Over the last 100 years urbanization has been a dominant driver of biodiversity change (McDonald et al. 2020). This change is, in part, due to humanity’s transition from rural to urban life. Since the start of the 20th century, the proportion of people living in cities has increased from 10% to over 50% (Citation). As a result, the amount and intensity of urban land cover has increased worldwide (Citation) and the largely negative influence urbanization has had on biodiversity has become undeniable (Citation). Cities are, after all, often located in biodiversity hotspots and the associated habitat loss that comes along with urbanization is a main driver of species extirpations or extinctions (Citation). Yet, some non-human species persist, and at times thrive, in the pockets of green space within cities. As such, the selective pressures of urban environments provide a unique opportunity to explore ecological processes across space or through time.

One underexplored process in urban environments is how species composition changes through time. Certainly, long-term comparisons of natural habitat turned urban often demonstrate a decrease in species richness and substantial community turnover as the historically present species are lost and replaced by urban-adapted species (Tingley et al. 2013, other papers). Yet, a forest turned suburban lot over 100 years likely tells a different story than a city over the same time frame. Cities have become larger and more densely populated over the last century (citation), which may heighten their selective pressures and filter out additional species, for example by further favoring species with generalist diets (citation, probably aronson hierarchical filtering paper, Pysek et al. 2004). If this is the case, species richness should decrease and species present that remain may have shared traits that increase their ability to persist in urban environments. Conversely, changes in species composition through time could instead reflect large-scale changes in the distribution and abundance of the regional species pool (Ward study?, Murgui 2014). If this is the case, species turnover within a city should reflect regional trends over time. As urbanization has been a primary driver of biodiversity change, understanding how species respond to urban environments through time will likely help THOUGHTS ON THIS? To make such comparisons through time, however, historical data must be available.

To explore how species composition changes over time in a city we replicated a historical bird survey in Lincoln Park, the largest park in Chicago, IL, USA. The original surveyors, Herbert and Alice Walter, surveyed Lincoln Park from March to May between 1898 and 1903 and wrote a field guide for city bird watchers (Walter and Walter 1904). Along with accurate species descriptions, the Walter’s book summarized their field notes for future comparison. The Walter’s survey was then continued by William Dreuth – a stock clerk, amateur naturalist, and expert birder – between 1927 and 1932 (Clark and Nice 1950). Decades later, a colleague found a copy of The Walter’s book in a Chicago antique store, and subsequent investigations into the Walter’s survey led us to William Dreuth’s original field notes in the back halls of the Chicago Academy of Sciences. These discoveries inspired us to continue these Lincoln Park surveys between 2012 and 2015. Our goals were to determine 1) how this bird community changed over a century and 2) if the temporal trends observed in Lincoln Park mimic those observed throughout Illinois. The extreme levels of urbanization Chicago experienced in the last century make this survey unique. Since the Walter’s bird counts in 1898, the height of the tallest building in downtown Chicago more than quadrupled, the average human population density doubled, and automobiles replaced horses (Randall et al. 1999, Sovacool 2009). Replicating these historical surveys highlights frequency changes of the resident and migrant birds that comprise Chicago’s avian community across 100 years of urbanization in a large metropolitan city.

# Methods

## Study Area

Lincoln Park is mostly linear and located along the western shore of lake Michigan, roughly 4.5 km north of downtown Chicago (Figure 1). The park has increased in size since the original bird survey from about 125 ha in 1898 to 481 ha currently (Figure 1). The land added has been north of Lincoln Park’s original boundary (Clark and Nice 1950; Figure 1). Primarily built for recreation, Lincoln Park has ponds, sports fields, nature areas, and expanses of turf grass peppered with mature trees and shrubs. These qualities have changed little since 1898. The Lincoln Park Zoo, which lies in the middle of the original park boundaries, was also present across all three survey periods.

Throughout Chicago, temperatures increase from near freezing at the start of the migratory season to roughly 15 °C by May (NOAA 2019). In March, sunrise begins near 7:30 AM and, by May, advances to about 5:00 AM. Regarding precipitation, snow is common in March. Sometimes it snows in April (NOAA 2019).

## Replicating the historical bird surveys

There is little information about the routes that the Walters and Dreuth walked to count birds. Thus, we assumed the Walters (1898 - 1903) did not follow a standard methodology, surveyed the entire park, and counted species by sight and sound because their birding guide included apt descriptions of birds and their song (Walter and Walter 1904). Dreuth, the surveyor between 1927 – 1933, included the count’s date, time, and end points of the path traveled in his field notes. Neither historical survey described the path traveled in the park, the distance at which birds were identified from the path, or the speed at which a surveyor walked.

At a minimum, historical surveys included the number of days per year counts were conducted. The Walters averaged 75.66 (min = 66, max = 87) counts per year. Because the Walters included a figure of observed species richness per day between May 7 and 20 in their birding guide – a time they deemed the height of migration – we assumed they counted birds each weekday in March and April but every day in May (Walter and Walter 1904). Conducting counts this way results in about 71 counts per year. Dreuth averaged 57.66 (min = 17, max = 85) counts per year.

We replicated these surveys using the available information but followed a standard methodology to increase repeatability. We delineated a 2.45 km line-transect from the northern-most point to the southern-most point of the original park boundaries (Figure 1). In the last century, many walking paths have not changed. Therefore, we used walking paths whenever possible. To count birds, one trained observer (Author’s initials) walked the transect at about 2 km hour-1 and started roughly one hour past local sunrise during clear weather. Species were identified by sight and sound within 50 m of the transect. A count’s start point was switched daily and started in the north or the south of the park (Figure 1). We conducted counts each weekday of March and April and then daily in May. This protocol more so replicates the Walter’s survey. We did this because their book was the basis for our survey. Lincoln Park birders made us aware of Dreuth’s survey after our surveys began.

Some species were not included in the historical surveys. House sparrows (*Passer domesticus*) were, and still are, abundant in Lincoln Park but the Walters and Dreuth did not count them. The Walters also excluded gannets and grebes, cranes and rails, ducks and geese, loons, terns and gulls, and pelicans, from their counts. Dreuth documented all species but the house sparrow. While we recorded all species encountered on our counts, we were only able to compare with what the historical surveyors recorded.

## Statistical analysis

To quantify differences in alpha diversity among survey periods and determine if species richness decreased over time we used the vegan package in R v. 4.0.0 (CITATION, R core team 2020). To do so, we created a community matrix for the three survey periods (1898 – 1903, 1922 – 1927, and 2012 – 2015) with species along the columns and survey years along the rows. If a species was detected on a given survey year the associated cell would equal 1 in a matrix, otherwise it was 0. To account for potentially missed species in a survey period and correct for differences in sampling effort we estimated alpha diversity from the communities matrices with the specpool function in vegan, which provides a point estimate of species richness as well as standard errors of the point estimate (O’Hara 2005).

To determine if, over time, generalist species became more frequent or if species trajectories reflected statewide changes in occupancy we used a binomial generalized linear mixed model (GLMM). Our response variable for this analysis was the proportion of days a species was observed per year weighted by the number of counts conducted per year. We used this response variable, which is comparable within but not among species (Royle and Nichols 2003), because abundances were not available in the first time period. We quantified each species generalist ability by deriving their diet and foraging breadth from a global bird and mammal functional trait database (Wilman et al. 2014). For diet breadth we counted the number of different diet categories that each species was known to forage on which could include invertebrates, mammals and birds, reptiles, fish, vertebrates, carrion, fruit, nectar, seeds, or other plant material such as seedlings, weeds, or lichen (n = 10, mean = 1.6, min = 1, max = 6). For foraging breadth, we counted the number of different strategies a species was known to use to forage which could include foraging below water, on water, on the ground, in the understory, in the middle to high levels of trees, in the tree canopy, or in the air (n = 7, mean = 2.4, min = 1, max = 5). To determine if changes in species relative frequency tracked statewide trends over time, we used statewide estimates of the absolute change in occupancy of breeding birds throughout Illinois between 1906 and 2008 (Ward et al. 2018). Only 55% of the birds in our analysis were represent in the Ward et al. (2018) dataset. Birds not represented, such as non-breeding birds, were given a value of 0.

Our GLMM included a few other independent variables. In addition to the three forementioned continuous variables (foraging breadth, diet breadth, and statewide trends) we also included survey period (1898 – 1903, 1922 – 1927, and 2012 – 2015) as categorical variables. Given the large time span between our own sampling and the middle survey relative to the time between the first and second sampling periods, we only included interactions between the 2012-2015 period and the three continuous variables to quantify if the response to these variables changed over time. For the random effect structure of the model we allowed all three survey periods to vary by species. This model was fit in V 2.19.1 of Stan (Stan Development Team 2018) using the stan\_glmer function in rstanarm (Goodrich et al. 2018). Following a 1,000 step burn-in the posterior was sampled a total of 24,000 times across 6 chains. The intercept of the model was given a vague Cauchy(0,5) prior while slope terms were given a Cauchy(0, 2.5) prior (Gelman citation). Some data had to be censored from our GLMM analysis. For example, a species was removed if they were only observed in one of the three survey periods, but were included if they were observed in at least two of the three survey periods. Likewise, waterbirds, shorebirds, and house sparrows were excluded from this analysis due to their omission from the first survey period.

Finally, we also quantified how much variability survey period explained in community composition. To do so we fit a PERMANOVA to a combined version of the three community matrices used for the species richness analysis except instead of representing species presence / absence as binary we input the proportion of days they were detected each year (citation). Survey period was used as a predictor in the PERMVANOVA.