p<0.001





Results: Urbanization and Biodiversity