Supplementary Materials for

Gentrification drives patterns of alpha and beta diversity in cities

Mason Fidino *et al.*

Corresponding author: Mason Fidino, mfidino@lpzoo.org

The PDF file includes:

Methods
Occupancy, alpha, and beta diversity models
Quantifying gentrification (Tables S1 to Sx)
Map of cities (Fig Sx)
References (x - y)

Methods

Biological sampling

We used data from 23 UWIN cities in the United States (U.S.) for this study. Each city followed the same systematic study design, placing motion-triggered camera traps in urban greenspace along an urbanization gradient (see Magle et al. 2019 for details). Mammal data for this study came from 12 distinct sampling periods between 2019 and 2021. Camera deployments in each sampling period were about 35 days (sd = 13.01) and began on the first of January, April, July, or October of each year. Because UWIN cities joined the network at different times, the number of sampling periods among cities varied (median = 7, minimum = 2, maximum = 12). The median number of unique camera-trapping sites per city was 35 (minimum = 23, maximum = 104).

Mammals in camera trap images were identified to species by trained experts. However, flying squirrel, gray squirrel, and cottontail rabbit species were summarized to either the subgenus or genus level given challenges in identifying them to the species level from camera trap images (Kays et al. 2022). For each camera deployment we counted the number of days each species was detected and the number of operational camera days, which were then used to estimate species occupancy and detectability within our multi-city multi-species occupancy model (Sutherland et al. 2016, Magle et al. 2021).

Overall, 48 mammal species were photographed but our multi-city, multi-species occupancy model converged when limited to the 21 species that were most frequently detected (a minimum of 75 detection days across at least 3 cities). Gray squirrels (Sciurus carolinensis or Sciurus griseus) were detected most often (~41,300 detection days) while flying squirrels (Glaucomys sp.) were detected the least (79 detection days). See Table S1 for a summary of species detected across cities.

Table S1: The number of sites species were detected and number of detections per city.

City	Species	Sites detected	Days detected
Athens, GA	armadillo	19	156
Athens, GA	bobcat	7	31
Athens, GA	cottontail sp	14	175
Athens, GA	coyote	17	191
Athens, GA	gray fox	19	131
Athens, GA	gray squirrel sp	27	839
Athens, GA	raccoon	24	335
Athens, GA	red fox	10	64
Athens, GA	striped skunk	3	9
Athens, GA	virginia opossum	20	172
Athens, GA	white tailed deer	27	1321
Bay Area, CA	bobcat	5	41
Bay Area, CA	california ground squirrel	3	45
Bay Area, CA	cottontail sp	10	180
Bay Area, CA	coyote	39	954
Bay Area, CA	fox squirrel	8	132
Bay Area, CA	gray fox	6	208
Bay Area, CA	gray squirrel sp	24	480
Bay Area, CA	mule deer	1	6
Bay Area, CA	north american river otter	1	1
Bay Area, CA	raccoon	41	1770
Bay Area, CA	red fox	4	10
Bay Area, CA	striped skunk	30	283
Bay Area, CA	virginia opossum	9	271

Bay Area, CA	white tailed deer	10	425
Boston, MA	cottontail sp	23	385
Boston, MA	coyote	21	243
Boston, MA	eastern chipmunk	14	123
Boston, MA	fisher	16	46
Boston, MA	flying squirrel sp	1	2
Boston, MA	gray fox	6	9
Boston, MA	gray squirrel sp	27	1608
Boston, MA	north american mink	1	1
Boston, MA	raccoon	26	436
Boston, MA	red fox	8	23
Boston, MA	red squirrel	4	13
Boston, MA	striped skunk	5	16
Boston, MA	virginia opossum	11	41
Boston, MA	weasel sp	1	2
Boston, MA	white tailed deer	22	361
Boston, MA	woodchuck	3	4
Chicago, IL	cottontail sp	84	1814
Chicago, IL	coyote	78	914
Chicago, IL	eastern chipmunk	28	194
Chicago, IL	flying squirrel sp	9	17
Chicago, IL	fox squirrel	78	1159
Chicago, IL	gray fox	2	2
Chicago, IL	gray squirrel sp	99	5516
Chicago, IL	muskrat	1	1
Chicago, IL	north american beaver	1	2
Chicago, IL	north american mink	10	16
Chicago, IL	raccoon	94	3062
Chicago, IL	red fox	10	26
Chicago, IL	striped skunk	46	365
Chicago, IL	virginia opossum	86	1554
Chicago, IL	weasel sp	4	5
Chicago, IL	white tailed deer	55	2563
Chicago, IL	woodchuck	3	3
Denver, CO	black bear	1	1
Denver, CO	black tailed prairie dog	9	29
Denver, CO	bobcat	3	3
Denver, CO	cottontail sp	18	107
Denver, CO	cougar	2	2
Denver, CO	coyote	29	117
Denver, CO	elk	3	10
Denver, CO	fox squirrel	37	1170
Denver, CO	mule deer	13	138
Denver, CO	north american beaver	2	2
Denver, CO	raccoon	27	252
Denver, CO	red fox	24	147
Denver, CO	red squirrel	2	3
Denver, CO	striped skunk	9	27
Denver, CO	virginia opossum	1	1
Denver, CO	white tailed deer	2	2

Houston, TX	armadillo	14	221
Houston, TX	bobcat	10	23
Houston, TX	cottontail sp	16	460
Houston, TX	coyote	18	68
Houston, TX	fox squirrel	23	489
Houston, TX	gray fox	3	9
Houston, TX	gray squirrel sp	23	434
Houston, TX	raccoon	21	360
Houston, TX	striped skunk	4	31
Houston, TX	virginia opossum	22	536
Houston, TX	white tailed deer	2	49
Indianapolis, IN	cottontail sp	40	1798
Indianapolis, IN	coyote	37	422
Indianapolis, IN	eastern chipmunk	11	66
Indianapolis, IN	fox squirrel	39	3786
Indianapolis, IN	gray squirrel sp	21	155
Indianapolis, IN	muskrat	3	6
Indianapolis, IN	north american badger	1	1
Indianapolis, IN	north american beaver	6	15
Indianapolis, IN	north american mink	9	37
Indianapolis, IN	raccoon	40	4658
Indianapolis, IN	red fox	31	220
Indianapolis, IN	red squirrel	19	181
Indianapolis, IN	striped skunk	2	3
Indianapolis, IN	virginia opossum	40	1743
Indianapolis, IN	white tailed deer	33	1091
Indianapolis, IN	woodchuck	16	240
Iowa City, IA	bobcat	6	9
Iowa City, IA	cottontail sp	35	1659
Iowa City, IA	coyote	34	309
Iowa City, IA	eastern chipmunk	19	311
Iowa City, IA	flying squirrel sp	1	1
Iowa City, IA	fox squirrel	28	690
Iowa City, IA	gray fox	2	2
Iowa City, IA	gray squirrel sp	37	3337
Iowa City, IA	muskrat	1	1
Iowa City, IA	north american badger	4	4
Iowa City, IA	north american beaver	1	1
Iowa City, IA	north american mink	19	61
Iowa City, IA	raccoon	37	3878
Iowa City, IA	red fox	35	652
Iowa City, IA	striped skunk	12	34
Iowa City, IA	virginia opossum	36	1202
Iowa City, IA	white tailed deer	36	3281
Iowa City, IA	woodchuck	20	116
Jackson, MS	armadillo	19	150
Jackson, MS	bobcat	5	13
Jackson, MS	cottontail sp	12	82
Jackson, MS	coyote	14	46
Jackson, MS	fox squirrel	4	13

Jackson, MS	gray fox	13	51
Jackson, MS	gray squirrel sp	34	2469
Jackson, MS	raccoon	34	1162
Jackson, MS	red fox	26	397
Jackson, MS	striped skunk	3	8
Jackson, MS	virginia opossum	34	641
Jackson, MS	white tailed deer	32	1054
Little Rock, AR	armadillo	24	466
Little Rock, AR	bobcat	10	32
Little Rock, AR	cottontail sp	25	655
Little Rock, AR	coyote	23	89
Little Rock, AR	eastern chipmunk	5	55
Little Rock, AR	flying squirrel sp	6	19
Little Rock, AR	fox squirrel	26	652
Little Rock, AR	gray fox	20	304
Little Rock, AR	gray squirrel sp	29	2285
Little Rock, AR	north american beaver	1	1
Little Rock, AR	north american mink	3	6
Little Rock, AR	raccoon	28	1720
Little Rock, AR	red fox	11	56
Little Rock, AR	striped skunk	11	101
Little Rock, AR	virginia opossum	29	1647
Little Rock, AR	white tailed deer	22	803
Little Rock, AR	woodchuck	8	36
Madison, WI	cottontail sp	16	218
Madison, WI	coyote	7	23
Madison, WI	eastern chipmunk	3	13
Madison, WI	gray fox	1	4
Madison, WI	gray squirrel sp	22	322
Madison, WI	north american beaver	1	1
Madison, WI	north american mink	1	3
Madison, WI	raccoon	13	101
Madison, WI	striped skunk	4	16
Madison, WI	virginia opossum	13	52
Madison, WI	weasel sp	1	5
Madison, WI	white tailed deer	7	46
Madison, WI	woodchuck	1	1
Metro LA, CA	bobcat	26	394
Metro LA, CA	california ground squirrel	20	386
Metro LA, CA	cottontail sp	28	1914
Metro LA, CA	cougar	3	8
Metro LA, CA	coyote	39	1535
Metro LA, CA	eastern chipmunk	2	5
Metro LA, CA	fox squirrel	16	79
Metro LA, CA	gray fox	8	76
Metro LA, CA	gray squirrel sp	7	127
Metro LA, CA	mule deer	21	322
Metro LA, CA	raccoon	23	114
Metro LA, CA	striped skunk	23	216
Metro LA, CA	virginia opossum	18	330

Metro LA, CA	white tailed deer	1	1
National Capital	cottontail sp	30	485
National Capital	coyote	29	82
National Capital	eastern chipmunk	21	346
National Capital	flying squirrel sp	3	6
National Capital	gray fox	2	18
National Capital	gray squirrel sp	73	4872
National Capital	north american mink	2	2
National Capital	north american river otter	2	2
National Capital	raccoon	75	3157
National Capital	red fox	75	2422
National Capital	virginia opossum	45	710
National Capital	white tailed deer	69	3349
National Capital	woodchuck	19	105
Phoenix, AZ	antelope ground squirrel	1	1
Phoenix, AZ	bobcat	48	585
Phoenix, AZ	cottontail sp	65	5190
Phoenix, AZ	cougar	8	20
Phoenix, AZ	coyote	82	3855
Phoenix, AZ	gray fox	40	655
Phoenix, AZ	harris antelope squirrel	21	612
Phoenix, AZ	hooded skunk	8	44
Phoenix, AZ	jackrabbit sp	46	2627
Phoenix, AZ	javelina	42	356
Phoenix, AZ	kit fox	11	198
Phoenix, AZ	mule deer	16	154
Phoenix, AZ	north american badger	8	77
Phoenix, AZ	north american beaver	1	3
Phoenix, AZ	north american river otter	6	25
Phoenix, AZ	raccoon	46	429
Phoenix, AZ	rock squirrel	49	758
Phoenix, AZ	round tailed ground squirrel	10	53
Phoenix, AZ	striped skunk	17	181
Phoenix, AZ	white tailed deer	1	1
Portland, Oregon	california ground squirrel	1	3
Portland, Oregon	cottontail sp	18	183
Portland, Oregon	coyote	19	159
Portland, Oregon	douglas squirrel	7	32
Portland, Oregon	fox squirrel	19	238
Portland, Oregon	gray squirrel sp	19	370
Portland, Oregon	mule deer	12	122
Portland, Oregon	north american beaver	1	1
Portland, Oregon	north american river otter	1	7
Portland, Oregon	raccoon	14	82
Portland, Oregon	striped skunk	2	4
Portland, Oregon	virginia opossum	8	73
Portland, Oregon	western chipmunk	2	15
Rochester, NY	cottontail sp	12	108
Rochester, NY	coyote	13	53
Rochester, NY	eastern chipmunk	13	253

Rochester, NY	fisher	2	5
Rochester, NY	flying squirrel sp	3	17
Rochester, NY	gray fox	5	11
Rochester, NY	gray squirrel sp	22	1308
Rochester, NY	north american beaver	1	1
Rochester, NY	north american mink	2	2
Rochester, NY	raccoon	22	300
Rochester, NY	red fox	23	506
Rochester, NY	red squirrel	8	40
Rochester, NY	striped skunk	4	5
Rochester, NY	virginia opossum	19	126
Rochester, NY	white tailed deer	20	440
Rochester, NY	woodchuck	8	58
Sanford, FL	armadillo	16	116
Sanford, FL	black bear	2	2
Sanford, FL	bobcat	8	17
Sanford, FL	cottontail sp	10	54
Sanford, FL	coyote	16	54
Sanford, FL	flying squirrel sp	3	3
Sanford, FL	gray fox	1	1
Sanford, FL	gray squirrel sp	26	439
Sanford, FL	raccoon	21	316
Sanford, FL	red fox	3	14
Sanford, FL	virginia opossum	16	159
Sanford, FL	white tailed deer	13	230
Salt Lake City, UT	black bear	1	2
Salt Lake City, UT	bobcat	12	15
Salt Lake City, UT	cottontail sp	4	11
Salt Lake City, UT	cougar	9	10
Salt Lake City, UT	coyote	31	71
Salt Lake City, UT	elk	4	6
Salt Lake City, UT	fox squirrel	1	1
Salt Lake City, UT	moose	27	66
Salt Lake City, UT	mule deer	67	570
Salt Lake City, UT	muskrat	2	2
Salt Lake City, UT	north american badger	1	1
Salt Lake City, UT	north american beaver	3	8
Salt Lake City, UT	north american porcupine	9	13
Salt Lake City, UT	raccoon	31	150
Salt Lake City, UT	red fox	15	53
Salt Lake City, UT	red squirrel	11	36
Salt Lake City, UT	rock squirrel	11	63
Salt Lake City, UT	striped skunk	25	79
Salt Lake City, UT	weasel sp	1	1
Salt Lake City, UT	yellow bellied marmot	1	1
Seattle, WA	black bear	3	15
Seattle, WA	bobcat	8	15
Seattle, WA	cottontail sp	20	438
Seattle, WA	coyote	23	339
Seattle, WA	douglas squirrel	5	14

Seattle, WA	elk	1	12
Seattle, WA	flying squirrel sp	1	2
Seattle, WA	gray squirrel sp	28	1510
Seattle, WA	mule deer	13	326
Seattle, WA	north american river otter	2	2
Seattle, WA	raccoon	28	646
Seattle, WA	striped skunk	1	1
Seattle, WA	virginia opossum	18	195
Saint Louis, MO	armadillo	8	224
Saint Louis, MO	bobcat	4	4
Saint Louis, MO	cottontail sp	25	581
Saint Louis, MO	coyote	27	350
Saint Louis, MO	eastern chipmunk	11	81
Saint Louis, MO	flying squirrel sp	2	2
Saint Louis, MO	fox squirrel	28	640
Saint Louis, MO	gray fox	8	21
Saint Louis, MO	gray squirrel sp	35	7115
Saint Louis, MO	north american mink	6	9
Saint Louis, MO	north american river otter	2	2
Saint Louis, MO	raccoon	35	2978
Saint Louis, MO	red fox	24	356
Saint Louis, MO	striped skunk	9	75
Saint Louis, MO	virginia opossum	32	1510
Saint Louis, MO	white tailed deer	25	3037
Saint Louis, MO	woodchuck	14	79
Tacoma, WA	black bear	1	2
Tacoma, WA	bobcat	3	12
Tacoma, WA	cottontail sp	29	637
Tacoma, WA	coyote	36	574
Tacoma, WA	douglas squirrel	11	231
Tacoma, WA	eastern chipmunk	5	29
Tacoma, WA	elk	5	92
Tacoma, WA	fox squirrel	1	1
Tacoma, WA	gray squirrel sp	35	2226
Tacoma, WA	jackrabbit sp	9	79
Tacoma, WA	mountain beaver	1	2
Tacoma, WA	mule deer	27	879
Tacoma, WA	raccoon	40	1333
Tacoma, WA	striped skunk	8	136
Tacoma, WA	virginia opossum	31	548
Tacoma, WA	weasel sp	5	5
Tacoma, WA	western chipmunk	5	58
Urbana, IL	bobcat	3	3
Urbana, IL	cottontail sp	31	1582
Urbana, IL	coyote	31	336
Urbana, IL	eastern chipmunk	18	399
Urbana, IL	flying squirrel sp	1	1
Urbana, IL	fox squirrel	27	1577
Urbana, IL	gray squirrel sp	33	2268
Urbana, IL	muskrat	1	3

Urbana, IL	north american badger	1	3
Urbana, IL	north american beaver	2	2
Urbana, IL	north american mink	18	83
Urbana, IL	north american river otter	2	3
Urbana, IL	raccoon	35	2519
Urbana, IL	red fox	18	115
Urbana, IL	striped skunk	16	160
Urbana, IL	virginia opossum	34	1642
Urbana, IL	weasel sp	3	3
Urbana, IL	white tailed deer	33	2267
Urbana, IL	woodchuck	16	250
Wilmington, DE	cottontail sp	23	431
Wilmington, DE	coyote	6	16
Wilmington, DE	eastern chipmunk	6	80
Wilmington, DE	flying squirrel sp	3	9
Wilmington, DE	gray squirrel sp	29	3898
Wilmington, DE	muskrat	2	10
Wilmington, DE	north american beaver	7	68
Wilmington, DE	north american mink	8	32
Wilmington, DE	north american river otter	3	8
Wilmington, DE	raccoon	29	2579
Wilmington, DE	red fox	29	2161
Wilmington, DE	red squirrel	2	32
Wilmington, DE	striped skunk	13	59
Wilmington, DE	virginia opossum	28	688
Wilmington, DE	weasel sp	1	1
Wilmington, DE	white tailed deer	28	2221
Wilmington, DE	woodchuck	19	184

TABLE S1 HERE.

Social-environmental variables

We calculated two independent variables and included both in all models. First, to represent a gradient of urban intensity we calculated the percent impervious cover within 1 km of each site from the 2019 National Land Cover Database imperviousness dataset (Dewitz and U.S. Geological Survey, 2021). Second, we determined if each site was within 500 m of a gentrifying Census tract. To quantify gentrification across a wide range of cities we modified a two-step process described by Chapple et al. (2017). For the first step, we identified Census tracts that were vulnerable to gentrification in 2010 as tracts with at least 500 residents and two of these three qualities: 1) a median income less than the city's average income, 2) a proportion of college-educated residents less than the city average, and 3) a proportion of nonwhite residents greater than the city average. To calculate gentrification vulnerability we used the 2010 US decennial Census data via the tidycensus package in R v 4.2.0 (Walker 2022, R Core Team 2022). For the second step, we used the 2019 American Community Survey (U.S. Census Bureau 2012) data to determine if a vulnerable Census tract became gentrified. Here, vulnerable tracts from the first step were identified as gentrified if they experienced a greater increase in median income between 2010 and 2019 than the average change across a city–after correcting for inflation–as well as one of two qualities: a change in college-educated residents or a change in the proportion of non-Hispanic white residents between 2010 and 2019 that exceeded the average change across the city. For additional details and summaries regarding this gentrification metric see the quantifying gentrification section of the supplemental material.

Associations between gentrification and our social-environmental variables

Among cities, on average, 25% (sd = 11%) of camera sites were within 500m of a gentrified Census tract. At sites near gentrified Census tracts, 46% (sd = 20%) of land cover was impervious, on average, while sites not near gentrified Census tracts had an average of 25% (sd = 21%) impervious cover. Within cities, Urbana, Illinois had the lowest percent of sites within 500m of a gentrified Census tract (3%) and Phoenix, Arizona had the highest (50%). See Supplemental Material X1 for additional information on within-city variation in impervious cover and maps of gentrified and non-gentrified Census tracts.

With respect to the 2019 distribution of the variables we used to quantify gentrification across cities, the median per capita income of gentrified Census tracts (mean = $68,785, sd = $28,193) was roughly $30,000 less than non-gentrified Census tracts (mean = $98,678, sd = $50,777). The proportion of non-Hispanic white residents living in gentrified Census tracts (mean = 0.28, sd = 0.26) was lower than non-gentrified Census tracts (mean = 0.48, sd = 0.30), and the proportion of people with a college degree in gentrified Census tracts (mean = 0.34, sd = 0.18) was slightly lower than non-gentrified Census tracts (mean = 0.48, sd = 0.23). Thus, gentrified Census Tracts still have lower incomes, fewer non-Hispanic white residents, and fewer college educated residents than non-gentrified Census Tracts. However, gentrified Census tracts saw greater than average shifts in these variables over time, such that the residents living there have become whiter, richer, and more educated.

Gentrification may also be associated with an increase in urban greenspace. As such, we quantified whether gentrified Census tracts had a greater increase in the proportion of greenspace (i.e., developed, open space from NLCD data) over the same time frame we used to quantify gentrification (i.e., 2010 to 2019). We did not find this to be true. After averaging the proportional increase in urban greenspace across gentrified and not-gentrified census tracts in each city, the among-city range in both types of Census tracts was effectively zero (min = -0.01, max = 0.00).

Statistical modeling

We used a meta-analytic approach to quantify associations between gentrification and impervious cover and patterns of alpha and beta diversity across U.S. cities, using a Bayesian approach for all models. However, unlike more-common meta-analyses, which must contend with issues of publication bias that can distort results (Nakagawa et al. 2022), our analysis used all available UWIN data to parameterize both alpha- and beta diversity models, resulting in a more unbiased and data-driven evaluation of our hypothesis.

To do so, we first fitted a Bayesian multi-city, multi-species occupancy model that included a first-order autoregressive term to account for repeat sampling across primary sampling periods within each city (Sutherland et al. 2016, Royle and Dorazio 2008, Magle et al. 2021). This model had three separate logit-linear predictors: one to indicate a species presence within a city's species pool, one for site-level occupancy, and one for site-level detection probability. Following Magle et al. (2021), we included the distance of each city to the known margin of a species' geographic range in the first linear predictor, with positive and negative numbers respectively indicating cities within and outside a species range. Range data came from IUCN red list data (IUCN 2020). For site-level occupancy and detection, we included impervious cover, gentrification, and the interaction between the two as slope terms in the model. All species-level parameters shared information among species and cities via their random effect structure. For a full description of this occupancy model see the occupancy, alpha, and beta diversity models section of the supplemental material. Following a 1,000 step adaptation phase and a 125,000 step burn-in, we sampled the posterior 120,000 times across 4 chains. We thinned chains by 3 for a total of 40,000 posterior samples. We assessed model convergence through a visual inspection of traceplots and ensured Gelman-Rubin diagnostics were < 1.10 (Gelman et al., 2013). Following model convergence, we simulated species occupancy at each site across the entire study area from 5000 random samples of the occupancy model's posterior distribution.

For the alpha diversity model we calculated 1) the expected species richness at each site and 2) the standard deviation in this estimate across the 5,000 posterior samples. To limit the effect of individual years on these estimates we calculated species richness at a site across all possible sampling periods. This resulted in one estimate per site across cities. We then fitted a varying intercept, varying slope log-linear model to these data, which treated species richness as the response variable but also incorporated the associated uncertainty in this estimate (Kery and Royle citation). Intercept and slope terms were treated as city-level random effects. We included impervious cover, gentrification, and the interaction between the two as covariates. See Supplemental Material 2 for a full description of the alpha diversity model.

For the beta diversity model we calculated 1) pairwise community dissimilarity between pairs of sites within each city (i.e., Sørensen's dissimilarity index) and 2) the standard deviation in this estimate across the 5,000 posterior samples (Legendre & Legendre 2012). Like the alpha diversity model, beta diversity estimates were made across all primary sampling periods. We then fitted a varying intercept, varying slope generalized dissimilarity model to these data, which treated pairwise dissimilarity as the response variable (GDM CITATION). This model used a clog link function and had an inverse link function of $1 - \exp(-\mu)$, where μ is the linear predictor for one data point. Similar to the alpha diversity model, the beta diversity model incorporated the associated uncertainty in the beta diversity estimate. Intercepts and slopes were treated as city-level random effects. Because community composition may be more similar in nearby sites, we controlled for geographic distance between site pairs by including it as a covariate. We also included differences in impervious cover and gentrification between sites as covariates. However, because this model uses I-spline basis functions to incorporate possible non-linear responses we could not include an interaction between gentrification and impervious cover in this model (GDM CITATION). See the occupancy, alpha and beta diversity models section of the supplemental material.

Occupancy, alpha, and beta diversity models

The multi-city multi-species autologistic occupancy model

This model is almost exactly the same model as we had used in Magle et al. (2022). For s in $1, ..., S$ species and c in $1, ..., C$ cities, π_{sc} is the probability species s is within city c. Futher, let x_{sc} be a Bernoulli random variable that equals 1 if the species is within that city and is otherwise zero such that $x_{sc} \sim \text{Bernoulli}(\pi_{sc})$. We made π_{sc} a function of one covariate–the distance a city is from the known edge of a species extent–using the logit link. This covariate was positive if a city was within a species extent and negative if it was outside. We compiled range information from IUCN red list data (IUCN, 2020). Thus the linear predictor for this level of the model was

$$\text{logit}(\pi_{sc}) = \boldsymbol{d}_s \boldsymbol{h}_c$$

where $\boldsymbol{d}_s$ is a vector of species-specific covariates while $\boldsymbol{h}_c$ is a vector of conformable regression coefficients where the leading element is 1 to accomodate the model intercept.

As the first level of the model estimates a species presence at the city-level, the next level estimates species presence within cities conditional on their presence in a city. Given that the number of sites and sampling periods varies across cities, we add a city subscript to define i_c in $1, ..., I_c$ sites and t_c in $1, ..., T_c$ sampling periods. However, for simplicity we drop these specific subscripts while we explain the model. Additionally, let z_{scit} be a Bernoulli random variable and ψ_{scit} be the probability of occupancy. As such, $z_{scit} \sim \text{Bernoulli}(\psi_{scit} x_{sc})$. As with the previous level of the model, ψ_{scit} can be made a function of covariates via the logit link. As a depature from the Magle et al. (2022) parameterization, we added a first-order autologistic term to account for any temporal dependence in occupancy status between adjacent sampling periods within a city. Thus, for the first time period in a city, the logit-linear predictor was

$$\text{logit}(\psi_{scit=1}) = \boldsymbol{\beta}_{sc} \boldsymbol{f}_{ci}$$

where $\boldsymbol{\beta}_{sc}$ is a vector of species and city-specific parameters and $\boldsymbol{f}_{ci}$ is a vector of conformable covariates whose first value is 1 for the model intercept. After the first sampling season, we added our autologistic term

$$\text{logit}(\psi_{scit}) = \boldsymbol{\beta}_{sc} \boldsymbol{f}_{ci} + \theta_{sc} z_{scit-1}, \text{ for t} > 1.$$

For the data model, y_{scit} was the number of days species s was detected at city c and site i on sampling season t. Given j_{cit} days of sampling, we assumed y_{scit} is a binomial random variable conditional on species presence

$$y_{scit}|z_{scit} \sim \text{Binomial}(j_{cit}, \rho_{scit} z_{scit})$$

where ρ_{scit} is the daily probability of detection that can be made a function of covariates with the logit link,

$$\text{logit}(\rho_{scit}) = \boldsymbol{\alpha}_{sc}\boldsymbol{g}_{ci}$$

where $\boldsymbol{\alpha}_{sc}$ is a vector of parameters and $\boldsymbol{g}_{ci}$ is a vector of conformable covariates where the leading value is 1 to accommodate the intercept.

Given that some species occur across multiple cities, there are multiple levels in which species could partially share information. At the top-level of the model we have the simplest hierarchical parameterization for parameters associated to π_{sc}, namely that there is a community mean for each parameter of which species-level coefficients vary around. We show this for the model intercept with the understanding that the same parameterization applies to all logit-scale covariates in this part of the model.

$$\begin{aligned}
\bar{d}_{0s} &\sim \text{Cauchy}(0, 2.5) \\
\sigma_{d_0} &\sim \text{Inv-Gamma}(1, 1) \\
d_{0s} &\sim \text{Normal}(\bar{d}_{0s}, \sigma_{d_0})
\end{aligned}$$

For the rest of the latent state model we add an additional hierarchical level to the model. However, this parameterization also aligns with the data model, and as such so we only describe it here once for the model intercept of the latent-state model. For example, for the model intercepts we begin with a community-level average among species and cities ($\bar{\beta}_0$). This parameter partially informs a species-level average among cities ($\bar{\beta}_{0s}$), which then informs species-specific coefficients in all cities (β_{0sc}).

$$\begin{aligned}
\bar{\beta}_0 &\sim \text{Cauchy}(0, 2.5) \\
\sigma_{\beta_0} &\sim \text{Inv-Gamma}(1, 1) \\
\bar{\beta}_{0s} &\sim \text{Normal}(\bar{\beta}_0, \sigma_{\beta_0}) \\
a_{\beta_0} &\sim \text{Uniform}(0, 10) \\
b_{\beta_0} &\sim \text{Uniform}(0, 10) \\
\sigma_{\beta_{0s}} &\sim \text{Inv-Gamma}(a_{\beta_0}, b_{\beta_0}) \\
\beta_{0sc} &\sim \text{Normal}(\bar{\beta}_{0s}, \sigma_{\beta_{0s}})
\end{aligned}$$

We added the hyperparameters for the shape (a_{β_0}) and rate (b_{β_0}) of the Inv-Gamma distribution to account for the fact that some cities may only have one sampling period of data. Note that the above parameterization also applies to the autologistic term of the model (θ_{sc}).

Finally, the latent state and data model included one other set of parameters to account for variation in occupancy or detectability across sampling seasons. As we have already added hierarchical structure via the centered parameterization of the other model parameters, we usd a non-centered parameterization here for this level of variability. Again, we show this for the latent state, but with a swapping of subscripts this could readily be applied to the data model as well.

$$\begin{aligned}
a_{\psi} &\sim \text{Uniform}(0, 10) \\
b_{\psi} &\sim \text{Uniform}(0, 10) \\
\sigma_{\psi c} &\sim \text{Inv-Gamma}(a_{\psi}, b_{\psi}) \\
\beta_{sct} &\sim \text{Normal}(0, \sigma_{\psi c})
\end{aligned} \tag{1}$$

With this parameterization, β_{sct} is a difference term that represents the logit-scale difference in occupancy for species s at city c from their average β_{0sc} (i.e., the model intercept). Again, like with the other parameters in this part of the model, we used hyperparameters for the Inv-Gamma distribution because not every city had more than one season of data.

The alpha diversity meta-analytic model

From our occupancy model we created a posterior distribution for the latent state of each species at each site across all cities (z_{scit}). From this, we derived two quantities for this model:

1. The mean expected species richness at each site across all sampling periods (r_{ci}). To do so, we calculated the number of unique species detected in city c and site i across the t sampling seasons and took the median across 5000 posterior samples.
2. The standard deviation of the first quantity across those 5000 posterior samples (σ_{ci}). This quantifies our level of uncertainty with the first estimate.

Following this, let $\boldsymbol{\beta}_c$ be a vector of city-specific regression coefficients and $\boldsymbol{x}_{ci}$ be a vector of city and site specific covariates where the leading element is 1 to account for the intercept. Within the linear predictor we also added an additional residual variation term, ϵ_{ci}, which was given a $\sim$ Inv-Gamma$(1, 1)$ prior. Thus, the log-linear predictor was

$$\log(\mu_{sci}) = \boldsymbol{\beta}_c \boldsymbol{x_{ci}} + \epsilon_{ci}$$

and following Kery and Royle (YEAR), we accounted for variability in the response variable with an additional level of the model

$$r_{ci} \sim \text{Normal}(\mu_{sci}, \sigma_{ci})$$

We treated the intercept and slope terms as city-level random effects. For example, the prior for the model intercept was

$$\begin{aligned} \bar{\beta}_0 &\sim \text{Cauchy}(0, 2.5) \\ \sigma_{\beta_0} &\sim \text{Inv-Gamma}(1, 1) \\ \beta_{0c} &\sim \text{Normal}(\bar{\beta}_0, \sigma_{\beta_0}) \end{aligned}$$

and the same specification was also applied to the slope terms as well (though not described here).

The beta diversity meta-analytic model

From our occupancy model, we created a posterior distribution for the latent state of each species at each site across all cities (z_{scit}). From this, we derived three quantities:

1. The mean pairwise Sorensen dissimilarity between pairs of sites within each city across all sampling periods (v_{cik}, where the subscript k denotes one of the sites in city c that is not the ith site). To do so, we calculated the number of unique species detected in city c and site i across the t sampling seasons. Following this, we calculated the Sorensen dissimilarity metric among all pairs of sites within each city using `vegan` in `R` across 5000 posterior samples. Finally, we took the median across all posterior samples for each site-pair
2. The standard deviation of the first quantity across those 5000 posterior samples (σ_{cik}). This quantifies our level of uncertainty with the first estimate.
3. The mean expected number of unique species between a site-pair (w_{cik}).

To estimate pairwise dissimilarity as a function of covariates we modified a generalized dissimilarity model to account for parametetric uncertainity of the response variable v_{cik} (GDM citation). To do so, GDMs estimate the relationship between dissimilarity and environmental or spatial differences between pairs of sites with a clog link (Mokany et al. 2022):

$$d_{cik} = 1 - \exp(-\eta_{cik})$$

where d_{cik} is the biological dissimilarity between sites i and k within city c and η_{cik} is the predicted ecological distance between (i.e., the linear predictor). Given p in 1,..,P covariates, the ecological distance between sites is

$$\eta_{cik} = b_0 + \sum_{p=1}^{P} |f_p(x_{cip}) - f_p(x_{ckp})|$$

where b_0 is the model intercept (i.e., the expected pairwise dissimilarity between sites with identical environments). Covariates are further transformed within GDMs which 1) uses I-spline basis functions (Ramsay, 1988) and 2) constrains slope terms to be non-negative. Doing so allows the effect of each covariate to non-linearly vary while also ensuring that beta diversity increases monotonically as sites are more different from one another (a core assumption of this model). More specifically, the I-spline basis function for predictor p with 3 basis functions (the default used for GDMs) is

$$fp(x_{cip}) = \sum_{j=1}^{3} a_{cpj} I_{pj}(x_{cip})$$

where a_{cpj} is a non-negative coefficient for the jth I-spline and I_{pj} is the jth I-spline of the covariate x_{cip}. For our own model, the binary gentrification status of site-pairs was not sent through an I-spline basis function. Instead, we used a dummy variable that took the value of 1 if a pair of sites differed in their gentrification status and was otherwise 0. Regardless, all slope terms were constrained to be non-negative. Thus, given η_{cik} and the total number of species between site pairs (w_{cik}), the first level of our model is

$$\begin{aligned} d_{cik} &= 1 - \exp(-\eta_{cik}) \\ \sigma_{cik} &= \sqrt{\frac{d_{cik} - (1 - d_{cik})}{w_{cik}}} \\ \mu_{cik} &\sim \text{Half-Normal}(d_{cik}, \sigma_{cik}) \end{aligned} \tag{2}$$

where σ_{cik} is the binomial variance function (Mokany et al. 2022) and the half-Normal is constrained to be non-negative. Following this, we account for variation in the measurement of our response variable

$$v_{cik} \sim \text{Half-Normal}(\mu_{cik}, \sigma_{cik})$$

where again σ_{cik} is measured based on the output of the occupancy model and provided as data to this model.

We treated all coefficients in the linear predictor as city-level random effects. For example for the model intercept the prior specification would be $b_c \sim \text{Half-Normal}(\bar{b}, \sigma_b)$ where $b \sim \text{Half-Normal}(0, 10)$ and $sigma_b \sim \text{Inv-Gamma}(1, 1)$. The Half-Normal distributions ensure that the coefficients will be non-negative.