DR MARK ALLAN DITMER (Orcid ID : 0000-0003-4311-3331)

MS FABIOLA IANNARILLI (Orcid ID : 0000-0002-7018-3557)

Article type : Research Article

Handling Editor: Dr Laura Prugh

Artificial night light helps account for observer bias in citizen science monitoring of an expanding large mammal population

Mark A. Ditmer[1]*, Fabiola Iannarilli[2], Andrew N. Tri[3], David L. Garshelis[3], Neil H. Carter[1]

[1]School for Environment and Sustainability, University of Michigan, Ann Arbor, MI 48109, USA
[2]Department of Fisheries, Wildlife, and Conservation Biology, University of Minnesota, Saint Paul, MN, 55108, USA
[3]Minnesota Department of Natural Resources, Grand Rapids, MN, 55744, USA
*mditmer@umich.edu

Abstract

1. The integration of citizen scientists into ecological research is transforming how, where, and when data are collected, and expanding the potential scales of ecological studies. Citizen-

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/1365-2656.13338

science projects can provide numerous benefits for participants, while educating and connecting professionals with lay audiences, potentially increasing acceptance of conservation and management actions. However, for all the benefits, collection of citizen-science data is often biased towards areas that are easily accessible (e.g. developments and roadways), and thus data are usually affected by issues typical of opportunistic surveys (e.g. uneven sampling effort). These areas are usually illuminated by artificial light at night (ALAN), a dynamic sensory stimulus that alters the perceptual world for both humans and wildlife.

2. Our goal was to test whether satellite-based measures of ALAN could improve our understanding of the detection process of citizen scientist-reported sightings of a large mammal.

3. We collected observations of American black bears (*Ursus americanus;* n = 1,315) outside their primary range in Minnesota, USA, as part of a study to gauge population expansion. Participants from the public provided sighting locations of bears on a website. We used an occupancy modelling framework to determine how well ALAN accounted for observer metrics compared to other commonly used metrics (e.g. housing density).

4. Citizen scientists reported 17% of bear sightings were under artificially-lit conditions and monthly ALAN estimates did the best job accounting for spatial bias in detection of all observations, based on AIC values and effect sizes ($\grave{\beta}$ = 0.81, 0.71 – 0.90 95% CI). Bear detection increased with elevated illuminance; relative abundance was positively associated with natural cover, proximity to primary bear range and lower road density. Although the highest counts of bear sightings occurred in the highly illuminated suburbs of the Minneapolis-St. Paul metropolitan region, we estimated substantially higher bear abundance in another region with plentiful natural cover and low ALAN (up to ~375% increased predicted relative abundance) where observations were sparse.

5. We demonstrate the importance of considering ALAN radiance when analyzing citizen scientist-collected data, and we highlight the ways that ALAN data provides a dynamic snapshot of human activity.

Keywords: bears, geographic expansion, human-wildlife interactions, occupancy model, spatial bias, species monitoring

Introduction

The integration of citizen science into research provides numerous benefits to the public and the research community. Citizen scientists benefit from project participation by receiving an authentic learning experience, place-based interactions with nature that deepen connections to the study area, and a sense of involvement in the research and management process (Dickinson et al., 2012; Newman et al., 2017). Developing projects that involve citizens can also serve as an outreach tool that increases the public's knowledge of species and helps connect lay people with scientists (Bonney et al., 2009). Moreover, researchers in citizen science projects can assess or monitor ecological processes and environmental change at greater spatio-temporal scales than would otherwise be possible (Dickinson, Zuckerberg, & Bonter, 2010). In addition to increased ecological inference, methods that leverage citizen scientists realize several practical advantages, including decreased data processing time (Swanson, Kosmala, Lintott, & Packer, 2016) and costs of data collection (Sullivan et al. 2009), as well as increased quantities of data collected (Bonney et al. 2009). Indeed, the number of studies that use citizen scientist-collected data has grown dramatically (Follett & Strezov, 2015; Silvertown, 2009). Wildlife monitoring applications include the assessment of changes in species' geographic ranges (Wilson, Anderson, Wilson, Bertram, & Arcese, 2013), population trends (Massimino, Harris, & Gillings, 2018), and biodiversity (Tulloch, Possingham, Joseph, Szabo, & Martin, 2013).

Gaining inference about ecological processes using opportunistically-collected citizen scientist data must account for the inherent biases of data collection (Altwegg & Nichols, 2019; Isaac, Strien, August, Zeeuw, & Roy, 2014). To maximize participation, citizen-science projects require data collection protocols to be simple (Dickinson et al., 2012), typically allowing for passive and observational data, resulting in large heterogeneity in how, where, and when data are collected (Kelling et al., 2015). As a result, observations tend to occur where people are present and able to detect the target species, oftentimes in areas that are the most accessible (Tulloch et al., 2013), yielding a spatial bias in the observations such that they do not reflect the spatial distribution or abundance of the species (Hugo & Altwegg, 2017). While numerous efforts have been made to account for observer effort and filter errant observations by modelling the observation process (Kelling et al., 2015), spatial bias remains an area ripe for improvement.

One important source of spatial bias is artificial light at night (ALAN). Areas most accessible by citizen observers, such as near residential developments or roads, are also those with night lighting (e.g. porch, streetlights). Lighted areas directly increase the probability of detection during crepuscular and nighttime hours. ALAN has become pervasive globally (Gaston, 2018), extending far from urban areas into protected areas (Garrett, Donald, & Gaston, 2019) and fundamentally altering the perceptual landscape for both humans and wildlife. Thus, models of wildlife distributions that omit ALAN run the risk of under or overestimating wildlife occurrence, especially along the wildland-urban interface, producing a bias that increases as ALAN increases (Kyba et al., 2017). To date, however, the degree to which ALAN biases wildlife data collected by citizen scientists has not been assessed, nor has that bias been incorporated into predictive models of wildlife occurrence.

Here we utilized an occupancy-modelling framework to test whether spatially explicit estimates of ALAN improve modelled detection processes in opportunistically collected wildlife observations by citizen scientists. Occupancy models explicitly account for detection bias while estimating species occurrence by separating ecological processes from detection processes within the same model (MacKenzie et al., 2017). Occupancy models are well suited to citizen science projects because they can test the influence of covariates that may influence either process (Kéry et al., 2010; Strien, Swaay, & Termaat, 2013), such as bias in animal detections by participants (Sun, Royle, & Fuller, 2019). However, even when the source of variation in detection process is known, such as ambient noise in avian surveys (Simons, Alldredge, Pollock, & Wettroth, 2007), there is often no spatially-explicit estimate that can be collected across large spatial scales or with regular frequency; this is especially problematic for highly mobile animals. Some studies have incorporated spatially-explicit estimates of sampling bias to better account for greater site accessibility of citizen science observers, such as proximity to roads, urban areas (Reddy & Dávalos, 2003; Warton, Renner, & Ramp, 2013), and human population density (Mair & Ruete, 2016). However, these metrics of the human footprint are static and do not capture its spatially and temporally dynamic nature, nor the changes to the sensory landscape created by ALAN associated with human activities. Choosing a variable that can be collected regularly and that accounts for observer bias at large spatial scales would be especially useful for tracking the spatio-temporal dynamics of observer bias.

Here, we capture the dynamic changes of the human footprint across the landscape by using recently developed estimates of human-generated night light (Román et al., 2018) produced by the US National Aeronautical and Space Administration (NASA). ALAN radiance levels correlate with spatial changes in human activity (Gaston, Bennie, Davies, & Hopkins, 2013), such as population and economic growth (e.g., natural gas drilling), and are collected at relatively fine scales (~ 1km^2) on a daily basis, dynamically representing seasonal shifts in human space use (e.g., ski resorts that are operational during only a few months; changes in traffic volume patterns). Quantifying and mapping ALAN may additionally identify areas in which nighttime lighting increases the chance of observing species during crepuscular and nocturnal periods.

We apply our occupancy modelling framework to a citizen-science project aimed to investigate range expansion of American black bears (*Ursus americanus*) in Minnesota, USA. Black bear population abundance and geographic range have been steadily increasing throughout much of North America (Scheick & McCown, 2014), owing to the bear's mobility, relatively high level of tolerance for human presence (and vice versa), and ability to exploit anthropogenic food sources (e.g. crops, trash, bird feeders; Tri et al., 2016, Evans et al. 2017). This propensity to forage for calorically-rich anthropogenically-sourced foods can bring them into close proximity to humans and result in human-bear conflicts (Wilton, Belant, & Beringer, 2014). Understanding where bears are expanding their range, and consequently elevating the risk of conflict with humans is of particular interest to wildlife managers (e.g. Evans, Hawley, Rego, & Rittenhouse, 2014).

In 2018, the Minnesota Department of Natural Resources (MNDNR) launched an online citizen-science data collection program, asking the public to report sightings of bears outside the forested, northern portion of the state, which constitutes primary bear range. The goal was to track the expansion of the population into less forested regions, after anecdotal reports suggested an increasing number of bear sightings outside the primary range. The data collection portal formalized collection of these observations into a monitoring tool, and also provided a means for citizens to view the distribution of sightings as they accumulated, and thus learn more about bear occurrences in the state. Black bears are an ideal species for citizen-scientist participation because they are a large-bodied, relatively easily-identifiable, iconic and charismatic species, which results in high levels of public participation, minimal species misidentification, and positive media attention for the project. Importantly, bears' tolerance for humans (including attraction to human-related food sources) enabled us to examine how several metrics of the human footprint influence detection of bears at moderately high levels of human presence.

Our objectives were to test how ALAN may influence the detection process of opportunistically-collected bear observations from citizen scientists. We compared how well spatially-explicit, monthly estimates of ALAN data explained variance in the detection process of bears within our occupancy models relative to factors that may be more commonly used (e.g., housing density) and assessed how ALAN impacted our results. We sought to understand whether quantification of ALAN, which has become a pervasive part of the modern global landscape (Kyba et al., 2017), helped to address observation bias. Properly accounting for observation bias is critical for fully realizing the potential benefits to ecological inference offered by citizen science projects.

Methods

Study area

Minnesota marks the westernmost edge of the eastern black bear population in the United States. Primary bear range in Minnesota matches the region of extensive forest cover in the north (~100,000 km^2), but bears also thrive along the edge of this primary forested range by exploiting agricultural crops (e.g. corn, sunflowers) and abundant wild fruits and nuts along edges of small, isolated patches of forest (Ditmer, Garshelis, Noyce, Haveles, & Fieberg, 2016). For example, in the mid-1990s, bears rapidly colonized the far northwestern corner of the state, a region that is over 50% agriculture and less than 20% forested, yet individuals living there are some of the physically largest and most fecund in the state due to abundant forage in the region (Ditmer, Noyce, Fieberg, & Garshelis, 2018). The total population is estimated at 12–15,000 bears, of which ~2,000 bears reside along the periphery of primary bear range, where the forest is much more fragmented (secondary bear range ~21,500 km^2; Garshelis & Tri, 2019; **Fig. 1**). Bear harvest occurs in the fall and a large portion of the primary bear range is regulated with a relatively conservative hunting quota system; the peripheral regions outside primary range are hunted more liberally, intended to prevent population increase and thus control the extent of bear damage to property or agricultural crops.

Natural landcover (forest, shrublands, wetlands) generally decreases moving farther from primary bear range. In the northern half of the state, human population density is sparse, and most roads have relatively low traffic volume. Along the edge of primary bear range, the landcover is a heterogenous composition of highly developed lands, with a high human population density in the Minneapolis-Saint Paul metropolitan region (population: >3.5 million), and extensive suburban developments in all directions, along with high-volume highways (e.g., Interstate 94) running approximately along the transition from primary to non-primary range. The southern half of the state is dominated by agricultural lands. Outside the primary bear range, swaths of forest occur in river corridors, which bears use as travelways (Ditmer et al., 2018; **Fig. 2**).

Our study area within Minnesota was restricted to a band within 55 km along the edge of (and including) secondary bear range (~115,000 km^2). This area contained > 90% of sites with bear observations and 95% of sites with recurring bear observations (bear observed in the same site more than once; **Fig. S1**; see Statistical analysis for more details). We did not aim to predict bear occurrence in regions with few observations of bears. Landcover within the designated study area rapidly changes from east to west, and beyond the western limit, there is little natural vegetation.

Figure 1. Map of Minnesota, USA, showing primary black bear range (where no citizen science observations were collected), secondary range (with more fragmented habitat and lower-density bear populations), the study region (including and within 55km of the secondary bear range), and all citizen scientist-collected observations, 2018-2019. Purple lines = counties containing the Minneapolis-Saint Paul metropolitan area.

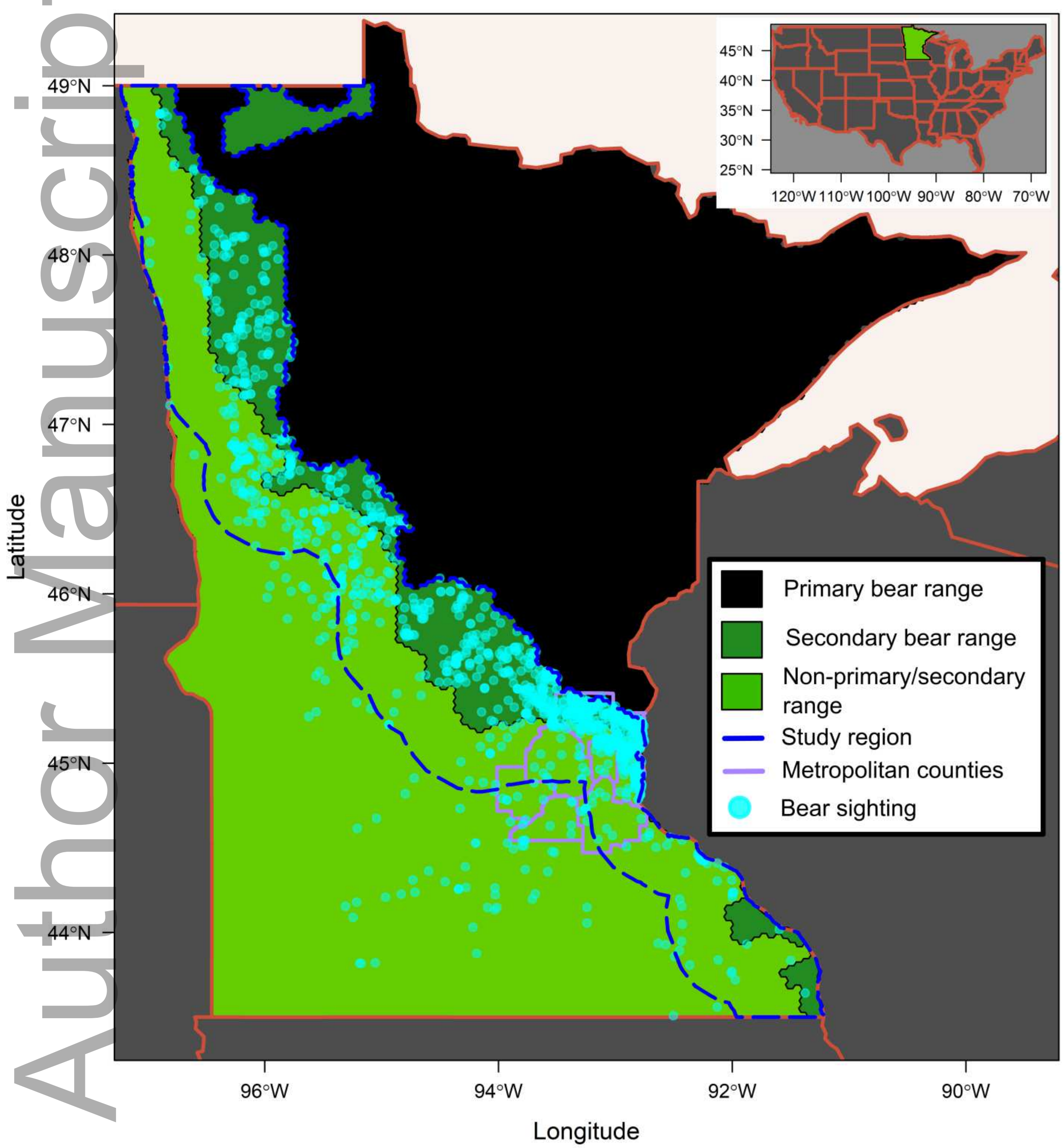
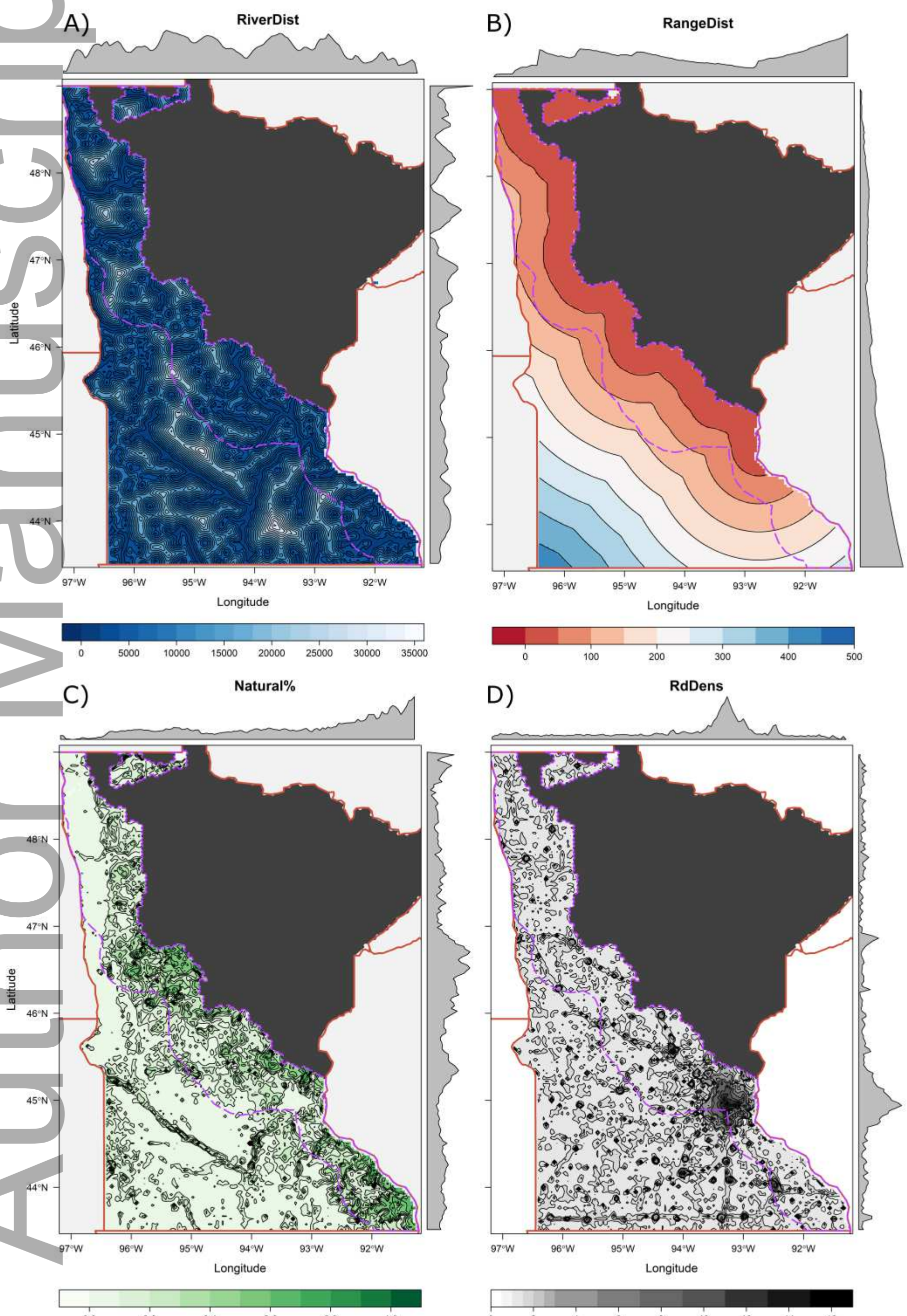


Figure 2. Maps of abundance covariates used in occupancy models of citizen scientist-collected observations of black bears outside of their primary range in Minnesota, USA, including: A) distance to nearest river (m), B) distance to primary bear range (km), C) natural landcover (% cover), and D) road

density (road length [km]/area [km^2]),. Each cell represents the mean value for a 25-km^2 area (the size of our sites). Histograms were created using mean values along the latitudinal and longitudinal axes of values within Minnesota not associated with primary bear range. Gray areas: primary bear range. Green lines: study area (<55 km from secondary bear range). The Minneapolis-Saint Paul metropolitan area is located between 44–45°N and 93–94°W.



Bear sighting collection

During 2018 - 2019, citizens were asked to report sightings of bears outside primary range (**Fig. 1**) by entering the location and answering questions about the bear's activity on a MNDNR-hosted website (https://www.dnr.state.mn.us/hunting/bear/bear-sightings.html). To facilitate accurate reporting, participants could enter a street address, coordinates, or click on the location using an interactive map (Survey 123 for ArcGIS, ESRI, Inc., Redlands, CA). If the participant entered an address or coordinates, the map automatically zoomed to the location for verification before submission. All bear sightings were publicly available for viewing on the website, except during bear baiting and hunting seasons (mid-August – mid-October). However, observations were collected during all months (rarely during November – March, when bears are generally hibernating). Observers entered their name and contact details to enable verification of unusual sightings, but we did not contact any observers, and all personal information was removed from the database before we began analysis.

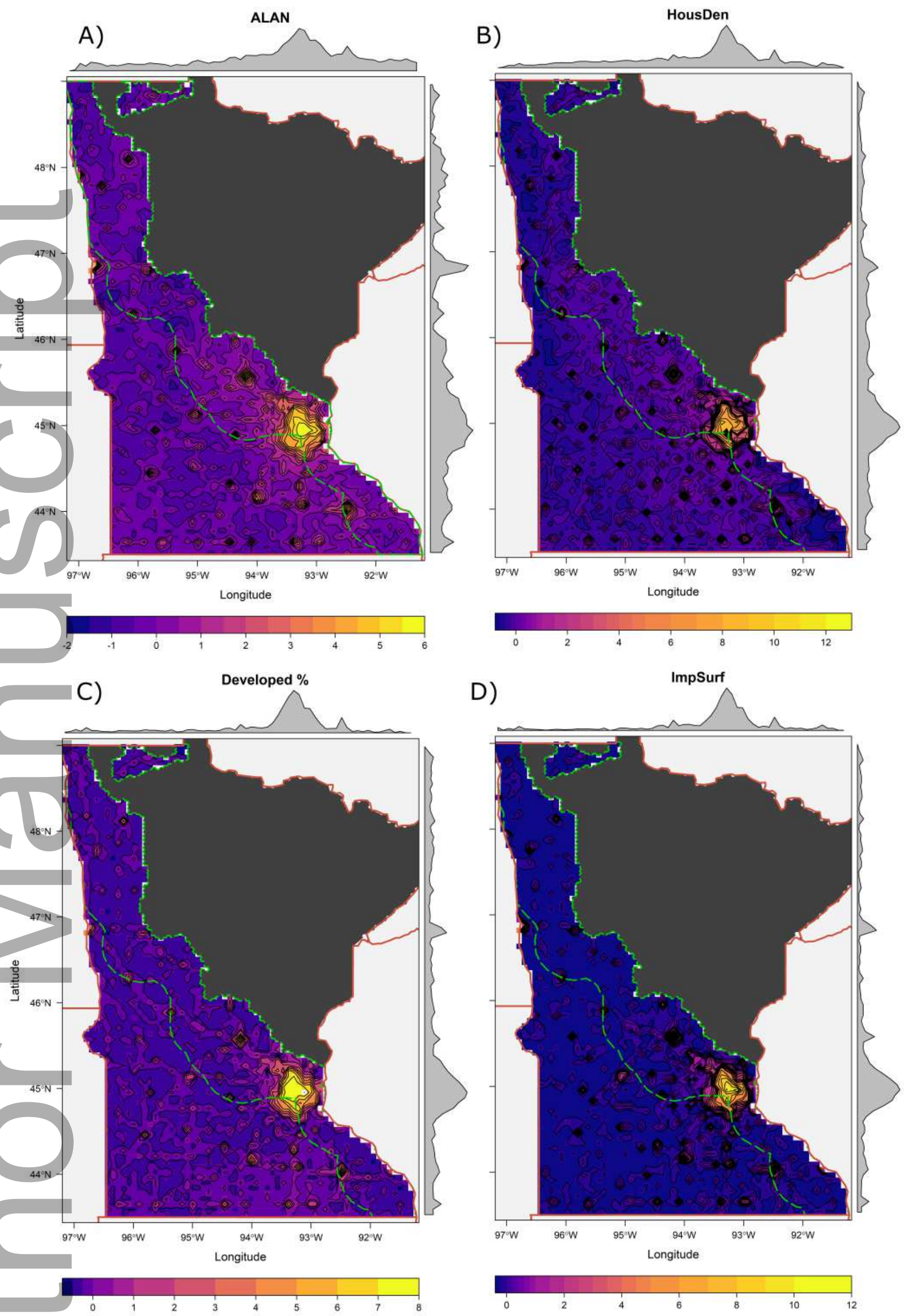
We were primarily interested in the number of sighting events (i.e., disregarding the number of bears reported in each sighting), and the date and location of each. However, we used the responses to other questions in each report, such as "What was the bear doing?" and "If the bear was eating, what was it eating?" to help ensure the response was valid. We excluded sightings that were within primary bear range or outside the state of Minnesota.

In 2019, we added two questions to the reporting website to get a better idea about the light conditions. The first question asked "What period of the day did you see the bear(s)?", and provided six options: 1) 01:00 – 05:00, 2) 05:00 – 09:00, 3) 09:00 – 13:00, 4) 13:00 – 17:00, 5) 17:00 – 21:00, and 6) 21:00 – 01:00.The second added question asked "What were the light conditions during the sighting?" with four options 1) Daylight, 2) Dawn or Dusk (low light), 3) Nighttime aided by artificial lights (streetlights, headlights, porch light, etc), 4) Nighttime with no artificial lighting.

Detection variables

We characterized five aspects of the human footprint outside primary bear range: 1) artificial light at night averaged over the months considered in the study (ALAN_ave.), 2) monthly ALAN estimates (ALAN_monthly), 3) housing density (HousDen), 4) developed land (Developed%), and 5) impervious surface (ImpSurf). These spatially explicit estimates were applied as covariates for the detection process of our occupancy model to account for biased detection and sampling efforts by citizen scientists observing bears **(Fig. 3)**. Estimates of nighttime radiance values were derived from data collected by NASA-NOAA's Suomi National Polar-Orbiting Partnership Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band. Data from the VIIRS sensor were lunar BRDF-corrected (bidirectional reflectance distribution function), and provided as 1 km^2 radiance values that remove the contributions of moonlight, clouds, terrain, wildfire, seasons, atmospheric effects, snow, and stray light, thus resulting in contributions of anthropogenic point source emissions only (Román et al., 2018). We used monthly ALAN estimates from the most current year available (2016). HousDen data were based on 2010 estimates at a 100 m^2 resolution (National Park Service, 2010). The Developed% layer was derived from the 2011 National Land Cover Database (NLCD) classification (U.S. Geological Survey, 2014). NLCD data are provided at a 30-m^2 resolution and we assigned a "1" for any "developed" classification (open-space – high intensity; class/value: 21 – 24), while all other landcover types were reclassified as "0". The ImpSurf estimates provide a percentage of impervious landcover (e.g. roads, energy production, urban areas) at a 100-m^2 resolution (Xian et al., 2011).

Figure 3. Maps of detection covariates used in occupancy models of citizen scientist-collected observations of black bears outside of their primary range in Minnesota, USA, including A) average artificial light at night, B) housing density, C) developed landcover, and D) impervious surface. We created mean scaled and centered values based on aggregated cells equal to four sites (100 km^2) for comparison and visualization. Each cell represents the mean value for a 25 km^2 area (the size of our sites). Histograms were created using mean values along the latitudinal and longitudinal axes of values within Minnesota not associated with primary bear range. Green lines: study area (<55 km from secondary bear range). The Minneapolis-Saint Paul metropolitan area is located between 44–45°N and 93–94°W.



Abundance/biological variables

We developed layers that estimated: 1) distance to nearest river (m; RiverDist), 2) distance to primary bear range (km; RangeDist), 3) road density (RdDens), and 4) percentage of area with natural cover (Natural%; **Fig. 2**). We hypothesized that each of these variables would help predict relative bear abundance within the non-primary bear range. Non-primary bear range was dominated by agricultural lands, and we believed bears would be more abundant in the areas that contained relatively high levels of natural cover despite their ability to persist in areas dominated by agriculture (Ditmer et al., 2018). We also hypothesized that bears would preferentially use riparian zones because of the associated natural cover, providing conduits for movement. We expected lower relative bear abundance farther from primary bear range (population source) and in areas with high road density. While bears do use lower traffic volume roads for movement and roadside forage, high road densities and traffic volume reduce cover and increase mortality risk (Brody & Pelton, 1989).

We created RiverDist using the National Hydrography Dataset (NHD; U.S. Geological Survey, 2015) obtained from the Minnesota Geospatial Commons (https://gisdata.mn.gov/dataset/water-national-hydrography-data). We used the shapefile for river features (NHDArea) and created a 30-m^2 resolution raster layer by calculating the distance to the nearest river from the centroid of each raster cell using the “Euclidian Distance” tool in the Spatial Analyst extension of ArcMap (v.10.6; ESRI, 2017). We used the same process to calculate the Euclidian distance to primary bear range. Road density estimates (1-km^2 resolution) were developed by the National Park Service (National Park Service Inventory and Monitoring Division - Modeling, Analysis, and Synthesis Group, 2014). To create the Natural%, we assigned any classifications from the NLCD raster layer associated with water, developed areas, barren areas or agriculture as “0”, and assigned landcover classifications associated with forest, shrubland, herbaceous, and wetlands (class/value: 40 – 74 & 90 – 95) as a “1”.

For all values associated with detection and abundance covariates, we calculated values that aligned with the same resolution of the ALAN data (1 km^2). If a given raster layer could not be aligned with the same extent of the ALAN data, we used the package ‘raster’ (Hijmans, 2019) in program R to convert each cell to a point (function ‘rasterToPoints’) based on its centroid and retained the value associated with each cell. For layers with binary data (0–1; Developed% and Natural%), we calculated the percentage of centroid points equal to 1, corresponding to if the cell was assigned a “0” or “1” as developed or natural, within each 1 km^2 cell associated with the ALAN data (i.e., % of each landcover category within ALAN raster cell). For all other layers, which were continuous values, we averaged the values of the points within each 1-km^2 cell and again associated it with the corresponding ALAN layer cell. Finally, we overlaid the locations of reported bear sightings, combining both years, onto the corresponding monthly ALAN raster layer to create monthly layers of bear sightings (BearSight). Over 97% of bear sightings occurred during April–October, when bears are not hibernating, so we only considered BearSight raster layers from those associated months.

Statistical analysis

The number of bear sightings depended both on the abundance of the species in the area and on factors affecting the detection process (Dénes, Silveira, & Beissinger, 2015). To assess which characterization of human footprint best described the detection process, we applied latent N-mixture models (Kéry, Royle, & Schmid, 2005; Royle, 2004) using the *pcount* function in the R-package 'unmarked' (Fiske & Chandler, 2011; vers.0.13-0) in program R (R Core Team, 2019).

The hierarchical structure of the N-mixture occupancy models explicitly accounts for imperfect detection and consists of two parts, one describing the ecological process determining the abundance of the species, and one describing the conditional detection process (Royle 2004). We fit a series of N-mixture models to our spatially replicated counts of bear observations and absences (no bears observed at the site in a given month) by altering the covariates describing the detection process with 1) Intercept only (NULL), 2) ALAN_ave., 3) ALAN_monthly, 4) Developed%, 5) HousDen, and 6) ImpSurf. However, we always included the same four covariates in the abundance portion of the model: RiverDist + RangeDist + RdDens + Natural%. All covariate values were scaled and centered for fit and comparison purposes. We tested for collinearity in our models using variance inflation factors via the "vif" function in the package 'unmarked'. The resulting variance inflation factors were all < 2, so we determined collinearity was not a problem (Dormann et al. 2013).

In order to better meet the closure assumption of occupancy models (Kéry & Royle 2016), we aggregated all spatial layers, for both detection and abundance, from 1 km^2 to 25 km^2. We refer to these 25 km^2 areas as "sites". We assumed that sites were large enough such that if a bear was detected there one month, it was occupied during all seven months (although individual bears may have used more than one site). In very sparsely-occupied parts of the state, where bears roam widely between distant food sources and adequate patches of habitat (Ditmer et al., 2018), this closure assumption might not hold, which is why we restricted our study to the region near the secondary range, where bear density was higher and food sources closer together. However, because of the potential for violation of the closure assumption, we interpret the estimates of our occupancy models to represent the total number of bears associated with a site during the overall period of study (referred to as 'relative abundance' in the Results and Discussion sections), rather than those permanently residing in the cell. Nevertheless, we tested the same models at 1-km^2 resolution (where closure was likely to be violated) and found similar results (**Tables S1 & S2)**.

We used the zero-inflated Poisson (ZIP) mixture to fit all the models, due to the instability of negative binomial mixture models applied to data with numerous zeros (Dénes et al., 2015; Knape et al., 2018). To assess fit we used QQ plots of site-sum randomized quantile residuals from the R package ‘nmixgof’ (Knape et al., 2018; **Fig. S2**). We determined that parameter estimates were stable at a K value (index of integration) of 200 (**Table S3**) and compared model fit among the different detection covariates using Akaike Information Criterion (AIC) values.

The residuals from our top model were autocorrelated based on spatial correlograms (by month) and Moran’s I. To account for the autocorrelation we used the package ‘spdep’ (Bivand et al. 2018) and created an autocovariate term (Crase et al. 2012) that was added to our top model (see **Table S4** for model values) prior to making predictions of relative bear abundance. We fit a variety of neighborhood radius distances (15 – 100 km) into our autocovariate variable and used AIC values to determine the distance that reduced variance the most (60 km). For both the detection and abundance component of the model, we applied the function *predict* to plot the effect of a specific variable while holding the others constant at their scaled mean. We created spatially-explicit predictions of relative bear abundance based on the top model, converting the values to a raster (function *rasterFromXYZ* in the ‘raster’ package).

Results

We received 1,081 reports of black bear sightings in 2018 and 811 in 2019 (2-year total = 1,892 sightings). After removal of invalid or unusable reports, and those outside of our defined study area, we retained 1,315 sightings for use in our analysis **(Fig. 1)**. Citizen-scientists observed bears most frequently during crepuscular hours (44% of 755 reports with the question included; 05:00 – 09:00 & 17:00 – 21:00; total 8 hours), and at nighttime (31%; 21:00 – 05:00; total 8 hours), while the diurnal period had the fewest observations (25%; 09:00 – 17:00). Of the 745 bear sighting reports that included a response to the question, “What were the lighting conditions during the sighting?”, 17% reported “nighttime aided by artificial lights”, 10%: nighttime with no artificial lights, 21%: dawn or dusk (low light), and 52% = daylight (many in the crepuscular period).

Bear Detection

Changing monthly estimates of ALAN best explained detection probability (AIC weight = 1.00) of citizen scientists encountering black bears relative to the null model and models including the other detection covariates **(Table 1; Fig. 4)**. ALAN_monthly had the greatest effect on detection within our occupancy models ($\grave{\beta}$ = 0.81, 0.71 – 0.90 95% CI) relative to ALAN_ave ($\grave{\beta}$ = 0.67, 0.57 – 0.76 95% CI), Developed% ($\grave{\beta}$ = 0.48, 0.34 – 0.61 95%CI), HousDens ($\grave{\beta}$ = 0.22, 0.03 – 0.40 95%CI) and ImpSurf ($\grave{\beta}$ = -0.20, -0.36 – -0.04 95%CI; **Fig. 4A**). Predicted detection probability increased from 0.05 (0.02 - 0.10 95% CI) at the lowest observed ALAN values to 29.1 (18.2 – 43.3 95% CI) in the most illuminated areas **(Fig. 4B)**. In contrast, the detection covariate Developed%, which was the most supported non-ALAN detection variable, was predicted to increase detection from 1.20 (0.57 – 2.11 95% CI) at its lowest observed values, to 16.6 (9.9 – 26.4 95% CI) at its largest.

Ecological Relationships

Based on the best-fitting model, which included ALAN_monthly as the covariate in the detection process, relative bear abundance increased with greater percentages of natural landcover (non-urban, crop, or barren), proximity to primary bear range and riparian areas, although RiverDist had a 95% confidence interval overlapping zero in our best fitting model **(Table 2; Fig. 2)**. RdDens had a negative relationship with relative bear abundance in models containing ALAN in the detection component, but a positive relationship with relative bear abundance in the Null model and models including HousDen and ImpSurf in the detection process **(Table 2; Fig. 2D)**.

A large percentage of the bear sighting reports were from the suburban section north of the Minneapolis-Saint Paul metropolitan area **(Fig. 1 & Fig. 5A)**. However, detection in this area was high, so predicted abundance of bears was lower than indicated by the large number of sightings. Likewise, predicted bear abundance was higher in northwestern Minnesota, where sightings were fewer, but detection was also far lower **(Fig. 5B & 5C)**. This area in the northwest has low levels of ALAN, low-to-medium road density, and is one of the few regions close to primary bear range with a large percentage of natural land cover **(Fig. 2)**, all characteristics that favor the establishment of bears. The monthly ALAN model, corrected for bias in detection and autocorrelation of the residuals, predicted an expected relative abundance for this area up to ~375% higher than indicated by the number of reported sightings.

Table 1. Ranking of occupancy models of citizen scientist-collected observations of black bears outside of their primary range in Minnesota, USA. All models contained the same covariates for bear abundance, but each contained a different explanatory variable for the detection process. nPars: number of parameters; AIC: Akaike Information Criterion; ΔAIC: AIC relative to top-ranked model; AICwt: model weight

Model	nPars	AIC	ΔAIC	AICwt
p(ALAN_monthly)	8	9115.7	0.0	1.00
p(ALAN_average)	8	9208.7	93.0	0.00
p(Developed%)	8	9340.4	224.7	0.00
p(HousDen)	8	9372.4	256.8	0.00
p(ImpSurf).	8	9373.7	258.1	0.00

p(Null)	7	9377.2	261.6	0.00

Figure 4. A) Mean coefficient and 95% confidence intervals of scaled and centered detection covariates from occupancy models of citizen scientist-collected observations of black bears outside of their primary range in Minnesota, USA. **B)** Predicted detection probability of monthly ALAN values from our top model (values were centered and scaled). All other variables within the occupancy model were at their mean values.

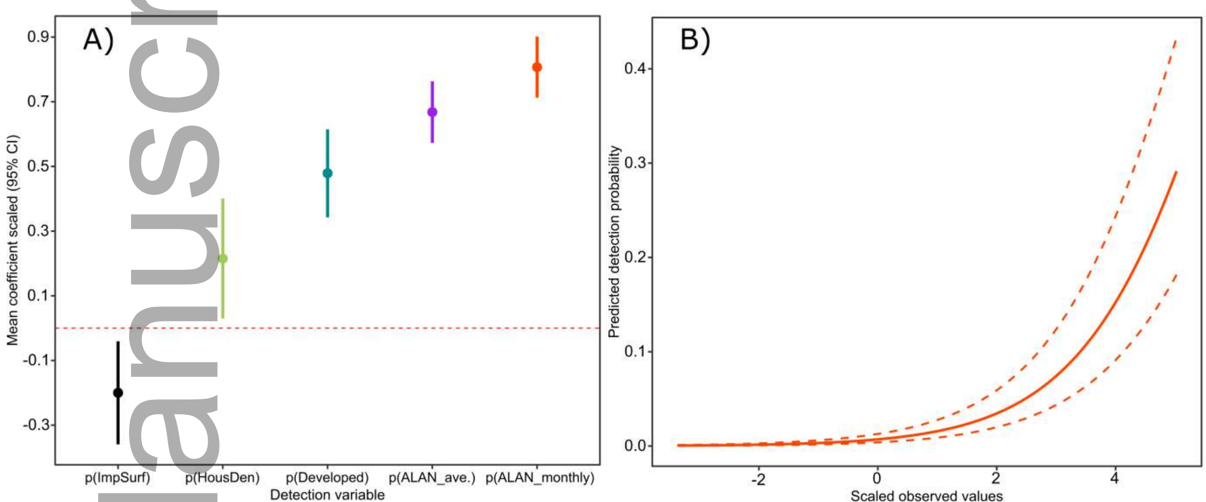
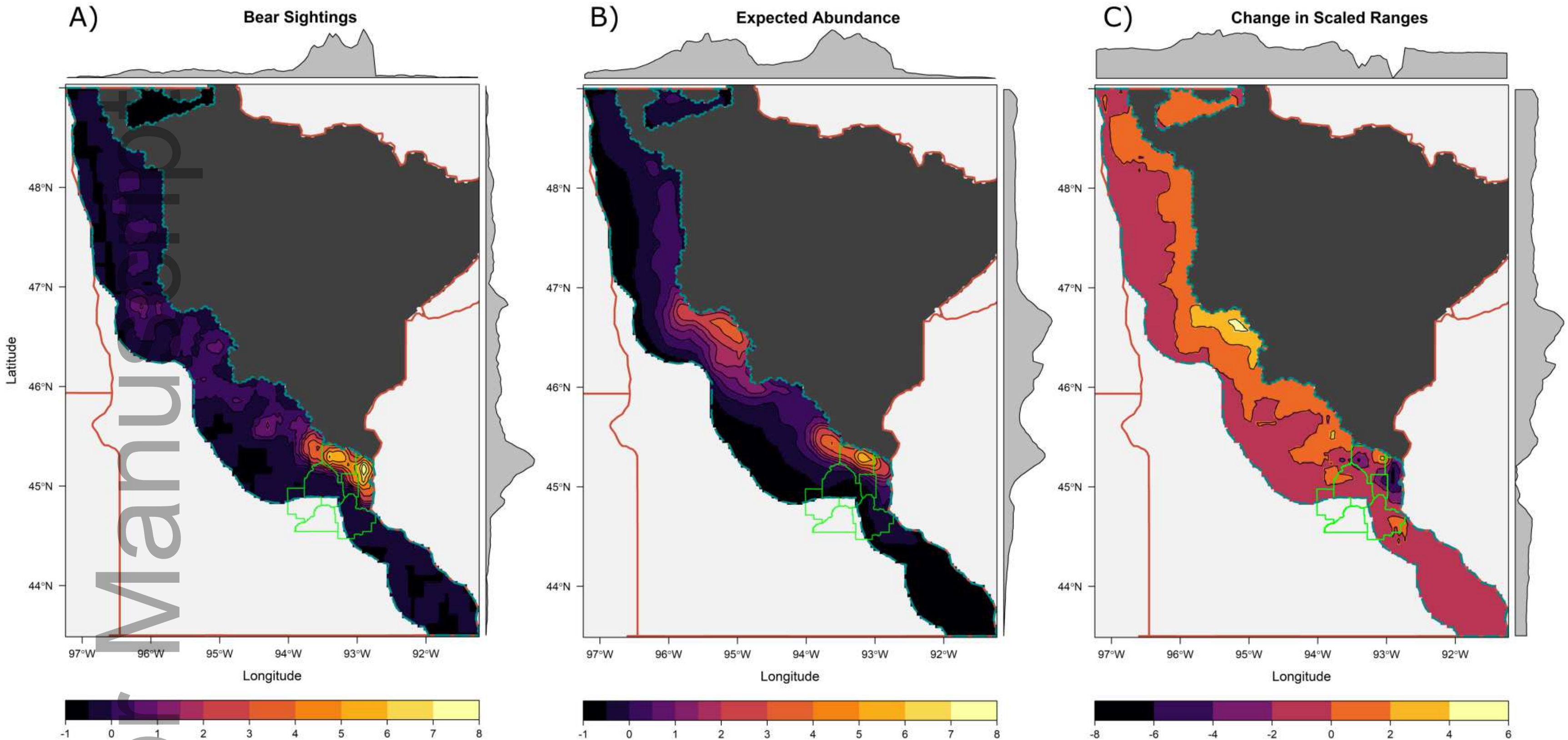


Table 2. Parameter estimates and standard errors (in parenthesis) from occupancy models fit to citizen scientist-collected observations of black bears outside of their primary range in Minnesota, USA. All models contained the same covariates for bear abundance, but each contained a different explanatory variable for the detection process. p, ψ, λ: influence on detection, occupancy and abundance, respectively.

Model	p: Intercept	ψ	λ: Intercept	Λ: RiverDist	Λ: RdDens	Λ: RangeDist	Λ: Natural%
p(ALAN_monthly)	-4.94 (0.30)	0.32 (0.08)	1.87 (0.31)	-0.05 (0.04)	-0.24 (0.05)	-0.97 (0.06)	0.55 (0.03)
p(ALAN_average)	-4.75 (0.30)	0.34 (0.08)	1.74 (0.31)	-0.06 (0.04)	-0.18 (0.05)	-0.98 (0.06)	0.55 (0.03)
p(Developed%)	-4.17 (0.27)	0.5 (0.08)	1.31 (0.28)	-0.12 (0.04)	-0.07 (0.07)	-0.99 (0.06)	0.59 (0.03)
p(HousDen)	-3.69 (0.20)	0.49 (0.08)	0.86 (0.21)	-0.14 (0.04)	0.14 (0.07)	-1.00 (0.07)	0.56 (0.03)
p(ImpSurf).	-3.59 (0.20)	0.48 (0.08)	0.77 (0.22)	-0.14 (0.04)	0.48 (0.08)	-0.98 (0.07)	0.55 (0.03)
p(Null)	-3.62 (0.21)	0.50 (0.08)	0.81 (0.22)	-0.14 (0.04)	0.30 (0.03)	-1.00 (0.07)	0.56 (0.03)

Figure 5. A) Centered and scaled counts of bears sighted outside of their primary range by citizen scientists. **B)** Scaled predicted expected relative abundance of bears based on our spatial autocorrelation-corrected best-fitting occupancy model which included monthly estimates of ALAN as the covariate in the detection portion of the model. **C)** The scaled difference between panel A and panel B. All raster cells ($5km^2$) were smoothed using a 7×7 moving window [function focal in package raster] to enhance visualization. Histograms were created using mean values along the latitudinal and longitudinal axes of values within Minnesota not associated with primary bear range. Green lines = counties containing the Minneapolis-Saint Paul metropolitan area. Gray areas: primary bear range.

429


A)
Bear Sightings
B)
Expected Abundance
C)
Change in Scaled Ranges
Latitude
Longitude
48°N
47°N
46°N
45°N
44°N
97°W
96°W
95°W
94°W
93°W
92°W
-1
0
1
2
3
4
5
6
7
8
-8
-6
-4
-2
0
2
4
6

Discussion

Citizen scientists have become an integral and powerful aspect of many ecological research and monitoring projects, yet due to the opportunistic nature of data collection, spatial biases in sampling arise. These must be accounted for in order to make accurate inferences from the data. We demonstrated that spatially-explicit estimates of ALAN, a growing environmental pollutant strongly correlated to human development and activities (Gaston et al., 2013), is a powerful source of data for reducing sampling bias driven by detection heterogeneity. Elevated ALAN radiance was associated with a greater detection probability among citizen scientists participating in an effort to assess range expansion of black bears across a large area (>115,000 km^2) including many privately-owned lands. In our occupancy models, ALAN provided the best proxy that combined presence of citizens with their ability to see bears; ALAN not only directly aided at least ~17% of bear sightings, but also explained the distribution of potential observers better than other surrogates (e.g., housing density). Accounting for ALAN reduced sampling biases, and improved predictions related to associations between ecological factors and animal presence, which in turn created more accurate and biologically-realistic predictions of species' relative abundance at a broad spatial scale.

Using opportunistic observations for monitoring population expansion has the disadvantage that animals can be seen only where people are present and sighting conditions are favorable for detection. The intent of modelling detectability using ALAN is to account for this inherent bias, so clusters of observations, or blank spots where observations are scarce or absent, can be compared even if levels of detectability by people are different. For example, we found relatively dark areas that, based on landscape characteristics, likely had a higher presence of bears than indicated by the observational data. Instead of including precise measurements of human presence, some studies have suggested ways to improve the data collection protocols for citizen scientist projects to strengthen inference (Altwegg & Nichols, 2019), such as accounting for completeness and individual ability to identify species (Kelling et al., 2015). However, for projects like ours that simply extend requests for participation to the public, keeping the process simple was key to maximizing the level of participation and number of reported bear sightings.

As expected, the ecological factors associated with the expansion of this bear population were low road density, high natural landcover, and proximity to riparian areas. American black bears are known to be relatively human-tolerant, opportunistic omnivores, and throughout their range, they are colonizing or recolonizing areas that were once assumed to contain too little natural habitat to support bears (Scheick & McCown, 2014). Bears may be enticed to leave their primary range and seek out new areas in search of mating opportunities or caloric hotspots (Noyce & Garshelis, 2010), such as garbage or birdfeeders in more developed settings (Merkle, Robinson, Krausman, & Alaback, 2013), or crops in agricultural areas (Ditmer et al., 2016). In northwestern Minnesota, Ditmer et al., (2018) found that male bears moved westward in late summer and fall into areas with very little forest cover to exploit crops such as corn. However, they required some forest cover near the feeding site, and typically returned to areas with more cover to den and feed the following year before crops again ripened. Female bears are more reluctant to venture far from forest, but green corridors along rivers may provide avenues for their expansion.

Potential Applications and Caveats for use of ALAN in Citizen Science-focused Projects

The ability to detect and monitor species' range shifts, contractions, or expansions is increasingly important due to rapid changes in climate (Chen, Hill, Ohlemüller, Roy, & Thomas, 2011), land use (Jetz, Wilcove, & Dobson, 2007), and human tolerance for species that share the landscape (Carter & Linnell, 2016). Currently, several large carnivore species are colonizing, or re-colonizing large regions in North America and Europe (Chapron et al., 2014). These species typically occur at relatively low densities, are extremely vagile, and require intensive monitoring to manage for human-wildlife conflicts. Our approach may be useful for monitoring changes in range and anticipating potential conflict hotspots. For polarizing species, connecting professionals with the public through a citizen science program enhances two-way information exchange, which is likely to enable more potential to mitigate potential conflict. Participation in a project can increase the public's receptiveness to management and conservation actions because participants have been part of the research process (Backstrand, 2003; Dvornich, Tudor, & Grue, 1995). Weckel, Mack, Nagy, Christie, & Wincorn (2010) found that surveying the public about their feelings of risk amidst increasing human-coyote (*Canis latrans*) interactions in suburban New York City, USA provided a low-cost tool for reducing conflict via outreach, modifying behavior, and improving understanding of coyote space use.

Previous studies have accounted for the sampling bias in opportunistically collected data through a variety of ways, such as changes in detection across time (Kéry & Schmid, 2004), observer effort (Mair & Ruete, 2016), spatial correlation of observations (Clement, Hines, Nichols, Pardieck, & Ziolkowski, 2016), habitat factors (Paolino et al., 2018), and spatial estimates of human presence. However, the spatial metrics used in these studies are static, rarely updated, and often do not reflect temporal trends such as seasonal traffic volumes. The same was true for the NLCD data we used to derive % developed surface within each site

(NLCD layers updates ~ every 5 years; 2001, 2006, 2011, 2016). Our model with monthly estimates of ALAN explained variability in the detection process better than the average ALAN composite, further highlighting that capturing temporal changes in the human footprint can also improve inference. ALAN is collected daily at ~1-km^2 resolution, and while cloud cover presents challenges with data loss, as with any other remotely-sensed product, its strength is the ability to detect changes through time (currently only available as a monthly composite). Combining a dynamic occupancy modelling framework with ALAN generation estimates at fine temporal scales could be especially informative for studies in regions undergoing rapid changes, locations in extreme latitudes where ALAN is prevalent for longer periods of most daily cycles for part of the year, and in areas that experience drastic fluctuations of human population density (e.g. National Parks in summer months) or are hotspots for human-wildlife conflict.

Wildlife species often alter their activity patterns towards crepuscular and nocturnal periods in areas with high human activity or urbanization (Gaynor, Hojnowski, Carter, & Brashares, 2018). These areas are illuminated by ALAN, which blurs the lines between day and night (Hölker, Wolter, Perkin, & Tockner, 2010), and makes species that would have been previously unobservable more available for detection. However, it is not apparent how ALAN impacts the behavior (specifically movement/space use) or distribution of most wildlife species (i.e., ALAN may alter abundance patterns). A growing body of research is documenting the ways that ALAN can disrupt species (Hölker et al., 2010); however, these studies have been primarily conducted at fine scales, or in laboratory settings, with nearly all considering smaller-bodied and less-vagile species. Although some species, such as insectivorous bats, may aggregate at light sources to forage (Jung & Kalko, 2010), others might avoid highly illuminated areas (Bliss-Ketchum, de Rivera, Turner, & Weisbaum, 2016). In the case of black bears, the species is known to be attracted to human-related food sources, and may thrive in areas with high human density, but they typically alter their movements and activities so as to reduce encounters with people (Beckmann & Berger 2003, Evans et al. 2017, Zeller et al. 2019). Many other species are not as tolerant of human presence or activities, or not as adaptable, so ALAN may reduce their use of an area.

We also caution that ALAN may not always be closely related to human presence. In most cases, modern human activities and presence are strongly linked in developed parts of the world with features such as street lights, residential lighting and headlights from vehicles. However, in some regions, economic activities may generate large amounts of ALAN without associated increased detection probability (e.g. industrial sites), thus reducing the spatial correlation between ALAN and human presence. Because of this, researchers may want to model observation bias with ALAN (where appropriate) along with a suite of covariates that are carefully chosen for the terrain, specific human activities, and other potentially important factors, such as Mair & Ruete (2016) who constructed "ignorance" scores which quantified overall observation bias of citizen-scientist collected data throughout Sweden.

Although most established citizen science projects that involve data collection are aimed at bird species, mammal-focused projects are increasing (Massimino et al., 2018). Because many mammals are less easily detected than birds, the inherent sampling and detection bias of the observations requires proper accounting for the dynamic human footprint in order to make biologically-sound inference. More wildlife studies are using remotely-sensed products to capture ecological changes at fine spatial and temporal scales (e.g. forage in the form of vegetative greenness). Here, we highlight that remotely-sensed ALAN data can play a similar role in capturing an accurate snapshot of the human footprint at fine temporal and spatial scales. In the future, we expect ALAN data to continue to increase in resolution and for NASA's filtering technologies to be further refined, thus providing more accurate delineations of the human footprint. Incorporating citizen-scientists into ecological research has been linked with numerous practical, social, and scientific advantages; within this framework, we need to account for the fact that most data will invariably be collected from areas that are the most accessible to the public, or their full potential cannot be realized.

Acknowledgements

We thank Bruce Anderson of the Minnesota Department of Natural Resources for technological support on the survey website. The work was supported by the NASA Ecological Forecasting Grant NNX17AG36G to NHC.

Authors' Contributions

MD conceived the ideas of integrating ALAN data with the citizen science survey developed by DG and AT. NC and FI provided insights on the analysis and interpretation throughout the process. All authors contributed toward writing and editing.

Data Availability

Data available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.59zw3r25c (Ditmer et al. 2020).

References

Altwegg, R., & Nichols, J. D. (2019). Occupancy models for citizen-science data. *Methods in Ecology and Evolution*, *10*(1), 8–21. doi: 10.1111/2041-210X.13090

Backstrand, K. (2003). Civic science for sustainability: Reframing the role of experts, policy-makers and citizens in environmental governance. *Global Environmental Politics*, *3*(4), 24–41.

Beckmann, J. P., & Berger, J. (2003). Rapid ecological and behavioural changes in carnivores: the responses of black bears (*Ursus americanus*) to altered food. Journal of Zoology, 261(2), 207–212.

Bivand, R. S., & Wong, D. W. S. (2018). Comparing implementations of global and local indicators of spatial association. *TEST*, *27*(3), 716–748. doi: 10.1007/s11749-018-0599-x

Bliss-Ketchum, L. L., de Rivera, C. E., Turner, B. C., & Weisbaum, D. M. (2016). The effect of artificial light on wildlife use of a passage structure. *Biological Conservation*, *199*, 25–28. doi: 10.1016/j.biocon.2016.04.025

Bonney, R., Cooper, C. B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K. V., & Shirk, J. (2009). Citizen science: A developing tool for expanding science knowledge and scientific literacy. *BioScience*, *59*(11), 977–984. doi: 10.1525/bio.2009.59.11.9

Brody, A. J., & Pelton, M. R. (1989). Effects of roads on black bear movements in western North Carolina. *Wildlife Society Bulletin*, *17*(1), 5–10.

Carter, N. H., & Linnell, J. D. C. (2016). Co-adaptation is key to coexisting with large carnivores. *Trends in Ecology & Evolution*, *31*(8), 575–578. doi: 10.1016/j.tree.2016.05.006

Chapron, G., Kaczensky, P., Linnell, J. D. C., Arx, M. von, Huber, D., Andrén, H., … Boitani, L. (2014). Recovery of large carnivores in Europe’s modern human-dominated landscapes. *Science*, 346(6216), 1517–1519. doi: 10.1126/science.1257553

Chen, I.-C., Hill, J. K., Ohlemüller, R., Roy, D. B., & Thomas, C. D. (2011). Rapid range shifts of species associated with high levels of climate warming. *Science*, *333*(6045), 1024–1026. doi: 10.1126/science.1206432

Clement, M. J., Hines, J. E., Nichols, J. D., Pardieck, K. L., & Ziolkowski, D. J. (2016). Estimating indices of range shifts in birds using dynamic models when detection is imperfect. *Global Change Biology*, *22*(10), 3273–3285. doi: 10.1111/gcb.13283

Crase, B., Liedloff, A.C. & Wintle, B.A. (2012). A new method for dealing with residual spatial autocorrelation in species distribution models. *Ecography*, 35, 879–888.

Dénes, F. V., Silveira, L. F., & Beissinger, S. R. (2015). Estimating abundance of unmarked animal populations: accounting for imperfect detection and other sources of zero inflation. *Methods in Ecology and Evolution*, *6*(5), 543–556. doi: 10.1111/2041-210X.12333

Dickinson, J. L., Shirk, J., Bonter, D., Bonney, R., Crain, R. L., Martin, J., … Purcell, K. (2012). The current state of citizen science as a tool for ecological research and public engagement. *Frontiers in Ecology and the Environment*, *10*(6), 291–297.

Dickinson, J. L., Zuckerberg, B., & Bonter, D. N. (2010). Citizen science as an ecological research tool: Challenges and benefits. *Annual Review of Ecology, Evolution, and Systematics*, *41*(1), 149–172. doi: 10.1146/annurev-ecolsys-102209-144636

Ditmer, M. A., Garshelis, D. L., Noyce, K. V., Haveles, A. W., & Fieberg, J. R. (2016). Are American black bears in an agricultural landscape being sustained by crops? *Journal of Mammalogy*, *97*(1), 54–67. doi: 10.1093/jmammal/gyv153

Ditmer, M. D., Iannarilli, F., Tri, A. N., Garshelis, D. L., & Carter, N. H. (2020). Data from: Artificial night light helps account for observer bias in citizen science monitoring of an expanding large mammal population. Dryad Digital Repository, https://doi.org/10.5061/dryad.59zw3r25c

Ditmer, M. A., Noyce, K. V., Fieberg, J. R., & Garshelis, D. L. (2018). Delineating the ecological and geographic edge of an opportunist: The American black bear exploiting an agricultural landscape. *Ecological Modelling*, *387*, 205–219. doi: 10.1016/j.ecolmodel.2018.08.018

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., … Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*, 27–46.

Dvornich, K. M., Tudor, M., & Grue, C. E. (1995). "NatureMapping": Assisting management of natural resources through public education and public participation. *Wildlife Society Bulletin (1973-2006)*, *23*(4), 609–614.

ESRI. (2017). *ArcGIS Desktop: Release 10.6*. Redlands, CA: Environmental Systems Research Institute.

Evans, M. J., Hawley, J. E., Rego, P. W., & Rittenhouse, T. A. G. (2014). Exurban land use facilitates human-black bear conflicts. *The Journal of Wildlife Management*, *78*(8), 1477–1485. doi: 10.1002/jwmg.796

Evans, M. J., Rittenhouse, T. A. G., Hawley, J. E., & Rego, P. W. (2017). Black bear recolonization patterns in a human-dominated landscape vary based on housing: New insights from spatially explicit density models. *Landscape and Urban Planning*, *162*, 13–24. doi: 10.1016/j.landurbplan.2017.01.009

Fiske, I., & Chandler, R. (2011). unmarked: An R package for fitting hierarchical models of wildlife occurrence and abundance. *Journal of Statistical Software*, *43*(1), 1–23. doi: 10.18637/jss.v043.i10

Follett, R., & Strezov, V. (2015). An analysis of citizen science based research: Usage and publication patterns. *PLOS ONE*, *10*(11), e0143687. doi: 10.1371/journal.pone.0143687

Garrett, J. K., Donald, P. F., & Gaston, K. J. (2019). Skyglow extends into the world's Key Biodiversity Areas. *Animal Conservation*, 0. doi: 10.1111/acv.12480

Garshelis, D.L., & Tri, A.N. (2019). *Status of Minnesota black bears 2018: harvests, complaints, foods, population trends, and hunter survey*. Retrieved from Minnesota Department of Natural Resources website: https://files.dnr.state.mn.us/recreation/hunting/bear/2018_bearharvest.pdf

Gaston, K. J. (2018). Lighting up the nighttime. *Science*, *362*(6416), 744–746. doi: 10.1126/science.aau8226

Gaston, K. J., Bennie, J., Davies, T. W., & Hopkins, J. (2013). The ecological impacts of nighttime light pollution: a mechanistic appraisal. *Biological Reviews*, *88*(4), 912–927. doi: 10.1111/brv.12036

Gaynor, K. M., Hojnowski, C. E., Carter, N. H., & Brashares, J. S. (2018). The influence of human disturbance on wildlife nocturnality. *Science*, *360*(6394), 1232–1235. doi: 10.1126/science.aar7121

Hijmans, R. J. (2019). *raster: Geographic data analysis and modeling. R package version 3.0-7.* Retrieved from https://CRAN.R-project.org/package=raster

Hölker, F., Wolter, C., Perkin, E. K., & Tockner, K. (2010). Light pollution as a biodiversity threat. *Trends in Ecology & Evolution*, *25*(12), 681–682. doi: 10.1016/j.tree.2010.09.007

Hugo, S., & Altwegg, R. (2017). The second Southern African Bird Atlas Project: Causes and consequences of geographical sampling bias. *Ecology and Evolution*, *7*(17), 6839–6849. doi: 10.1002/ece3.3228

Isaac, N. J. B., Strien, A. J. van, August, T. A., Zeeuw, M. P. de, & Roy, D. B. (2014). Statistics for citizen science: extracting signals of change from noisy ecological data. *Methods in Ecology and Evolution*, *5*(10), 1052–1060. doi: 10.1111/2041-210X.12254

Jetz, W., Wilcove, D. S., & Dobson, A. P. (2007). Projected impacts of climate and land-Use change on the global diversity of birds. *PLOS Biology*, *5*(6), e157. doi: 10.1371/journal.pbio.0050157

Jung, K., & Kalko, E. K. V. (2010). Where forest meets urbanization: Foraging plasticity of aerial insectivorous bats in an anthropogenically altered environment. *Journal of Mammalogy*, *91*(1), 144–153.

Kelling, S., Fink, D., La Sorte, F. A., Johnston, A., Bruns, N. E., & Hochachka, W. M. (2015). Taking a 'Big Data' approach to data quality in a citizen science project. *Ambio*, *44*(4), 601–611. doi: 10.1007/s13280-015-0710-4

Kéry, M., & Royle, J. A. (2016). Chapter 6 - Modeling abundance with counts of unmarked individuals in closed populations: Binomial N-mixture models. In M. Kéry & J. A. Royle (Eds.), *Applied Hierarchical Modeling in Ecology* (pp. 219–312). doi: 10.1016/B978-0-12-801378-6.00006-0

Kéry, M., Royle, J. A., Schmid, H., Schaub, M., Volet, B., Häfliger, G., & Zbinden, N. (2010). Site-occupancy distribution modeling to correct population-trend estimates derived from opportunistic observations. *Conservation Biology*, *24*(5), 1388–1397. doi: 10.1111/j.1523-1739.2010.01479.x

Kéry, M., & Schmid, H. (2004). Monitoring programs need to take into account imperfect species detectability. *Basic and Applied Ecology*, *5*(1), 65–73. doi: 10.1078/1439-1791-00194

Knape, J., Arlt, D., Barraquand, F., Berg, Å., Chevalier, M., Pärt, T., … Żmihorski, M. (2018). Sensitivity of binomial N-mixture models to overdispersion: The importance of assessing model fit. *Methods in Ecology and Evolution*, *9*(10), 2102–2114. doi: 10.1111/2041-210X.13062

Kyba, C. C. M., Kuester, T., Sánchez de Miguel, A., Baugh, K., Jechow, A., Hölker, F., … Guanter, L. (2017). Artificially lit surface of Earth at night increasing in radiance and extent. *Science Advances*, *3*(11). doi: 10.1126/sciadv.1701528

MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L., & Hines, J. E. (2017). *Occupancy estimation and modeling: Inferring patterns and dynamics of species occurrence*. Elsevier.

Mair, L., & Ruete, A. (2016). Explaining spatial variation in the recording effort of citizen science data across multiple taxa. *PLOS ONE*, *11*(1), e0147796. doi: 10.1371/journal.pone.0147796

Massimino, D., Harris, S. J., & Gillings, S. (2018). Evaluating spatiotemporal trends in terrestrial mammal abundance using data collected during bird surveys. *Biological Conservation*, *226*, 153–167. doi: 10.1016/j.biocon.2018.07.026

Merkle, J. A., Robinson, H. S., Krausman, P. R., & Alaback, P. (2013). Food availability and foraging near human developments by black bears. *Journal of Mammalogy*, *94*(2), 378–385. doi: 10.1644/12-MAMM-A-002.1

National Park Service. (2010). *NPScape housing measure – Phase 1 metrics processing SOP: Current housing density, historic housing density, and projected housing density metrics* (No. Natural Resource Report. NPS/NRPC/IMD/NRR—2010/251. Published Report-2165448.). Fort Collins, Colorado: National Park Service, Natural Resource Program Center.

National Park Service Inventory and Monitoring Division - Modeling, Analysis, and Synthesis Group. (2014). *Road density: Explanatory Variable for Soundscape Model Training - CONUS*. Retrieved from https://irma.nps.gov/DataStore/Reference/Profile/2217356

Newman, G., Chandler, M., Clyde, M., McGreavy, B., Haklay, M., Ballard, H., … Gallo, J. (2017). Leveraging the power of place in citizen science for effective conservation decision making. *Biological Conservation*, *208*, 55–64. doi: 10.1016/j.biocon.2016.07.019

Noyce, K. V., & Garshelis, D. L. (2010). Seasonal migrations of black bears (Ursus americanus): causes and consequences. *Behavioral Ecology and Sociobiology*, *65*(4), 823–835. doi: 10.1007/s00265-010-1086-x

Paolino, R. M., Royle, J. A., Versiani, N. F., Rodrigues, T. F., Pasqualotto, N., Krepschi, V. G., & Chiarello, A. G. (2018). Importance of riparian forest corridors for the ocelot in agricultural landscapes. *Journal of Mammalogy*, *99*(4), 874–884. doi: 10.1093/jmammal/gyy075

R Core Team. (2019). *R: A language and environment for statistical computing.* Retrieved from https://www.R-project.org/

Reddy, S., & Dávalos, L. M. (2003). Special paper: Geographical sampling bias and Its implications for conservation priorities in Africa. *Journal of Biogeography*, *30*(11), 1719–1727.

Román, M. O., Wang, Z., Sun, Q., Kalb, V., Miller, S. D., Molthan, A., … Masuoka, E. J. (2018). NASA's Black Marble nighttime lights product suite. *Remote Sensing of Environment*, *210*, 113–143. doi: 10.1016/j.rse.2018.03.017

Scheick, B. K., & McCown, W. (2014). Geographic distribution of American black bears in North America. *Ursus*, 24–33. doi: 10.2192/URSUS-D-12-00020.1

Silvertown, J. (2009). A new dawn for citizen science. *Trends in Ecology & Evolution*, *24*(9), 467–471. doi: 10.1016/j.tree.2009.03.017

Simons, T. R., Alldredge, M. W., Pollock, K. H., & Wettroth, J. M. (2007). Experimental analysis of the auditory detection process on avian point counts.*The Auk*, *124*(3), 986–999.

Strien, A. J. van, Swaay, C. A. M. van, & Termaat, T. (2013). Opportunistic citizen science data of animal species produce reliable estimates of distribution trends if analysed with occupancy models. *Journal of Applied Ecology*, *50*(6), 1450–1458. doi: 10.1111/1365-2664.12158

Sun, C. C., Royle, J. A., & Fuller, A. K. (2019). Incorporating citizen science data in spatially explicit integrated population models. *Ecology*, *100*(9), e02777. doi: 10.1002/ecy.2777

Swanson, A., Kosmala, M., Lintott, C., & Packer, C. (2016). A generalized approach for producing, quantifying, and validating citizen science data from wildlife images. *Conservation Biology*, *30*(3), 520–531. doi: 10.1111/cobi.12695

Tri, A. N., Edwards, J. W., Strager, M. P., Petty, J. T., Ryan, C. W., Carpenter, C. P., … Carr, P. C. (2016). Habitat use by American black bears in the urban–wildland interface of the Mid-Atlantic, USA. *Ursus*, *27*(1), 45–56. doi: 10.2192/URSUS-D-15-00007.1

Tulloch, A. I. T., Possingham, H. P., Joseph, L. N., Szabo, J., & Martin, T. G. (2013). Realising the full potential of citizen science monitoring programs. *Biological Conservation*, *165*, 128–138. doi: 10.1016/j.biocon.2013.05.025

U.S. Geological Survey. (2014). NLCD 2011 Land Cover (2011 Edition, amended 2014) - National Geospatial Data Asset (NGDA) Land Use Land Cover (Version 2011 (2014 amended)). Sioux Falls, SD.

U.S. Geological Survey. (2015). *National Hydrography Dataset - High-Resolution*. Denver, CO.

Warton, D. I., Renner, I. W., & Ramp, D. (2013). Model-based control of observer bias for the analysis of presence-only data in ecology. *PLOS ONE*, *8*(11), e79168. doi: 10.1371/journal.pone.0079168

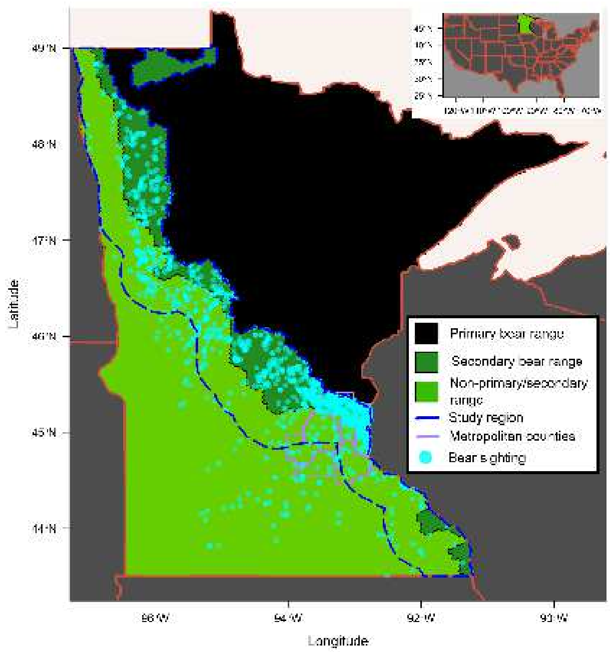
Weckel, M. E., Mack, D., Nagy, C., Christie, R., & Wincorn, A. (2010). Using citizen science to map human-coyote interaction in suburban New York, USA. *The Journal of Wildlife Management*, *74*(5), 1163–1171.

Wilson, S., Anderson, E. M., Wilson, A. S. G., Bertram, D. F., & Arcese, P. (2013). Citizen science reveals an extensive shift in the winter distribution of migratory western grebes. *PLOS ONE*, *8*(6), e65408. doi: 10.1371/journal.pone.0065408

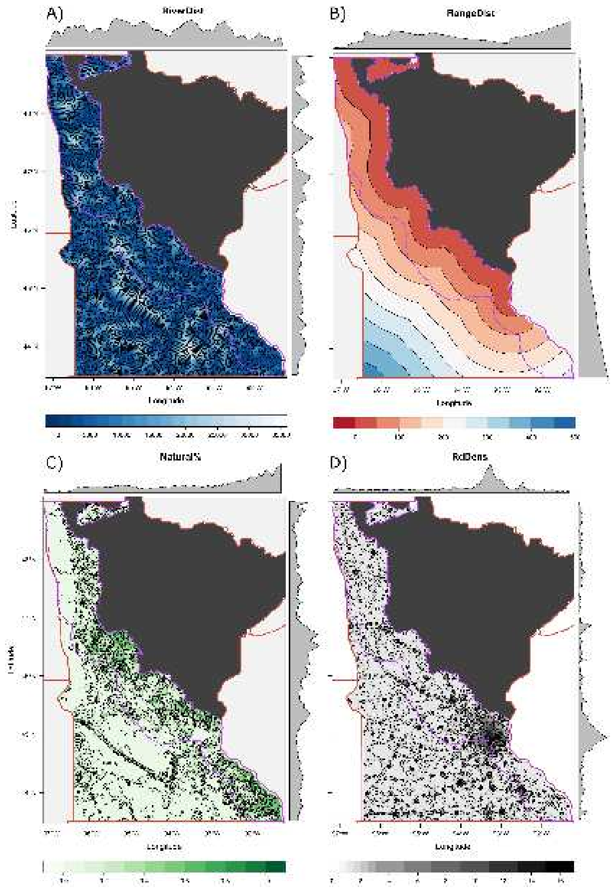
Wilton, C. M., Belant, J. L., & Beringer, J. (2014). Distribution of American black bear occurrences and human–bear incidents in Missouri. *Ursus*, *25*(1), 53–60. doi: 10.2192/URSUS-D-13-00017.1

Xian, G. Z., Homer, C. G., Dewitz, J., Fry, J., Hossain, N., & Wickham, J. (2011). Change of impervious surface area between 2001 and 2006 in the conterminous United States. *Photogrammetric Engineering and Remote Sensing*, *77*(8), 5.

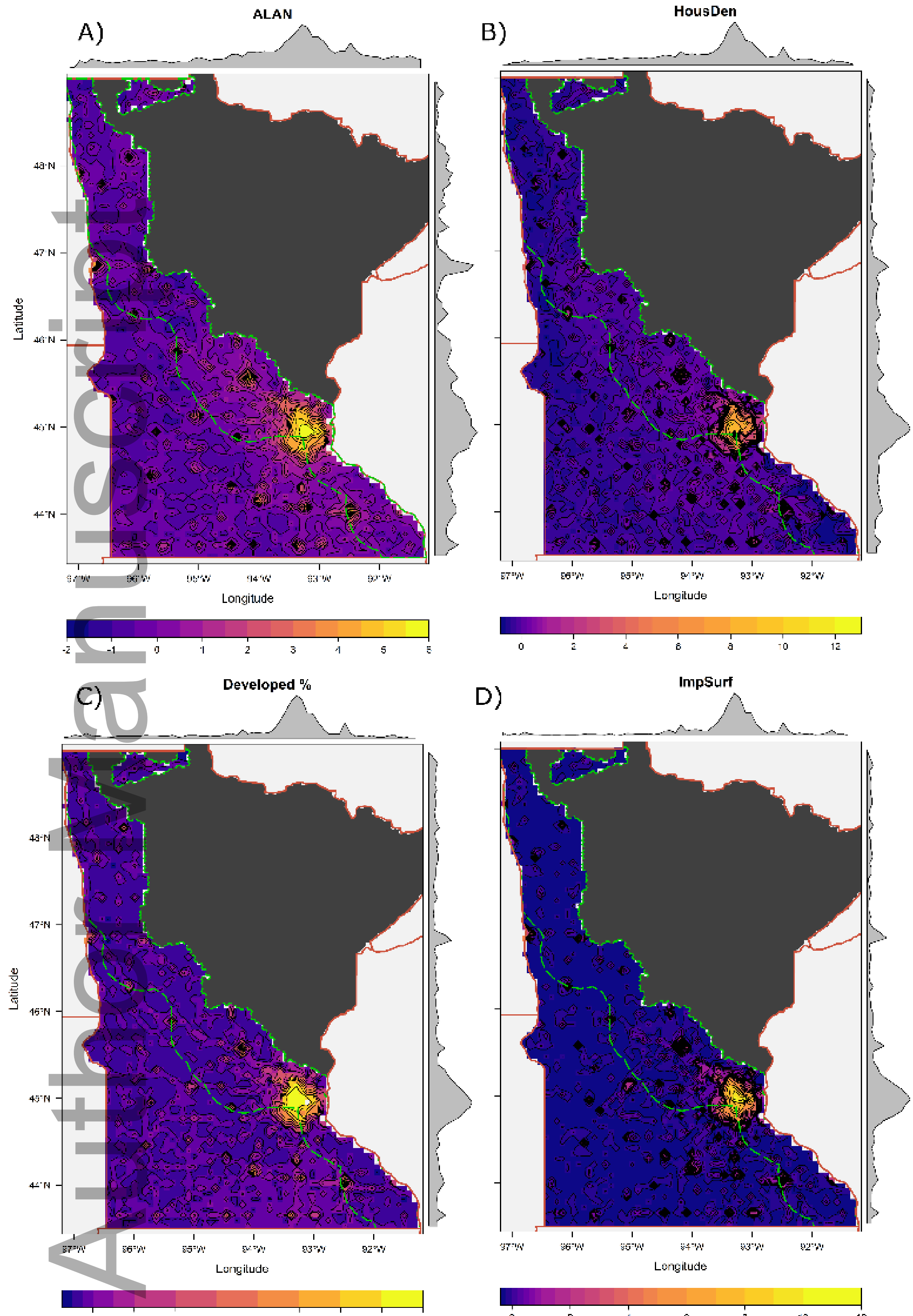
Zeller, K. A., Wattles, D. W., Conlee, L., & DeStefano, S. (2019). Black bears alter movements in response to anthropogenic features with time of day and season. *Movement Ecology*, 7(1), 19. doi: 10.1186/s40462-019-0166-4



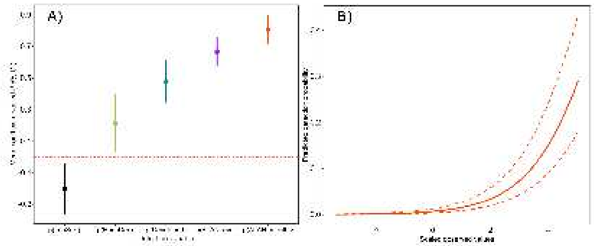
jane_13338_f1.png



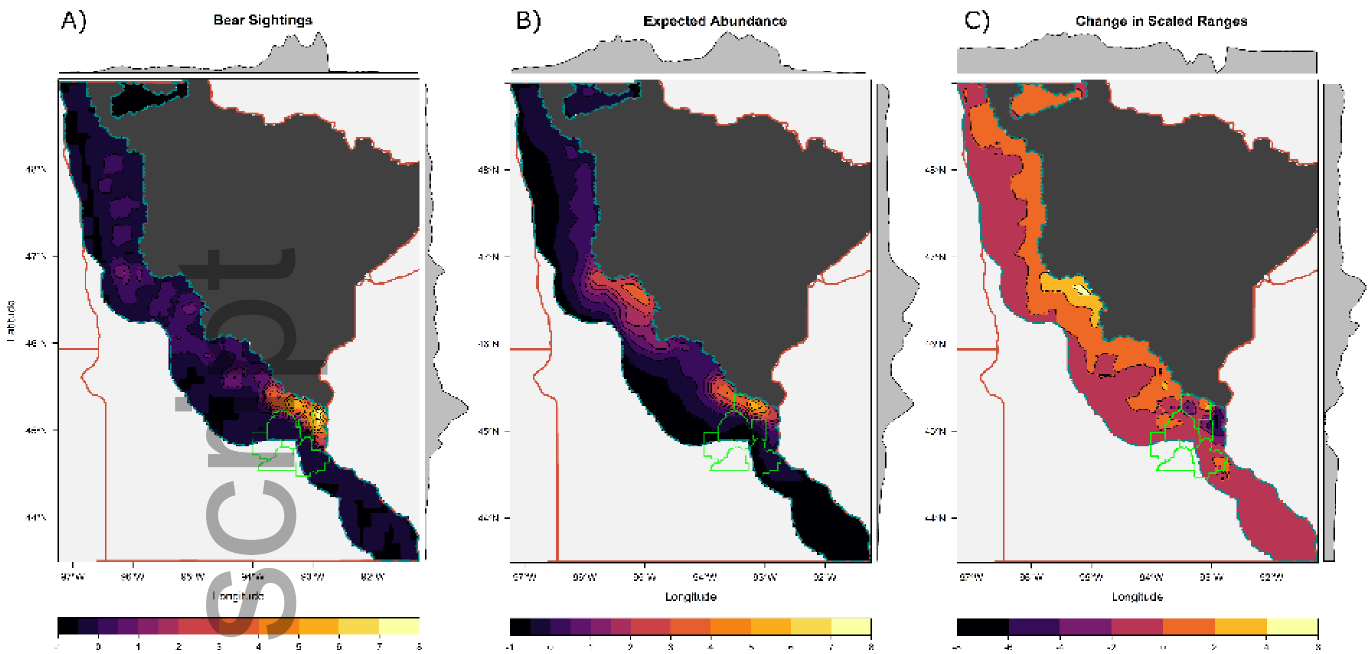
jane_13338_f2.png



jane_13338_f3.png



jane_13338_f4.png



jane_13338_f5.png