Beyond urbanization metrics: Using graphical causal models to investigate mechanisms in urban ecology and evolution

Jesse B. Borden[1], James P. Gibbs[2], John P. Vanek[1,3], Bradley J. Cosentino[1*]

[1]Department of Biology, Hobart and William Smith Colleges, 300 Pulteney Street, Geneva, NY, 14456, USA

[2]Department of Environmental Biology, The State University of New York College of Environmental Science and Forestry, 1 Forestry Drive, Syracuse, NY 13210, USA

[3]New York Natural Heritage Program, The State University of New York College of Environmental Science and Forestry, 625 Broadway, Albany, NY 12207, USA

[*]Corresponding author: Bradley J. Cosentino (cosentino@hws.edu)

Abstract

As the fastest-growing form of land use, urbanization drives profound environmental change that reshapes biodiversity. Studies in urban ecology and evolution often rely on generic indices of urbanization to characterize biodiversity patterns along urban-rural gradients. However, these indices do not permit explicit tests of causal hypotheses by which urbanization mediates ecological and evolutionary processes. Here, we show how a graphical causal modeling framework with directed acyclic graphs (DAGs) can be used to design clear conceptual models to better evaluate mechanistic hypotheses about the effects of urbanization on biodiversity. We introduce the basic structure of DAGs and illustrate their value, first with simulated data and then with a case study on coat color variation in eastern gray squirrels (*Sciurus carolinensis*) along an urbanization gradient in Syracuse, New York, USA. We show how univariate statistical models with generic urbanization predictors are difficult to interpret and can lead to misleading conclusions about mechanisms in urban ecology and evolution. In contrast, DAGs make causal assumptions transparent and can point to specific processes driving biodiversity patterns. When applied to our case study, analyses informed by a DAG revealed a surprising finding: although squirrel melanism was more prevalent in urban than rural populations, the prevalence of melanism was constrained by components of environmental change common to cities, namely roads, forest loss, and predator activity, in contrast to expectations. Managing biodiversity in an increasingly urbanized world will require a mechanistic understanding of how urbanization impacts biodiversity patterns; graphical causal models can provide a powerful approach to do so.

Key words: causal inference, directed acyclic graphs, eastern gray squirrel, *Sciurus carolinensis*, urban ecology, urban evolution

Introduction

Urban areas represent the fastest growing land use type and have profound effects on biodiversity (Aronson et al. 2014; Ritchie and Roser 2018; Szabó et al. 2023). Urbanization consists of a suite of abiotic and biotic changes driven by high concentrations of people and the physical environment they build (Fig. 1; Ritchie and Roser 2018), the ecological and evolutionary consequences of which are often studied using the urban-rural gradient approach. In this approach components of biodiversity are compared along a landscape-scale sampling area from the urban core to rural habitats (McDonnell and Pickett 1990; McDonnell and Hahs 2008). The urban-rural gradient approach has been used to document associations of population size, species distributions, community composition, and genetic and phenotypic trait variation with urbanization (Winchell et al. 2016; Moll et al. 2019; Cosentino and Gibbs 2022; Santangelo et al. 2022). Although this method has been helpful to clarify spatial patterns of biodiversity associated with urbanization, reliance on univariate metrics to estimate urbanization effects (e.g., impervious cover; Moll et al. 2019) can be problematic, leading to biased conclusions about the processes underlying patterns of urban biodiversity.

Gaining insight into the processes underlying urban patterns in ecology and evolution is complicated by two factors. First, environmental changes associated with urbanization cascade through multiple causal pathways (Alberti et al. 2020). For example, road networks can *directly* affect biodiversity through vehicular collisions and noise pollution (Francis et al. 2012; Bennett

2017; Kok et al. 2023), but roads also have *indirect* effects by contributing to habitat fragmentation and pollution runoff, altering predator-prey dynamics, or facilitating the spread of invasive species, which in turn lead to impacts on biodiversity (Bowman et al. 2010; Bennett 2017; Mumma et al. 2019; Cerqueira et al. 2021). When a univariate model is used to test how a biodiversity component is related to road cover, the estimated effect includes the sum of direct and indirect effects. In some cases, these effects can offset one another, obscuring different causal mechanisms and leading to the conclusion that road cover has no impact. Second, many environmental conditions covary along urbanization gradients – for example, road extent with housing density or light pollution – confounding the interpretation of the effect of any one axis of environmental variation (Moll et al. 2019). Taken together, estimates of urbanization effects with univariate models can include a mix of true causal effects and noncausal confounding, rendering a mechanistic interpretation nearly impossible.

The covariation of many environmental changes along urbanization gradients makes interpretation of statistical models for observational research designs particularly challenging (Fig. 1). Our goal is to explore how graphical causal models can be used to make stronger inferences about hypothesized urbanization effects and inform the design of statistical models to test causal mechanisms of interest (Rohrer 2018; McElreath 2020; Arif and MacNeil 2023; Arif and Massey 2023). Specifically, we highlight directed acyclic graphs (DAG) as one approach that can be useful for specifying causal assumptions about urban systems, including direct effects, indirect effects via mediators, and confounding paths. To do so, we first use a simple DAG with simulated data to show how univariate models of urban effects on biodiversity can have unclear interpretations. We then apply these methods to a case study in urban evolution,

testing hypotheses about urban effects on phenotypic variation in eastern gray squirrels (*Sciurus carolinensis*).

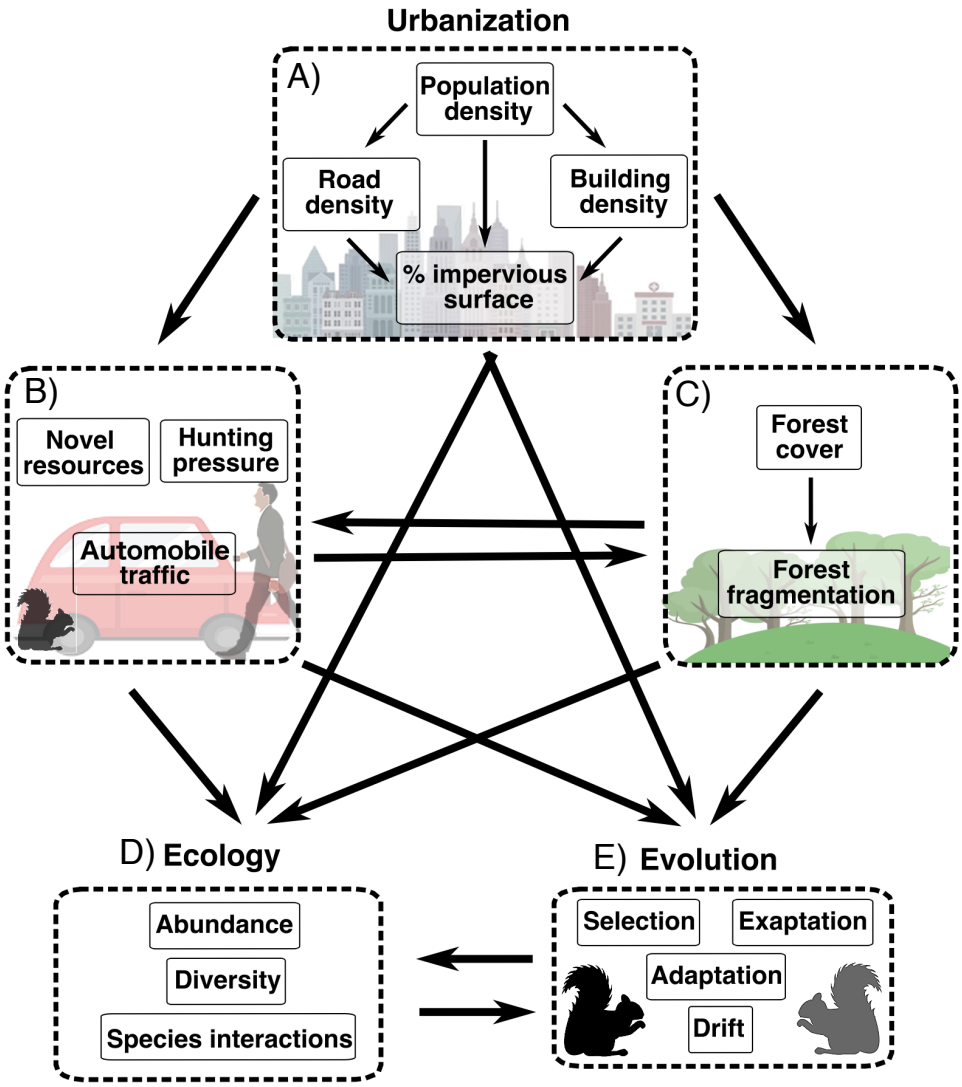

Fig. 1. A conceptual model illustrating a network of potential mechanisms by which urbanization affects phenotypic variation in eastern gray squirrels (*Sciurus carolinensis*), a species that shows urban-rural clines in coat color (gray or melanic) with multiple mechanistic hypotheses. (A) Urbanization consists of a suite of interacting anthropogenic drivers, which can represent distinct mechanisms of biotic change. This suite of changes directly alters (B) human-wildlife interactions, (C) habitat quality and configuration, (D) ecological processes, and (E) evolutionary processes. However, there are also many indirect pathways by which urbanization can impact ecological and evolutionary processes and shape urban-rural clines in coat color (gray or melanic) in eastern gray squirrels.

Directed acyclic graphs for urbanization effects

Directed acyclic graphs are visual diagrams representing hypothesized causal effects in a system that can help guide study design and statistical thinking (Arif and Massey 2023). DAGs consist of variables connected by arrows, with each arrow representing a direct causal effect. For example, consider an urban ecologist interested in the effect of housing density on bird species diversity. Figure 2A shows a simplified, hypothetical DAG related to this question. As shown in the DAG, housing density can affect bird diversity via two pathways. First, the causal path *Housing → Diversity* is a **direct effect,** which could represent the expectation that housing structures provide nesting habitat for some bird species. Second, the pathway *Housing → Greenspace → Diversity* represents an **indirect effect**. In this case, the causal effect of housing density on species diversity is transmitted by green space, which is a **mediator**. Housing density might be expected to decrease bird diversity indirectly by decreasing green space. The **total effect** of housing density on species diversity reflects the sum of these direct and indirect effects.

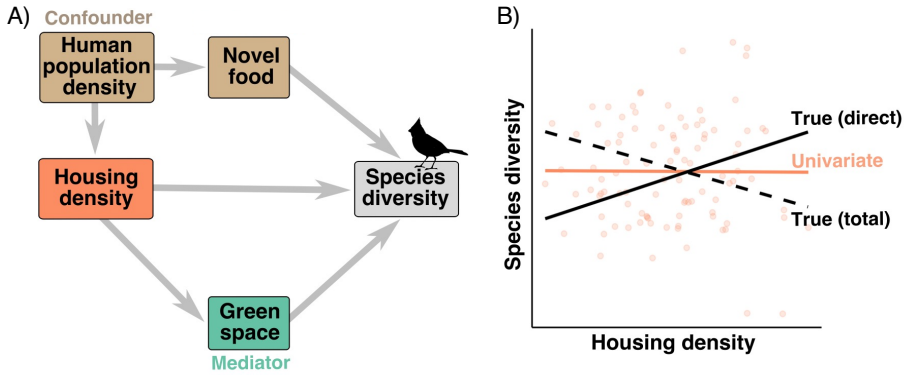

Fig. 1. (A) A hypothetical directed acyclic graph (DAG) illustrating causal relationships in a system where a researcher is interested in the causal effects of housing density on bird species diversity. The DAG identifies confounding and mediating pathways that need to be considered for valid causal inference. (B) Data simulation illustrating the problem of confounding when estimating total and direct effects of housing density on species diversity. Standardized regression coefficients (β) were chosen to

specify the direction and strength of causal effects in the system: 1) human population density increases housing density ($\beta = 0.8$), 2) housing density increases bird diversity ($\beta = 0.3$), 3) housing density indirectly reduces bird diversity due to a negative effect of housing density on greenspace ($\beta = -0.8$) and positive effect of greenspace on bird diversity ($\beta = 0.7$), 4) human density indirectly increases bird diversity due to positive effects of human density on food availability ($\beta = 0.9$) and food availability on bird diversity ($\beta = 0.4$). The simulation shows the true direct and total effects of housing density on species diversity are not recovered by a univariate model of housing density predicting species diversity.

Beyond providing a visualization of causal assumptions and hypotheses, DAGs help identify three elemental causal structures: confounders, mediators, and colliders (McElreath 2020). Identifying these causal structures is essential to inform how statistical models should be parameterized and interpreted to estimate causal effects of interest (McElreath 2020). For example, to correctly estimate the effects of housing density on bird species diversity, it is important to notice the "backdoor path" from housing density to species diversity through human population density and novel food (Fig. 2A; *Housing* ← *Human* → *Food* → *Diversity*). The path *Human* → *Housing* reflects the expectation that human density increases housing density, and the path *Human* → *Food* → *Diversity* could represent a positive indirect effect of people on bird diversity by provisioning of novel food resources (e.g., bird feeders). Here human population density functions as a **confounder** because it affects both housing density (the explanatory variable of interest) and species diversity (the response variable of interest). A simple univariate regression coefficient of housing density predicting species diversity would include a mix of the direct causal effect of housing density, the indirect effect through green space, and the noncausal associations with human population density and novel food, leading to a biased estimate that recovers none of the true effects of housing density as illustrated via data simulation (Fig. 2B).

Correct estimation of the effects of housing density on bird diversity requires specific parameterization of statistical models based on causal assumptions made explicit by the DAG.

For example, estimating the direct effect of housing density on species diversity requires adjusting for human population density and green space, such as by including these variables as covariates in a regression model. But care must be taken to decide which covariates to include in a model to estimate the effect of interest. For example, if the total effect of housing density was of interest, green space should *not* be adjusted, as it functions as a mediator of housing density and therefore contributes to the total effect. A correct estimate of the total effect of housing density on species diversity requires adjusting for human population density or novel food, but not green space. Determining appropriate adjustment covariates for any causal effect of interest from a DAG can be done with the help of software such as *DAGitty* (Textor et al. 2016).

While DAGs resemble structural equation models (SEM), unlike SEMs, DAGs do not allow for bidirectional relationships or cycles (i.e., they are acyclic), and they do not imply a specific form of the causal relationships between variables (e.g., positive-negative, linear-nonlinear, additive-interactive). A critical assumption when using a DAG is that it specifies the correct causal structure of the system. In practice, it is important to consider whether any non-trivial causal effect may be missing, but especially unobserved or unmeasured confounders of the causal relationships of interest. Thus, while DAGs are very useful for clearly stating causal assumptions and identifying appropriate adjustment covariates to estimate the causal effects of interest, it is important that researchers apply prior knowledge of their system to inform the hypothesized structure of the DAG and include important confounders, even when unobserved. There are approaches to address unobserved confounding effects, and we refer the reader to Byrnes and Dee (2025) for a detailed treatment. For a more complete overview of DAGs, including the implications of additional causal structures such as colliders, see Rohrer (2018).

Case study: coat color variation in eastern gray squirrels

Study species and hypothesized urbanization effects

Eastern gray squirrels (*Sciurus carolinensis*) provide an excellent study species to show how DAGs can be used to help disentangle mechanisms by which urbanization affects ecological and evolutionary outcomes. Gray squirrels are common across urban and rural landscapes in their native range of eastern North America. Individuals tend to have one of two coat colors – gray or black (melanic) — inherited in a simple Mendelian fashion (McRobie et al. 2009). The prevalence of melanism tends to be greater in cities than adjacent rural forests (Gibbs et al. 2019; Cosentino and Gibbs 2022; Cosentino et al. 2023). Multiple hypotheses have been proposed about the ecological and evolutionary processes maintaining these urban-rural clines (Cosentino and Gibbs 2022; Cosentino et al. 2023, Fig. 1). For example, the melanic morph is visually more conspicuous than the gray morph, such that predation may select for the more cryptic gray morph (Proctor et al. 2025). Thus, both predator activity and habitat context (e.g., sources of refuge, including trees and buildings) could alter the prevalence of melanism. Additionally, road mortality has been hypothesized to maintain clines in melanism, as melanic morphs are more visible to humans and underrepresented among roadkill compared to gray morphs (Gibbs et al. 2019; Proctor et al. 2025; Parlin et al. in press).

Generating an initial DAG

We created a DAG to represent our causal assumptions for how predator activity, components of habitat structure that affect refuge availability (building cover, forest cover, and forest fragmentation), and road cover affect spatial variation in the prevalence of melanism along an

urbanization gradient (Fig. 3A). We expected melanism would be greatest in areas of refugia (high building cover, high forest cover, low forest fragmentation) and high road cover, and lowest in areas of high predator activity. Critically, while these variables are all hypothesized to affect the prevalence of melanism directly (Tab. 1), they can also covary along urban-rural gradients and confound one another, highlighting the necessity of using a DAG to identify potential confounds and appropriate adjustments to estimate the effects of interest. For example, forest cover and forest fragmentation may directly affect the prevalence of melanism by altering crypsis or availability of refuge, but such habitat alteration could also indirectly affect the prevalence of melanism by changing predator activity (Bateman et al. 2012).

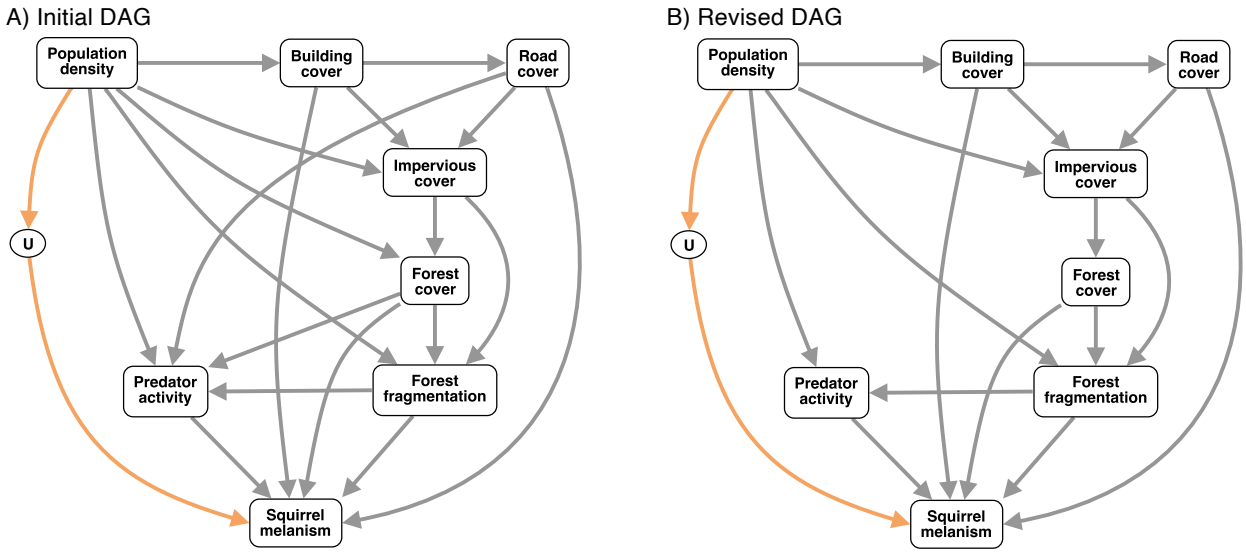

Fig. 3 Directed acyclic graphs (DAG) representing our scientific model, showing hypothesized causal relationships between variables that could affect variation in melanism in gray squirrels along urbanization gradients. An initial DAG was generated representing our original hypotheses (A), and the DAG was revised by dropping direct effects among explanatory variables that were not empirically supported (B) (see Results). Arrows indicate hypothesized direct effects. Squares represent measured variables, and circles represent unobserved variables. "U" represents the unobserved effects of people on the prevalence of squirrel melanism, including translocations and hunting pressure, which are known to occur but remain unmeasured. Arrow color distinguishes pathways that were estimable with measured variables (gray) from pathways that were not estimable due to unobserved processes (orange).

The foundation of our DAG was human population density, operating on the assumption that environmental change along urbanization gradients is ultimately driven by people. We included a direct effect of human population density on building cover (Glover and Simon 1975). Because roads are the main way to access buildings, we included a direct effect of building cover on road cover. Buildings and roads were assumed to be direct causes of impervious cover, and we included a direct effect of human population density on impervious cover to represent other types of impervious surface that people build (e.g., sidewalks, parking lots). We included a direct effect of impervious cover on forest cover, as the creation of infrastructure requires tree removal (Liu et al. 2014). Forest fragmentation has a known negative relationship with forest cover (Fahrig 2003, 2017), so we included a direct effect of forest cover on fragmentation. We also included a direct effect of impervious cover on forest fragmentation, as landscape elements like buildings and roads cause discontinuities in forest cover, regardless of their effects on the absolute amount of forest (Liu et al. 2014; Liu et al. 2019). Human population density was included as a direct effect of forest cover and fragmentation to account for other ways that people remove and fragment forested areas (e.g., gardening and open parks). We included direct effects of road cover, forest cover, and fragmentation on mammalian predator activity, as all three have been linked to urban carnivore distribution and behavior (Ordeñana et al. 2010; Bateman et al. 2012).

Although DAGs are useful for identifying and adjusting for confounders, unobserved or unmeasured confounders still create the potential for bias (Byrnes and Dee 2025). For example, the high prevalence of melanism among squirrels in cities may be due in part to historical introductions by people (i.e., propagule pressure) or reduced hunting pressure in cities (Cosentino et al. 2023), variables that were not possible to measure. Including unobserved

228 variables like these in a DAG can be important to make gaps in knowledge explicit and identify
229 confounding with other measured variables. As such, we added an unobserved variable to our
230 DAG to represent unmeasured mechanisms by which people could mediate coat color change
231 ("U"). We included a direct effect of human population density on U and a direct effect of U on
232 the prevalence of melanism (Fig. 3A). By including this unobserved variable in the DAG, we
233 were able to identify how backdoor paths including U as a confounder could be closed by
234 adjusting for human population density (Fig. 3A).
235
236 *Study area and sampling*
237 To parameterize the DAG, we used data on coat color variation in gray squirrels across an urban-
238 rural gradient in Syracuse, New York, USA (43.0469° N, -76.1444° W). Syracuse has an
239 estimated human population of 146,000 (in 2022) and occurs in a humid continental climate
240 (Koppen 1936) with four distinct seasons, including hot summers and cold winters. Tree cover
241 ranges from 4.5% to 47% across the city, averaging 27% (Nowak et al. 2016). The surrounding
242 rural landscape is a mosaic of suburban areas, small villages, agriculture, and forest
243 (Supplemental Fig. S1).
244 We used camera traps and point count surveys to estimate prevalence of gray squirrel
245 color morphs at 49 sites across the urban-rural gradient of Syracuse between 16 September 2021
246 and 13 October 2022 (Supplemental Fig. S1; Cosentino et al. 2023). Because squirrel
247 populations are limited by tree cover, sites were selected in urban greenspaces and rural
248 woodlots with mature deciduous trees (Parker and Nilon 2008; Merrick et al. 2016). Site
249 locations were chosen to span the urban-rural gradient ranging 1.3–23.7 km from the city center
250 (median = 9.3 km) and separated by at least 1 km. At each site we deployed a single camera trap

251 (Browning Strikeforce Pro XD) between September 2021 and March 2022, fastening the camera
252 to a tree 30-50 cm off the ground. Cameras were active for 46-379 days (median = 263 days),
253 with daily detection histories generated for each color morph by a single observer. We also
254 conducted visual point count surveys at each site between October 2021 and October 2022.
255 During each survey, a single observer recorded the number of squirrels seen of each color morph
256 during a 3-min period. Sites were surveyed 2-7 times each (median = 5 surveys).
257
258 *Quantifying measured variables*
259 All measured variables in the DAG other than predator activity were measured via remotely
260 sensed datasets with the *raster* and *exactextractr* packages in R (Tab. 1; Baston et al. 2022;
261 Hijmans et al. 2023; R Core Team 2024). We used a 300-m buffer to summarize variables
262 around each site to encompass a typical home range size of gray squirrels (<5 ha) (Koprowski et
263 al. 2016). Mammalian predator activity was indexed using our camera trap imagery and
264 calculated as the percentage of camera trap days mammalian predators were detected at each site
265 (Bu et al. 2016). Predators detected by camera traps included coyote (*Canis latrans*), gray fox
266 (*Urocyon cinereoargenteus)*, red fox (*Vulpes vulpes*), fisher (*Pekania pennanti*), domestic dog
267 (*Canis familiaris*) and domestic cat (*Felis catus*).
268
269 *Statistical analyses*
270 We tested hypotheses from our DAG (Fig. 3A) in two stages. First, we used Bayesian
271 generalized linear models to estimate direct effects among measured variables other than squirrel
272 melanism, selecting appropriate distributions for response variables and adjustment covariates
273 based on the DAG using the *DAGitty* R package (Textor et al. 2016). For example, to estimate

the direct effect of building cover on impervious cover, we assumed impervious cover follows a beta distribution and modeled it as a function of building cover, along with road cover as an adjustment. Models were fit with STAN via the R package *brms* (Bürkner 2017) with noninformative priors and standardized predictors. We generated a revised DAG (Fig. 3B) that dropped direct effects when posterior distributions broadly overlapped 0 (Supplemental Tab. S1). Next, we estimated the direct and total effects of each variable on squirrel melanism using an integrated hierarchical model combining point count and camera trap data (Kéry and Royle 2020; Cosentino et al. 2023). This model included a process submodel to estimate morph-specific abundances and proportion melanic at each site, and a detection submodel accounting for temperature effects on activity. We estimated direct and total effects of measured variables in our DAG on proportion melanic with linear models, adjusting for DAG-informed covariates as needed. We also fit univariate models with each measured variable as a predictor of melanism to compare inferences with direct and total effects. Hierarchical models were fit in JAGS via the *jagsUI* R package (Plummer 2017, Kellner and Meredith 2024) using mostly noninformative priors, but also with informative priors for detection and abundance intercepts based on prior knowledge. All models were run with four chains and 4000 retained iterations after enough warm-up or burn-in to reach convergence, which was confirmed visually with traceplots and R-hat < 1.01 (Gelman and Hill 2007). Posterior distributions were summarized with means and 95% credible intervals. Details on statistical models are provided in Supplemental Text S1.

Results and Discussion

Our case study's goal was to evaluate and refine the set of hypothesized factors that affect coat color variation in gray squirrels along an urbanization gradient. Aligning with previous studies showing a strong cline along the urbanization gradient (Cosentino and Gibbs 2022; Cosentino et al. 2023), we re-confirmed prevalence of squirrel melanism was greatest in the city, as inferred by a positive total effect of human population density on proportion melanic (Fig. 4, Tab. 2). However, our DAG-informed analyses provided multiple new insights into the potential mechanisms maintaining the urban-rural cline. Remarkably, the total effects of building cover, road cover, and impervious cover (all metrics of human impact commonly used to represent "urbanization effects") on the prevalence of melanic squirrels were each *negative* despite squirrel melanism being most prevalent in the city (Fig. 4, Tab. 2). These negative total effects suggest the distribution of the melanic morph is limited by various aspects of physical infrastructure, even within the city where melanic squirrels are generally most prevalent. Considering the positive total effect of human population density combined with negative total effects of physical infrastructure on squirrel melanism, our results suggest there remain unmeasured cause(s) represented by U in our DAG (Fig. 5) that play an important role in maintaining the urban-rural cline by favoring melanics in the city.

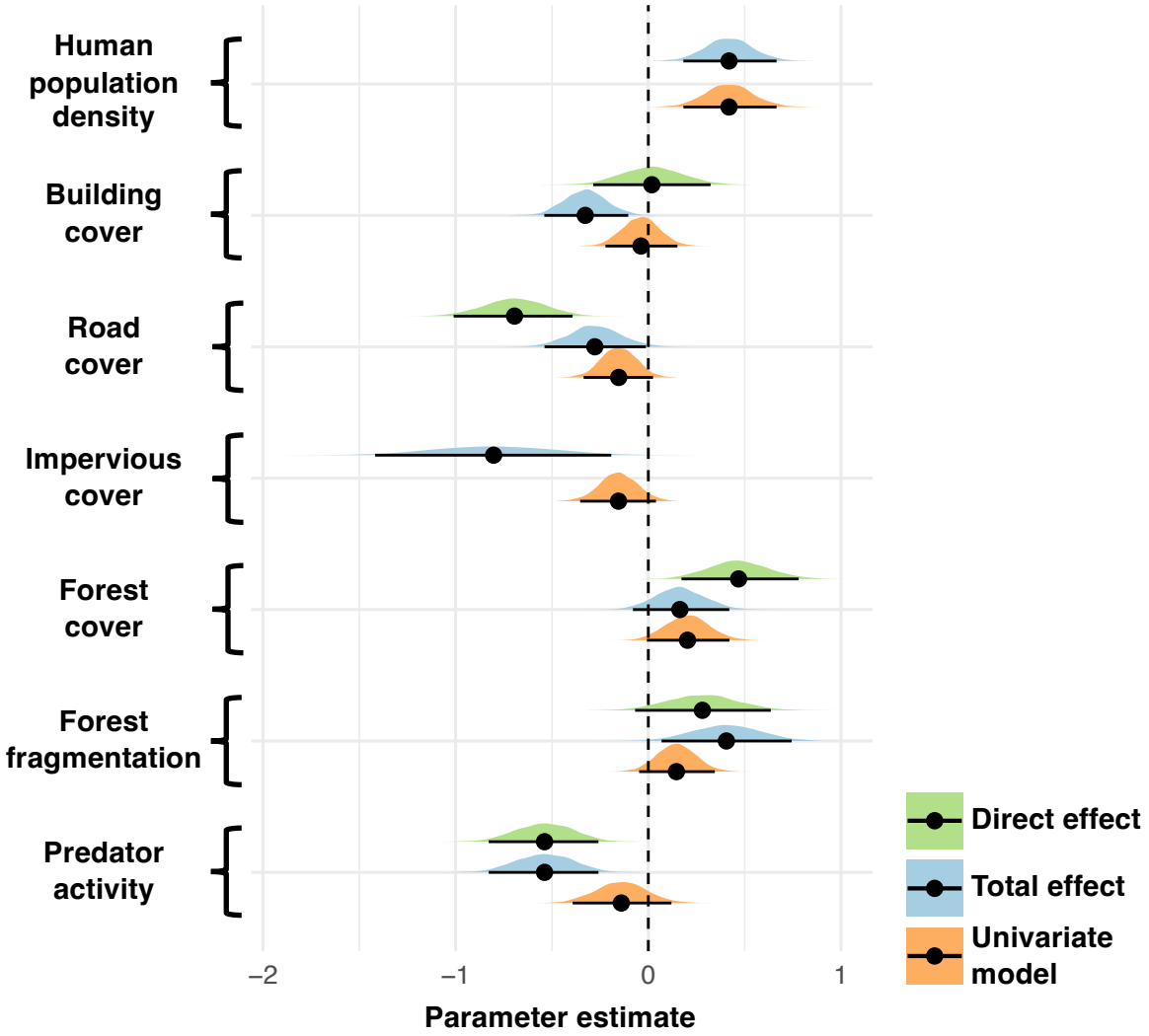

Fig. 4 Parameter estimates for direct effects (green), total effects (blue), and univariate models (orange) for predictors of coat color (proportion melanic) among eastern gray squirrels (*Sciurus carolinensis*). All parameter estimates were made with standardized predictors. Not all predictors have both direct and total effects based on our directed acyclic graph (Fig. 3B).

Despite potential unmeasured variables contributing to urban rural clines in melanism, the statistical models informed by our DAG provided novel insights into specific mechanisms maintaining the urban-rural cline in melanism. For example, prior research shows melanic squirrels are underrepresented among roadkill, relative to the gray morph—a finding attributed to humans more easily detecting and avoiding melanic squirrels against gray asphalt when driving

(Gibbs et al. 2019, Parlin et al. in press, Proctor et al. 2025). However, our findings show road cover had strong negative direct and total effects on the prevalence of melanism across the landscape (Figs. 4, 5; Table 2). This suggests melanic squirrels may be avoiding areas with greater road cover, and thus previously observed underrepresentation of melanic squirrels among roadkill (Gibbs et al. 2019, Parlin et al. in press) may be due in part to habitat selection or behavioral avoidance of roads by the melanic morph, rather than being a sole function of differences in visibility to people while driving. Isolating the direct and total effects of road cover while considering confounds (e.g., building cover) provided this new insight into why melanic squirrels are underrepresented among roadkill and contributes more evidence that within cities the prevalence of squirrel melanism is constrained by high disturbance.

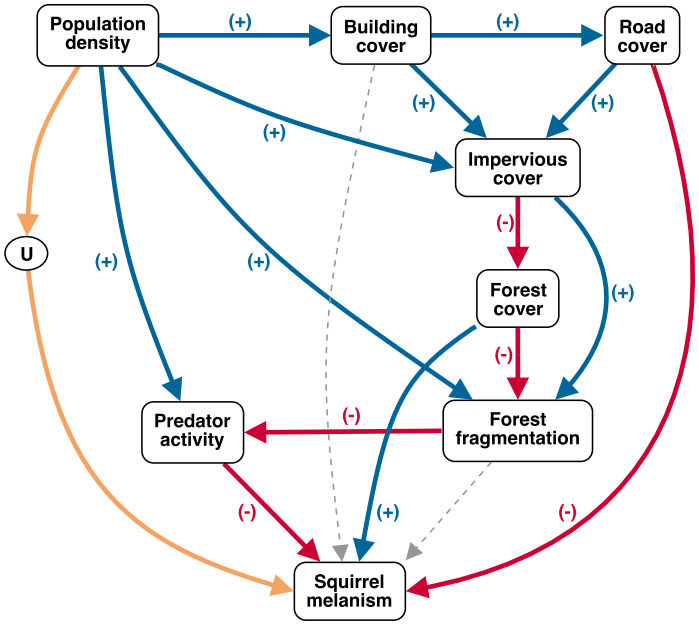

Fig. 5 Direction of estimated direct effects overlaid on our directed acyclic graph. Blue solid lines indicate positive direct effects, red solid lines indicate negative direct effects, and gray dotted lines indicate effects for which the posterior distribution of the parameter estimate broadly overlapped zero. Orange arrows show causal pathways which could not be estimated due to unobserved processes. Parameter estimates for effects on squirrel melanism can be found in Tab. 2, and parameter estimates among measured environmental variables can be found in Supplemental Tab. S1.

Another key finding from our models was the strong negative direct effect of predator activity on the prevalence of melanism (Figs. 4, 5; Tab. 2). A previous translocation experiment showed survival is lower for melanic morph than the gray morph in rural woodlands (Cosentino et al. 2023), and we know the melanic morph is generally more conspicuous than the gray morph (Proctor et al. 2025). As such, crypsis and predation pressure could contribute to patterns in melanism across the urban-rural gradient. This combination of forces might help explain the direct negative effect of predator activity *and* positive total effect of forest fragmentation on melanism. Although we did not find strong evidence of a direct effect of fragmentation on melanism, our analyses support the idea that fragmentation mediates the prevalence of melanism through its negative effect on predator activity (Fig. 5, Supplemental Tab. S1), leading to a positive total effect of fragmentation on melanism (Fig. 4, Table 2). The melanic morph may benefit from a suppressive effect of fragmentation on predator activity locally, but it is important to note there was a strong positive direct effect of human population density on predator activity (Fig. 5, Supplemental Tab. S1), indicating predator activity was greatest in the city where melanics are most common. It is possible the melanic morph uses behavioral strategies to mitigate their predation risk in densely populated areas of the city where predator activity is greatest (Sarno et al. 2015; Gaynor et al. 2019; Engel et al. 2020). Such strategies could include different activity patterns in space (including vertically from ground to canopy) and time (diel patterns), resulting in differing susceptibility to predation. Because vertical niche shifts have been linked to competition and predation, climate change, and urbanization (Rankin et al. 2018; Basham and Scheffers 2020; Borden et al. 2021; Gámez and Harris 2022), future research should investigate differences in spatial-temporal behavioral patterns between color morphs and how it may impact the landscape of fear.

367 Despite these novel insights regarding mechanisms contributing to the urban-rural cline
368 in melanism, it is intriguing that melanism is most prevalent in the city yet constrained by
369 components of physical infrastructure that are widespread in cities. This seemingly contradictory
370 pattern may be due in part to the interaction of historic and contemporary forces, including
371 ecological, evolutionary, and social (Des Roches et al. 2021). For example, melanic squirrels,
372 once common across the forested landscape, were extirpated in many areas during the period of
373 European colonialism when extensive hunting accompanied deforestation and logging in old
374 growth forests near settlements (Robertson 1973; Cronon 1991; Benson 2013; Thompson et al.
375 2013). Evidence suggests urban areas may have functioned as refugia for gray squirrels in
376 general, including the melanic morph (Benson 2013; Gibbs et al. 2019). Indeed, changing
377 attitudes about nature led to efforts to deliberately introduce gray squirrels to cities starting in the
378 late 1800s, celebrating them as "our most loved" species in cities, and establishing a social
379 attitude of tolerance and even admiration (Benson 2013). Despite well-known cases of
380 introductions of the melanic morphs to cities (e.g., Washington DC, Fischman et al. 2021), most
381 squirrel translocations are undocumented. We strongly suspect human-mediated transport and
382 bans on hunting in densely populated cities, including Syracuse, have played important roles in
383 contributing to the maintenance of urban-rural clines in squirrel melanism.
384 Although such social forces can be difficult to measure, our DAG illustrates how it can
385 be important to include unmeasured processes in a causal modeling framework. As just one
386 example, consider the direct negative effect of predator activity on melanism inferred from our
387 analyses. Based on the DAG in Fig. 3B, the adjustment set we used to infer the direct effect of
388 predator activity on melanism included human population density and fragmentation. Adjusting
389 for human population density was needed in part to close the backdoor path *Predator* ← *Human*

390 → *U* → *Melanism*. If the *Human* → *U* → *Melanism* pathway was *not* included in the DAG, the
391 direct effect of predator activity on squirrel melanism could be estimated without adjusting for
392 human population density. Doing so would lead to a biased estimate of the direct effect of
393 predator activity due to confounding with the unmeasured variable, U. Indeed, when we estimate
394 the direct effect of predator activity on melanism without adjusting for human population
395 density, the estimated effect is no longer negative (posterior mean = 0.04, 95% CI = -0.22, 0.35).
396 This positive bias is due to the positive effect of people on predator activity and the residual
397 positive association between human population density and melanism operating through the
398 unobserved processes. Graphical causal models, such as DAGs, highlight one way in which
399 unobserved processes in urban systems – known or unknown – can be explicitly considered
400 when estimating effects of interest.

401 A more striking demonstration of the value of graphical causal models is seen when
402 comparing estimated effects when appropriately adjusted based on a DAG versus univariate
403 models commonly seen in urban ecology and evolution. For example, had we relied solely on
404 univariate models to estimate the association between squirrel melanism and common
405 urbanization metrics, we might have erroneously concluded that melanism is unrelated to *any* of
406 the measured variables in our DAG other than human population density, as credible intervals
407 for all univariate predictors other than human density overlapped zero (Fig. 4; Tab. 2). This
408 finding raises significant concerns about using single-variable proxies of urbanization. For
409 example, impervious cover is commonly used to represent urbanization (Szulkin et al. 2020), yet
410 our univariate model with impervious cover resulted in a posterior distribution that was largely
411 negative, implying melanism is lowest in areas of high impervious cover (Fig. 4; Tab. 2). Had we
412 used impervious cover alone as an index of urbanization, we might have concluded there was no

evidence for an urban-rural cline, or even weak evidence for a reversed where melanism is greater in rural areas. Indeed, our own previous studies using impervious surface as a surrogate for urbanization likely underestimated true clines in melanism (Gibbs et al. 2019; Cosentino and Gibbs 2022). If using single variables as proxies of urbanization to describe biodiversity patterns along urbanization gradients cannot be avoided, our results suggest it is best to use variables that are causal ancestors of environmental change in cities, namely the density of people (Fig. 5). Distance to city center can also be effective for characterizing spatial patterns of biodiversity along urbanization gradients (Cosentino et al. 2023).

Overall, our research demonstrates how urban impacts on biodiversity are complex and involve a network of causal pathways. Environmental variables commonly used to generically represent urbanization may at best describe coarse patterns with respect to urbanization effects in ecology and evolution, but they are nearly impossible to interpret mechanistically (Moll et al. 2019). We show how DAGs can be an effective tool to clarify causal assumptions and make inferences about a variable's mechanistic effects in light of those assumptions. Causal modeling frameworks present great opportunity for disentangling the mechanisms at play in urban systems and understanding the causal drivers of biodiversity within and among cities.

Acknowledgments

Field assistance was provided by Joelee Tooley and Richard Rich. Access to field sites was provided by the New York Department of Environmental Conservation, City of Syracuse, and select private and municipal landowners.

435 **Table 1.** List of environmental variables in our models which are hypothesized to impact the
436 prevalence of melanism in eastern gray squirrels (*Sciurus carolinensis*).
437

Environmental variable	Rationale	References	Variable data source
Human population density	The foundational driver of urbanization, increasing building and road cover, driving deforestation and habitat fragmentation, and causing many unmeasured changes.	(Glover and Simon, 1975)	NASA's CIESIN gridded human population dataset (Center For International Earth Science Information Network-CIESIN-Columbia University 2017)
Building cover	Alters habitat structure and availability by reducing tree cover, but also vides nesting sites, novel microclimates, and shading that may be important for crypsis.	(Benson, 2013; Parker and Nilon 2008)	Cornell University Geospatial Information Repository
Road cover	Roadkill is a major driver of squirrel mortality in urban contexts particularly roads with ≥ 30 mph speed limit. Melanic squirrels are more visible to drivers on roads, and they have lower mortality rates from vehicular collisions than the gray morph.	(McCleery et al. 2008; Gibbs et al. 2019; Parlin et al. in press; Proctor et al. 2025)	Cover of roads at least 30 mph speed limit from NYS GIS Resources repository for public data (https://data.gis.ny.gov)
Forest cover	Increases resource availability, including food, nests, and refuge from predators.	(Parker and Nilon 2008; Koprowski et al. 2016; Merrick et al. 2016)	ESA WorldCover dataset derived from Sentinel-1 and Sentinel-2 satellite imagery (Zanaga et al. 2022)
Forest fragmentation	Restricts movements and may affect fitness of color morphs differently by altering predator activity and crypsis of each morph assuming fragmented areas are more visually open.	(Parker and Nilon, 2008; Koprowski et al. 2016)	Measured as disjunct core area density based on forest cover from ESA WorldCover dataset derived from Sentinel-1 and Sentinel-2 satellite imagery (Zanaga et al. 2022)
Mammalian predator activity	Mammalian predator activity alters predation risk as well as the landscape of fear. Differences in crypsis or behavior may cause morphs to be affected differently by predator activity.	(Gibbs et al. 2019; Cosentino and Gibbs 2022; Proctor et al. 2025)	Calculated from camera trap detections

438

Table 2. Posterior means, 95% credible intervals, and adjustment sets for parameter estimates of direct effects, total effects, and univariate models for predictors of coat color (proportion melanic) among eastern gray squirrels (*Sciurus carolinensis*). Separate models were fit to estimate direct and total effects of predictor variables on proportion melanic with appropriate adjustment covariates based on our directed acyclic graph (Fig. 3B). Univariate estimates were based on models with only the predictor variable and no adjustment covariates. Predictor variables included human population density (Human), building cover (Building), road cover (Road), impervious cover (Impervious), forest cover (Forest), forest fragmentation (Frag), and mammalian predator activity (Predator).

Direct Effects				
Predictor	*Mean*	*LCL*	*UCL*	*Adjustments*
Human	-	-	-	-
Building	0.02	-0.29	0.32	Human, Road, Forest, Frag
Road	**-0.69**	**-1.01**	**-0.39**	Human, Building, Forest, Frag
Impervious	-	-	-	-
Forest	**0.47**	**0.17**	**0.78**	Human, Building, Road, Frag
Frag	0.28	-0.07	0.64	Human, Impervious, Forest, Predator
Predator	**-0.54**	**-0.83**	**-0.26**	Human, Frag

Total Effects				
Predictor	*Mean*	*LCL*	*UCL*	*Adjustments*
Human	**0.42**	**0.18**	**0.67**	-
Building	**-0.33**	**-0.54**	**-0.10**	Human
Road	**-0.28**	**-0.54**	**-0.01**	Building
Impervious	**-0.80**	**-1.42**	**-0.19**	Human, Building, Road
Forest	0.16	-0.08	0.42	Impervious
Frag	**0.41**	**0.07**	**0.74**	Human, Impervious, Forest
Predator	**-0.54**	**-0.83**	**-0.26**	Human, Frag

Univariate Estimates				
Predictor	*Mean*	*LCL*	*UCL*	*Adjustments*
Human	**0.42**	**0.18**	**0.67**	**-**
Building	-0.04	-0.22	0.15	**-**
Road	-0.15	-0.34	0.02	**-**
Impervious	-0.15	-0.35	0.04	**-**
Forest	0.20	-0.01	0.42	**-**
Frag	0.15	-0.05	0.35	**-**
Predator	-0.14	-0.39	0.12	**-**

Literature cited

Alberti M, Palkovacs EP, Roches SD, De Meester L, Brans KI, et al (2020) The complexity of urban eco-evolutionary dynamics. BioScience 70(9):772–793

Arif S, MacNeil MA (2023) Applying the structural causal model framework for observational causal inference in ecology. Ecological Monographs 93(1):e1554

Arif S, Massey MDB (2023) Reducing bias in experimental ecology through directed acyclic graphs. Ecology and Evolution 13(3):e9947

Aronson MFJ, La Sorte FA, Nilon CH, Katti M, Goddard MA, et al (2014) A global analysis of the impacts of urbanization on bird and plant diversity reveals key anthropogenic drivers. Proc R Soc B 281(1780):20133330–20133330

Basham EW, Scheffers BR (2020) Vertical stratification collapses under seasonal shifts in climate. J Biogeogr 47:1–11

Baston D, ISciences, LLC (2022) exactextractr: Fast extraction from raster datasets using polygons (version 0.9.1) [computer software]. https://cran.r-project.org/web/packages/exactextractr/index.html

Bateman PW, Fleming PA, Fleming T, Comber SL (2012) Big city life: Carnivores in urban environments. J Zool 287(1):1–23

Bennett VJ (2017) Effects of road density and pattern on the conservation of species and biodiversity. Curr Landsc Ecol Rep 2(1):1–11

Benson E (2013) The urbanization of the eastern gray squirrel in the United States. J Am Hist 100(3):691–710

472 Borden JB, Bohlman S, Scheffers BR (2021) Niche lability mitigates the impact of invasion but
473 not urbanization. Oecologia 198:1–10
474 Bowman J, Ray JC, Magoun AJ, Johnson DS, Dawson FN (2010) Roads, logging, and the large-
475 mammal community of an eastern Canadian boreal forest. Can J Zool 88(5):454–467
476 Bu H, Wang F, McShea WJ, Lu Z, Wang D, et al (2016) Spatial co-occurrence and activity
477 patterns of mesocarnivores in the temperate forests of Southwest China. PLoS ONE
478 11(10):e016427
479 Bürkner P-C (2017) brms: An R package for Bayesian multilevel models using Stan. J Stat Softw
480 80:1–28
481 Byrnes JEK, Dee LE (2025) Causal inference with observational data and unobserved
482 confounding variables Ecol Lett 28:e70023
483 Center for International Earth Science Information Network-CIESIN-Columbia University
484 (2017) Gridded population of the world, version 4 (GPWv4): Population density, revision 11
485 [dataset]. Palisades, NY: Socioeconomic Data and Applications Center (SEDAC)
486 Cerqueira RC, de Rivera OR, Jaeger JAG, Grilo C (2021) Direct and indirect effects of roads on
487 space use by jaguars in Brazil. Sci Rep 11:22617
488 Cosentino BJ, Gibbs JP (2022) Parallel evolution of urban–rural clines in melanism in a
489 widespread mammal. Sci Rep 12:1752
490 Cosentino BJ, Vanek JP, Gibbs JP (2023) Rural selection drives the evolution of an urban–rural
491 cline in coat color in gray squirrels. Ecol Evol 13(10):e10544
492 Cronon W (1991) Nature's metropolis: Chicago and the Great West. WW Norton & Company,
493 New York

Des Roches S, Brans KI, Lambert MR, Rivkin LR, Savage AM, et al (2021) Socio-eco-evolutionary dynamics in cities. Evol Appl 14(1):248–267

Engel LD, Carlen EJ, Losos JB, Winchell KM (2020) Eastern gray squirrels (*Sciurus carolinensis*) differ in abundance and response to humans across urban habitats of St. Louis. Urb Nat 33: 1–16

Fahrig L (2003) Effects of habitat fragmentation on biodiversity. Annu Rev Ecol Evol Syst 34(1):487–515

Fahrig L (2017) Ecological responses to habitat fragmentation per se. Annu Rev Ecol Evol Syst 48:1–23

Fischman B, Cosentino B, Gibbs J (2021) City squirrels look different. Is evolution driving a color change? Discover Magazine, City Science Salon. https://www.discovermagazine.com/planet-earth/city-squirrels-look-different-is-evolution-driving-a-color-change

Francis CD, Kleist NJ, Ortega CP, Cruz A (2012) Noise pollution alters ecological services: Enhanced pollination and disrupted seed dispersal. Proc R Soc B 279(1739):2727–2735

Gámez S, Harris NC (2022) Conceptualizing the 3D niche and vertical space use. Trends Ecol Evol 37:953-962

Gaynor KM, Brown JS, Middleton AD, Power ME, Brashares JS (2019) Landscapes of fear: Spatial patterns of risk perception and response. Trends Ecol Evol 34(4):355–368

Gelman A (2006) Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). Bayesian Anal 1(3):515–534

Gelman A, Hill J (2007) Data analysis using regression and multilevel/hierarchical models. Cambridge University Press

Gibbs JP, Buff MF, Cosentino BJ (2019) The biological system—urban wildlife, adaptation, and evolution: Urbanization as a driver of contemporary evolution in gray squirrels (*Sciurus carolinensis*). In: Hall MP, Balogh S (eds) Understanding Urban Ecology. Springer, pp 269–286

Glover DR, Simon JL (1975) The effect of population density on infrastructure: The case of road building. Econ Dev Cult Change 23(3):453–468

Hijmans RJ, van Etten J, Sumner M, Cheng J, Baston D, et al (2023) raster: Geographic data analysis and modeling (version 3.6-20) [computer software].

Kellner K, Meredith M (2024) jagsUI: A wrapper around "rjags" to streamline "JAGS" analyses (version 1.6.2) [computer software]. https://cran.r-project.org/web/packages/jagsUI/index.html

Kéry M, Royle JA (2020) Applied hierarchical modeling in ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 2: Dynamic and advanced models. Academic Press

Kok ACM, Berkhout BW, Carlson NV, Evans NP, Khan N, et al (2023) How chronic anthropogenic noise can affect wildlife communities. Front Ecol Evol 11

Koppen W (1936) The geographical system of climate. In: Koppen W, Geiger R (eds) Handbook of climatology. Gebruder Borntraeger, Berlin

Koprowski J, Munroe K, Edelman A (2016) Gray not grey: The ecology of *Sciurus carolinensis* in their native range in North America. In: Shuttleworth M, Lurz P, Gurnell J (eds) Grey Squirrels: Ecology and Management of an Invasive Species in Europe. European Squirrel Initiative

Lemoine NP (2019) Moving beyond noninformative priors: Why and how to choose weakly informative priors in Bayesian analyses. Oikos 128(7):912–928

Liu J, Coomes DA, Gibson L, Hu G, Liu J, et al (2019) Forest fragmentation in China and its effect on biodiversity. Biol Rev 94(5):1636–1657

Liu S, Dong Y, Deng L, Liu Q, Zhao H, et al (2014) Forest fragmentation and landscape connectivity change associated with road network extension and city expansion: A case study in the Lancang River Valley. Ecol Indic 36:160–168

McCleery RA, Lopez RR, Silvy NJ, Gallant DL (2008) Fox squirrel survival in urban and rural environments. J Wild Mgmt 72:133–137

McDonnell MJ, Hahs AK (2008) The use of gradient analysis studies in advancing our understanding of the ecology of urbanizing landscapes: Current status and future directions. Landsc Ecol 23:1143–1155

McDonnell MJ, Pickett STA (1990) Ecosystem structure and function along urban-rural gradients: An unexploited opportunity for ecology. Ecology 71(4):1232–1237

McElreath R (2020) Statistical rethinking: A Bayesian course with examples in R and STAN. CRC Press

McRobie H, Thomas A, Kelly J (2009) The genetic basis of melanism in the gray squirrel (*Sciurus carolinensis*). J Hered 100(6):709–714

Merrick M, Evans K, Bertolino S (2016) Urban grey squirrel ecology, associated impacts, and management challenges. In: Shuttleworth M, Lurz P, Gurnell J (eds) Grey Squirrels: Ecology and Management of an Invasive Species in Europe. European Squirrel Initiative, Warwickshire, UK

561 Moll RJ, Cepek JD, Lorch PD, Dennis PM, Tans E, et al (2019) What does urbanization actually
562 mean? A framework for urban metrics in wildlife research. J Appl Ecol 56(5):1289–1300
563 Mumma MA, Gillingham MP, Johnson CJ, Parker KL (2019) Functional responses to
564 anthropogenic linear features in a complex predator-multi-prey system. Landsc Ecol
565 34(11):2575–2597
566 Nowak DJ, Hoehn RE, Bodine AR, Greenfield EJ, O'Neil-Dunne J (2016) Urban forest
567 structure, ecosystem services and change in Syracuse, NY. Urban Ecosyst 19(4):1455–1477
568 Ordeñana MA, Crooks KR, Boydston EE, Fisher RN, Lyren LM, et al (2010) Effects of
569 urbanization on carnivore species distribution and richness. J Mammal 91(6):1322–1331
570 Parker TS, Nilon CH (2008) Gray squirrel density, habitat suitability, and behavior in urban
571 parks. Urban Ecosyst 11(3):243–255
572 Parlin AF, Cosentino BJ, Lehtinen RM, McDonald JE, Sinclair ECC, Gibbs JP (in press) Road
573 mortality contributes to the evolution of an urban-rural cline in squirrel coat color.
574 Evolutionary Applications.
575 Plummer M (2017) JAGS version 4.3.0 user manual. https://sourceforge.net/projects/mcmc-
576 jags/files/Manuals/4.x/
577 Proctor J, Bryan A, Cosentino BJ, Gibbs JP (2025) Crypsis in a polymorphic mammal along an
578 urbanization gradient. Urban Ecosyst 28:94
579 R Core Team (2024) R: A language and environment for statistical computing. R Foundation for
580 Statistical Computing. Vienna, Austria. https://www.R-project.org/
581 Rankin EEW, Knowltonid JL, Gruner DS, Flaspohler DJ, Giardina CP, et al (2018) Vertical
582 foraging shifts in Hawaiian forest birds in response to invasive rat removal. PLoS ONE 10

Ritchie H, Roser M (2018) Urbanization. Our World in Data. https://ourworldindata.org/urbanization

Robertson GI (1973) Distribution of color morphs of *Sciurus carolinensis* in eastern North America. MS Thesis, University of Western Ontario, London

Rohrer JM (2018) Thinking clearly about correlations and causation: Graphical causal models for observational data. Adv Methods Pract Psychol Sci 1(1):27–42

Santangelo JS, Ness RW, Cohan B, Fitzpatrick CR, Innes SG, et al (2022) Global urban environmental change drives adaptation in white clover. Science 375:1275–1281

Sarno RJ, Parsons M, Ferris A (2015) Differing vigilance among gray squirrels (*Sciurus carolinensis*) along an urban–rural gradient on Long Island. Urban Ecosyst 18(2):517–523

Szabó B, Korányi D, Gallé R, Lövei GL, Bakonyi G, et al (2023) Urbanization decreases species richness, and increases abundance in dry climates whereas decreases in wet climates: A global meta-analysis. Sci Total Environ 859:160145

Szulkin M, Garroway CJ, Corsini M, Kotarba A, Dominoni D (2020) How to quantify urbanization when testing for urban evolution? In: Szulkin M, Munshi-South J, Charmantier A (eds) Urban Evolutionary Biology. Oxford University Press

Textor J, van der Zander B, Gilthorpe, MS, Liskiewicz M, Ellison GT (2016) Robust causal inference using directed acyclic graphs: the R package '*DAGitty*'. Int J Epidemiol 45(6):1887–1894

Thompson JR, Carpenter DN, Cogbill CV, Foster DR (2013) Four centuries of change in northeastern United States forests. PLoS ONE 8(9):e72540

Winchell KM, Reynolds RG, Prado-Irwin SR, Puente-Rolón AR, Revell LJ (2016) Phenotypic shifts in urban areas in the tropical lizard *Anolis cristatellus*. Evolution 70(5):1009–1022

Zanaga D, Van De Kerchove R, Daems D, De Keersmaecker W, Brockmann C, et al (2022) ESA WorldCover 10 m 2021 v200 (version v200) [dataset]. Zenodo https://doi.org/10.5281/zenodo.7254221

Statements and Declarations

Funding

This research was funded by the U.S. National Science Foundation (DEB-2018140, DEB-2018249).

Data Availability

Data and code are available in the following repository:
https://github.com/bcosentino/urbanCausalInference