

Western University

Scholarship@Western

---

Statistical and Actuarial Sciences Publications

Statistical and Actuarial Sciences Department

---

2023

## Invasion dynamics of the European Collared-Dove in North America are explained by combined effects of habitat and climate

Yiran Shao

Danielle Ethier

Simon Bonner

Follow this and additional works at: <https://ir.lib.uwo.ca/statspub>



Part of the [Statistics and Probability Commons](#)

---

Invasion dynamics of the European Collared-Dove are  
explained by combined effects of habitat and climate

Yiran Shao<sup>1</sup>      Danielle Ethier<sup>2</sup>      Simon Bonner<sup>1\*</sup>

<sup>1</sup>Department of Statistical and Actuarial Science, Western University,  
London, ON, Canada

<sup>2</sup>Birds Canada, Port Rowan, ON, Canada

\* Corresponding author: Simon Bonner, University of Western Ontario 1151

Richmond Street London, Ontario, N6A 5B7 Canada; E-mail:

sbonner6@uwo.ca.

## Abstract

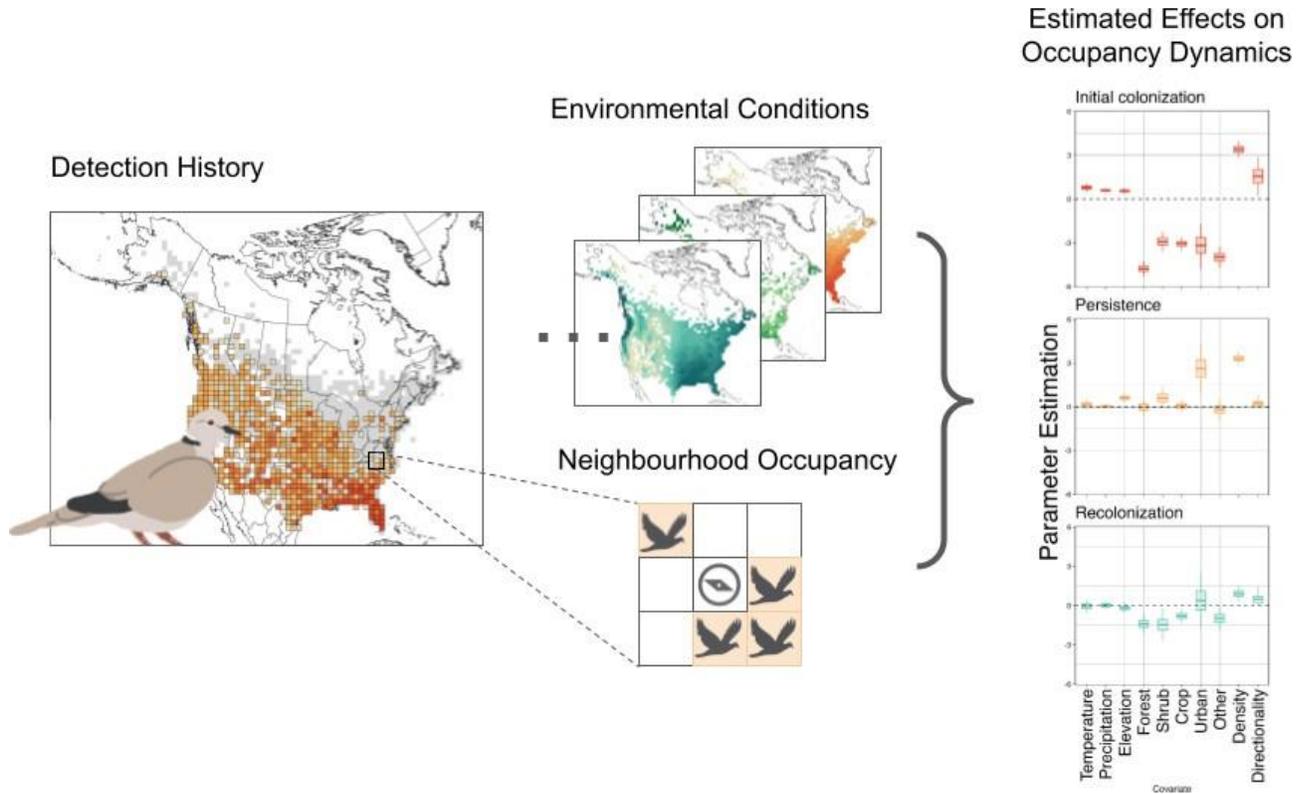
Global biodiversity is increasingly threatened by the spread of invasive species. Understanding the mechanisms influencing the initial colonization and persistence of invaders is therefore needed if conservation actions are to prevent new invasions or strive to slow their spread. The Eurasian collared-dove (*Streptopelia decaocto*, EUCO) is one of the most successful avian invasive species in North America; however, to our knowledge, no study has simultaneously examined the role that climate-matching, human activity, directional propagation, and local density have in this invasion process. Our research expands upon a cellular-automata-based hierarchical model developed to assess directional invasion dynamics to further quantify the impacts of climate, elevation, and land cover type on the spread of EUCO in North America. Our results suggest that EUCO's dispersal patterns can largely be explained by the effects of habitat, climate, and environmental conditions at different stages of the invasion process rather than some innate preferred north-westerly spread. Specifically, EUCO initially colonized warm and wet grassland habitats and tended to persist in urban areas. We also found that while EUCO were more likely to spread to the northeast of existing habitats, directional preference did not drive persistence and recolonization events. These findings highlight the importance of incorporating both neighbourhood effects and environmental factors in the modelling of range-expanding species, adding to the toolset available to researchers to model invasive species spread. Further, our research demonstrates that historical records of invasive species occurrences can provide the data resources needed to disentangle the characteristics driving species invasion and enable predictions that are of critical importance to resource managers.

*Keywords:* Bayesian modelling, Citizen science, Habitat condition, Invasive species, Management planning, Range expansion, *Streptopelia decaocto*

## Lay summary

- The Eurasian collared-dove is an invasive species in North America. Starting from the Bahamas in 1974, it has colonized much of the US and Canada.
- The dove tended to spread to the northwest, and previous studies suggested that this was due to an innate preference for this direction. However, these studies did not consider that this might also be explained by habitat selection.
- We studied doves' range expansion using data from the Project Feeder Watch citizen science study. Our findings suggest that they prefer grasslands and thrive in areas with higher temperatures, more precipitation, and higher elevations, with little regard for direction.
- These results add to our understanding of the invasion process of the collared-dove in North America, and broadly highlight the importance of historical records to disentangle the characteristics driving species invasion that are important to resource managers. The new methods can also be applied to understand the spread of other invasive species.

# Graphical Abstract



## 1 Introduction

Invasive species represent the second most common cause of extinction and pose a major threat to global biodiversity through competition, predation, disease transmission, and hybridization within native communities (Sala et al. 2000, Bellard et al. 2016, Davis 2003, Gurevitch and Padilla 2004). Despite their global influence, management efforts are often unable to limit the spread of invasive species (Cohen et al. 2022). Prediction of when and where species are likely to expand their ranges requires a quantitative understanding of the biotic and abiotic mechanisms responsible, such as geographic variation in the habitat and

environmental conditions (e.g., wind, precipitation, temperature) or the presence of preferential direction of spread (e.g., anisotropic; Morin et al. 2009). Determining the interplay between these characteristics is therefore essential to making effective management decisions to stop or slow the spread of invasive species.

Among all avian invaders, not many have been as successful as the Eurasian collared-dove (*Streptopelia decaocto*, EUCO) in expanding their range and establishing breeding populations, making EUCO an excellent case study for the spread of invasives. EUCO is a large dove with a distinctive black half-collar at the nape of the neck. EUCO are often found in suburbs, farmland, wood edges and open country, foraging seeds, berries and insects in flocks (Romagosa 2002). They are generally acknowledged to be nonmigratory, but young birds may disperse long distances which aids in the spread of the population (Romagosa 2002).

Native to South Asia and limited areas of southeastern Europe, these doves were initially brought to the Americas as a pet and were accidentally released to the Bahamas in 1974 (Romagosa and McEneaney 1999, Romagosa 2002). Over the past 48 years, they have expanded across the continental US, reaching Florida in 1982, California in 1992, and British Columbia, Canada, in 2000, where their expansion is still ongoing (fig. 1) (Romagosa and McEneaney 1999, Romagosa 2002, Koenig 2020). Although a direct impact of this invader on native birds through competition has not been observed, concerns have been raised about their potential to transmit disease and parasites (Poling and Hayslette 2006, Romagosa and Labisky 2000, Koenig 2020). Studies show that EUCOs are able to transmit a variety of pathogens, including the West Nile virus, *Chlamydia psittaci*, which causes endemic avian chlamydiosis, and *Trichomonas gallinae*, which causes avian trichomoniasis (Rappole et al. 2000, Panella et al. 2013, Donati et al. 2015, Stimmelmayer et al. 2012). All of these can be fatal to native birds. Therefore, models predicting their invasive spread across international borders must consider the factors driving both initial colonization and persistence in an area. EUCOs are known to live in a wide range of

habitats, but the association between habitat characteristics and invasion dynamics is not well understood. Previous studies found that EUCOs have a particular preference for human-altered landscapes including urban and agricultural areas (Fujisaki et al. 2010, Saâd et al. 2021, Scheidt and Hurlbert 2014). For example, EUCO occurrences in Europe are associated with suburban areas with a mixture of shrub and tree cover, often close to human-made structures with a constant food supply (Coombs et al. 1981). These findings support the “human-activity” hypothesis, which states that anthropogenic changes make new environments more invadable by altering native communities and reducing biotic resistance (Sax and Brown 2000). Human activity may also offer new niche opportunities such as supplementary feeding.

On the other hand, the “climate matching” hypothesis, states that invasive species have a higher probability of success if they are introduced into regions with a climate like that found in their native area (Abellán et al. 2017, Redding et al. 2019), as species tend to conserve their native climatic niche. Eraud et al. (2007) found that both detection and occurrence probabilities are low for EUCOs in elevated areas with high proportions of forest coverage. Further, the range of EUCOs is believed to be limited to warmer areas; however, it is unclear how much their climatic preference weighs on habitat selection (Fujisaki et al. 2010, Scheidt and Hurlbert 2014). According to their invasion history in Europe, they seem to predominate in areas with warmer temperatures and higher levels of precipitation, similar to their native subtropical habitats in Asia; however, their occurrence has also been reported in areas as cold as the Ural Mountains in Russia (Hagemeijer and Blair 1997).

In addition to the dynamic interplay between biotic and abiotic drivers of invasion dynamics, a commonly accepted empirical hypothesis originally described by Coombs et al. (1981) states that there exists a tendency for EUCO to expand its range in a north-westerly arc (see also Slager 2020, Fujisaki et al. 2010). The initial invasions of EUCO in Europe and North America both appeared to progress primarily northwestward. In Europe, EUCOs started their invasion in Turkey and the Balkans in southeast Europe, then reached Hungary

in 1932, Germany in 1943, France in 1950, and Britain in 1955 (Hudson 1965). Similarly, their Bahamas-Florida-California-British Columbia invasion route in North America also shows a general west-northwest direction (Hooten et al. 2007, Hooten and Wikle 2008). This “north-westerly” hypothesis was initially made based on these observations, and later quantitatively supported by the study of Bled et al. (2011) that included directionality as an explicit component of the expansion dynamics. However, the relationship between directional movement and environmental factors remains a critical uncertainty in predicting the invasion process, not just for EUCOs, but also for other invading species.

In this study, we take advantage of an international citizen-science dataset collected through Project FeederWatch (PFW) to model the invasion dynamics of the EUCO and examine how these patterns are affected by habitat and climatic factors across North America, while also assessing the role of directional spread. Hierarchical models provide users with flexible statistical tools that can incorporate biological hypotheses while accounting for sources of sampling bias within count data. Specifically, citizen-science data suffers from imperfect detection, and failure to account for imperfect detection can result in biased or misleading inferences (MacKenzie et al. 2002). We, therefore, adopt the use of an extended cellular-automata-based hierarchical occupancy model (Bled et al. 2011), which is a spatially explicit expansion of the dynamic occupancy model (Royle and Kery 2007) to model the invasion dynamics of collared-doves. Our modelling approach provides users with a straightforward way to express population dynamics parameters (colonization, persistence, and recolonization probabilities) as functions of predictor variables, while also considering the number of occupied sites in the neighbourhood, directional spread, and detection probability. Further, we compare the full model with reduced models to test specific hypotheses about the persistence of the doves once sites are occupied and the effects of directionality on dispersal. Although we expect that persistence is permanent for the doves so that no temporary extinction or recolonization occurs, including recolonization makes the modelling framework more broadly accessible for studying the

dispersal of other species. For example, plants may leave seeds which can germinate long after the initial individuals have disappeared, facilitating a new colonization event (Keeley 2006).

While our selection of habitat and environmental covariates was informed by the literature (hypotheses detailed previously), our analysis was largely exploratory with regard to their relative effects on the colonization and persistence processes. We predict that if the “human-activity” hypothesis is supported, occupancy, colonization, and persistence would have positive relationships with anthropogenically disturbed habitats. On the other hand, if the “climate matching” hypothesis is supported, we would predict occupancy, colonization, and persistence to be lower in areas with high elevation and proportions of forest coverage, and higher in warmer areas with greater amounts of precipitation. However, the high dispersal potential of the EUCO may result in temporal patterns of occupancy that are largely independent of landscape features, and more influenced by the spatial arrangement of previously occupied sites and northwest directional spread.

## **2 Methods**

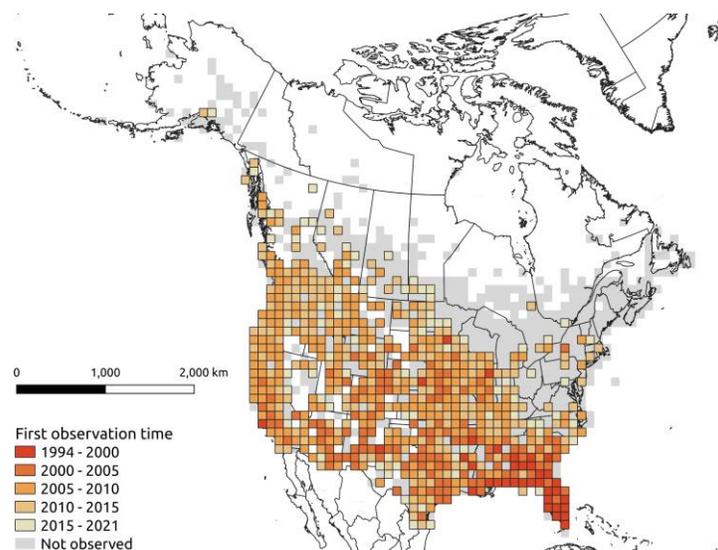
### **2.1 Data sources**

#### **2.1.1 Occurrence Records**

Project FeederWatch is a winter bird monitoring program jointly administered by the Cornell Laboratory of Ornithology and Birds Canada (Bontar and Cooper 2012). Wells et al. (1998) provides a detailed description of the PFW protocol. Briefly, volunteer participants record the highest number of each species seen at one time visiting feeders in their yards during survey periods of two consecutive days. These surveys occur at one-week intervals from

November through April, and survey days are selected ahead of time to avoid biasing counts toward days with high numbers. Participants also provide additional details about the location, date, and time of their observations. At the time of writing, the PFW sites were located in all US states and all provinces and territories in Canada. The raw PFW data collected between 1994 and 2021 contains 342,910 submissions of bird sightings obtained from 70,374 unique sites. EUCOs were identified in 18,660 (5.4%) records from 5,619 (8.0%) unique sites (fig. 1).

To understand the patterns of species occurrence at a regional resolution, we utilized grid cells as our sample units. We adapted the grid cell structure used by the Christmas Bird Count (CBC) to compute the annual total counts of observations reported from each cell. The CBC grid divides continental North America into 2,423 grid cells of approximately  $100 \times 100$  km (fig. 1) (Meehan et al. 2019). After removing cells covered mostly ( $> 50\%$ ) with water, we are left with 2,208 cells as our study area. Among the cells, 1,134 (51.36%) of them had at least one survey completed during the examined period and 646 (29.26%) of the cells had at least one detection record of EUCO. Our models consider the number of surveys that include observations of EUCO out of the total number



of surveys conducted in each cell for each year as the binomial response variable.

Figure 1: Map of the grid cell structure and the year of initial observation records of the Eurasian collared-dove (*Streptopelia decaocto*) from the Project FeederWatch database in North America. Eurasian collared-dove records were collected between 1994 and 2021 (18,660 observation reports out of 342,910 surveys) and were grouped by 5-year periods. Map produced using the NAD 1983 Albers North America projection (EPSG:102008). Data source for temperature and precipitation: ScienceBase North America Climate data (1950–2000); elevation: ScienceBase North America Elevation data; forest, shrub, crop, urban and other land cover types: North American Land Change Monitoring System (NALCMS) land cover map 2015 (2015). Scale: 1:30,000,000.

### **2.1.2 Environmental Variables**

We included a set of eight environmental covariates that have been shown in the literature to be biologically influential on the dispersion pattern of EUCCO or other similar bird species (Fujisaki et al. 2010, Scheidt and Hurlbert 2014, Banko et al. 2002, Davies et al. 2007). The two climate variables, mean annual temperature and precipitation, were derived from the monthly average measurements between 1950 and 2000 downloaded from the ScienceBase North America Climate data derived from WorldClim database, at a 30-arc-second-resolution (1-km) grid of values. We chose the North America elevation data with the same resolution from the ScienceBase North America Elevation data derived from the GTOPO30 dataset as a topography indicator. We summarized all variables by computing the average value over time in each grid cell. For land cover information, we obtained our raw data from the North America Commission for Environmental Cooperation’s North American Land Change Monitoring System (NALCMS) land cover map 2015 at 30 meters. To simplify the models, we grouped the 19 original land cover classes into five broad categories: forest, shrub, grassland, crop, and urban (see table S1 in the Supplementary materials). The proportion of each land cover category was calculated for each cell. Unclassified areas within a cell were tagged as “other”. We consider grasslands as the reference category for land cover in

our analysis. All environmental variables in our models were constant over time; however, further extensions could allow for time-varying covariates in order to investigate temporal effects including yearly fluctuations in weather and anthropogenic climate change. Plots of the spatial distribution of each covariate are provided in fig. S1 in the Supplementary Materials.

## 2.2 Hierarchical Modelling

### 2.2.1 Model Descriptions

The models applied in this study build on the hierarchical occupancy based cellular automata model proposed by Bled et al. (2011), which in turn is based on the methods of Royle and Kéry (2007). We largely follow their theory and notation in describing our extended models, which take into consideration environmental covariates, separate the effect of directional preference in dispersion from the neighbourhood effect and allow the strength of the directional influence to be estimated from the data.

Our full model allows for the incorporation of all three components of dynamic occupancy and allows for the separate effects of both neighbourhood density and directionality on each of these components. Let  $\mu_{i,t}$  denote the probability that site  $i \in \{1, \dots, M\}$  is occupied at time point  $t \in \{1, \dots, T\}$  based on the value of the covariates and whether it and its neighbouring sites are occupied at the previous time point. The dynamics of occupancy status are accounted for by three parameters: initial colonization (the probability that a site is occupied for the first time),  $\gamma$ ; persistence (the probability that a site remains colonized), denoted by  $\phi$ ; and recolonization (the probability that a once-occupied site that was abandoned is occupied again),  $\theta$ . The initial and recolonization events are distinguished by the availability of a site for first colonization,  $A_{i,t}$ , where  $A_{i,t} = 1$  indicates that the site has never been colonized before and, therefore, is available for

initial colonization, and  $A_{i,t} = 0$  otherwise. Let  $Z_{i,t}$  represent the true occurrence of the species at site  $i$  during time period  $t$ , so  $Z_{i,t} = 1$  when the site is occupied and  $Z_{i,t} = 0$  otherwise. Mathematically,

$$Z_{i,t} \mid Z_{i,t-1} \sim \text{Bern}(\mu_{i,t}) \quad (1)$$

with  $A_{i,1}=0$ ,

$$A_{i,t} = \begin{cases} 0, & \text{if } \sum_{s=1}^{t-1} Z_{i,s} > 0 \\ 1, & \text{otherwise} \end{cases} \quad (2)$$

for  $t=2, \dots, T$ , and

$$\begin{aligned} \mu_{i,t} = & \phi_{i,t-1} Z_{i,t-1} + \\ & \gamma_{i,t-1} (1 - Z_{i,t-1}) Z_{i,t-1} + \\ & \theta_{i,t-1} (1 - Z_{i,t-1}) (1 - A_{i,t-1}). \end{aligned} \quad (3)$$

Table 1: Notation for the variables in the cellular-automata-based hierarchical occupancy model for Eurasian collared-dove (*Streptopelia decaocto*) with prior distributions for the model parameters.

Parameter	Description	JAGS Notation	Prior Distribution
Input data			
$Z_{i,t}$	Site occupancy state	$z[i,t]$	
$Y_{i,t}$	Number of detections in a site	$Y[i,t]$	
$K_{i,t}$	Number of visits to a site	$K[i,t]$	
$X_i$	Environmental covariates	Elev_Mean[i], forest_per[i], ...	
$\delta_{\vec{j}}$	Angle of the vector joining the center of cell j and i	angle1[i, j]	
Derived parameters			
$\mu_{i,t}$	Site occupancy probability	$\text{muZ}[i,t]$	
$\gamma_{i,t}$	Initial colonization parameter	$\text{gamma}[i,t]$	
$\phi_{i,t}$	Persistence parameter	$\text{phi}[i,t]$	
$\theta_{i,t}$	Recolonization parameter	$\text{theta}[i,t]$	
$A_{i,t}$	Site availability for initial colonization	$A[i,t]$	
$w_{ij,t}$	Weight of the occupancy status of neighbour cell j on the future occupancy status of the central cell i	$w[i,j,t]$	
Model parameters			
$\gamma_{0t}, \phi_{0t}, \theta_{0t}$	Intrinsic initial colonization, persistence and recolonization probability	$\text{lgamma0},$ $\text{lphi0},$ $\text{ltheta0}$	$N(0, 1)$
$\gamma_1, \phi_1, \theta_1$	Parameters controlling	$\text{lNeigPhi1},$	$N^+(0, 1000)$

	neighbourhood effect		INeigGamma1, INeigTheta1	
$\gamma_2, \phi_2, \theta_2$	Parameters controlling directional effect		IDirePhi1, IDireGamma1, IDireTheta1	$N(0, 1000)$
$\gamma_{3...10},$ $\phi_{3...10},$ $\theta_{3...10}$	Parameters controlling environmental effects		IElevPhi1, IElevGamma1, IElevTheta1, ...	$N(0, 25)$
$\delta_0$	Direction of preference		dirSpread	$U(-\pi, \pi)$
$p$	Detection probability		$p$	$U(0, 1)$

---

The dependency on the occupancy of the surrounding cells and on the climate, elevation and land cover covariates is then incorporated into the three dynamic parameters. Specifically, we define the persistence parameter  $\phi$  as

$$\text{logit}(\phi_{i,t-1}) = \phi_{0t} + \phi_1 L_{1i,t-1} + \phi_2 L_{2i,t-1} + \sum_{k=1}^8 \phi_{k+2} X_{i,k} \quad (4)$$

where  $L_{1i,t}$  and  $L_{2i,t}$  are the covariates reflecting the occupancy density in the neighbourhood and the directional preference of dispersion (defined below), respectively,  $\phi_3$  through  $\phi_5$  are the coefficients for mean elevation, temperature, and precipitation, and  $\phi_6$  through  $\phi_{10}$  are the coefficients for the effect of the land cover classes (forest, crop, shrub, urban and other) with their values estimated with respect to the effect of grassland. Similarly, the initial colonization parameter  $\gamma$  and recolonization parameter  $\theta$  are expressed as

$$\text{logit}(\gamma_{i,t-1}) = \gamma_{0t} + \gamma_1 L_{1i,t-1} + \gamma_2 L_{2i,t-1} + \sum_{k=1}^8 \gamma_{k+2} X_{i,k} \quad (5)$$

$$\text{logit}(\theta_{i,t-1}) = \theta_{0t} + \theta_1 L_{1i,t-1} + \theta_2 L_{2i,t-1} + \sum_{k=1}^8 \theta_{k+2} X_{i,k} \quad (6)$$

To take into account the unexplained directional preference of spread, we included two covariates, the neighbourhood effect covariate  $L_{1i,t}$  and the directional preference covariate  $L_{2i,t}$ , into the calculations of the dynamics parameters. The neighbourhood effect covariate,  $L_{1i,t}$ , is calculated as the proportion of cells within the neighbourhood of  $i$  that are occupied at time  $t$ . The directional preference covariate,  $L_{2i,t}$ , is calculated as the averaged cosine of the angle of difference between the preferred direction of spread and the vector that points from the occupied neighbours to the central site  $i$  at time  $t$ . Let  $N_i$  be the set of  $n_i$  available neighbours for site  $i$ . Most cells in a square grid will have  $n_i = 8$  neighbours, though those at the edges will have only  $n_i = 5$  and those at the corners will have only  $n_i = 3$ . In practice,

square cells might have anywhere between 1 and 8 neighbours after masking the possible range (e.g., removing cells mostly covered by water). Then:

$$L_{1i,t-1} = \frac{\sum_{j \in N_i} Z_{j,t-1}}{n_i} \quad (7)$$

and

$$L_{2i,t-1} = \frac{\sum_{j \in N_i} Z_{j,t-1} \cos(\delta_0 - \delta_{ji})}{n_i} \quad (8)$$

where  $\delta_0$  is the angle of the vector defining the preferred direction of spread and  $\delta_{ji}$  is the angle of the vector joining the centres of the cells  $i$  and  $j$  (fig. 2). In our implementation, the prior and posterior distributions of the directionality factors,  $\phi_2$ ,  $\gamma_2$ , and  $\theta_2$ , are truncated at zero. This is necessary because a positive effect in one direction has exactly the same implication as a negative effect in the opposite direction (e.g., the persistence probabilities computed when  $\phi_2 = 1$  and  $\delta_0 = \pi/2$  or  $\phi_2 = -1$  and  $\delta_0 = -\pi/2$  are identical provided the other parameters remain the same). Truncating the distribution of these parameters ensures that the model is identifiable and that the posterior distributions are not bimodal.

Note that the range of possible values for  $L_2$  depends on the number of values for a cell,  $n_i$ . The maximum range is  $(-1.00, 1.00)$ , but this only applies to cells with 1 neighbour which accounted for just 3 (0.14%) of the cells in our grid. The vast majority of cells, 1,677 (75.95%), were completely surrounded by other cells so that  $n_i = 8$ . For these cells, the range of  $L_2$  is limited to  $(-0.30, 0.30)$ . These cells will have the most impact on the estimates of the directionality effect because they predominate, so we focus on the effect within this range in our results.

The model of Bled et al. (2011) includes directionality in a similar way, except that the effects of neighbourhood occupancy and the directional preference are combined into a single covariate. To recover their model of the initial occupancy probability from our full model,

for example, we need to set  $\gamma_1 = \gamma_2 = \gamma^*$  in equation (5). The equation then becomes

$$\text{logit}(\gamma_{i,t-1}) = \gamma_{0t} + \gamma_1^* D_{1i,t-1} + \sum_{k=1}^8 \gamma_{k+2} X_{i,k} \quad (9)$$

where

$$D_{1i,t-1} = L_{1i,t-1} + L_{2i,t-1} = \frac{\sum_{j \in N_i} w_{ij} Z_{j,t-1}}{n_i} \quad (10)$$

as given in Bled et al. (2011, eqn. 12) with  $w_{ij} \propto 1 + \cos(\delta_0 - \delta_{\bar{j}})$  as in Bled et al. (2011, eqn. 14). The same applies also to the definition of persistence and recolonization probabilities. A key restriction of this formulation is that it is not possible to separate the strength of the neighbourhood effect from the strength of directionality. The assumption of this equality means that if the occupancy dynamics of the central site are influenced by the proportion of occupied sites in its neighbourhood, then there must be an effect of the same strength caused by the directionality preference. By separating the neighbourhood effect  $L_{1i,t}$  and the directionality preference  $L_{2i,t}$  in our full model, we can independently assess the impact of what proportion of cells in a neighbourhood are occupied and exactly in which direction the neighbouring cells lie. Note that the model of Bled et al. (2011) does include the quadratic effect of  $D_{1i,t-1}$  in the linear predictors to account for possible non-linearity. However, adding the quadratic effects of both the neighbourhood and directionality effects, and potentially their interaction, would significantly increase the number of parameters, and we have chosen not to include these terms in our full model.

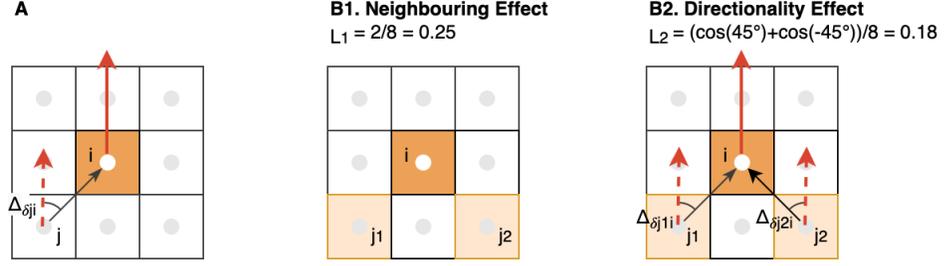


Figure 2: Graphical representation of the neighbourhood structure of cells used in the cellular-automata-based hierarchical occupancy model for Eurasian collared-dove (*Streptopelia decaocto*) A shows the vectors and angles used in directional preference estimation. B1 and B2 demonstrate an example of calculation for the neighbourhood effect factor  $L_1$  (B1) and the directionality effect factor  $L_2$  (B2). The red arrow indicates the preferred direction of spread and  $\Delta_\delta$  denotes the angle it makes with the vector  $\vec{j}i$  pointing from a neighbour  $j$  of the central cell  $i$ , indicated by black arrows.

Finally, our full model accounts for imperfect detection by incorporating the detection probability per survey,  $p$ , into the observation process. For the simplicity of the model, we assumed the probability of detection given that at least one dove is truly present in a cell to be constant for all surveys over all years. The number of detections  $Y_{i,t}$  in  $K_{i,t}$  surveys can then be modelled as

$$Y_{i,t} \mid Z_{i,t} \sim \text{Bin}(K_{i,t}, pZ_{i,t}). \quad (11)$$

### 2.2.2 Model Fitting

We estimated the parameters of the full model by fitting it to data obtained from all surveys in the years 1994 to 2021, and then tested the model by comparing the predicted detections of EUCO in 2021 with the actual observation records from the PFW dataset. In particular, we computed the area under the curve (AUC) for the reporter-operator curve (ROC) as a measure of prediction accuracy. A larger testing data set would be beneficial for assessing the

prediction accuracy of the model. However, the models are complex and the primary goal of our analysis was to examine the relationships between the predictors and the processes governing occupancy and to estimate the parameters of these models. For these reasons, we have limited the validation to a single year. If larger data sets are available then more years of validation could be used to assess the future predictions of occupancy.

We analyzed the models in the Bayesian framework. The parameters we incorporated with their corresponding notation and prior settings are summarized in table 1. Note that the prior distributions for the parameters are all assumed to be independent. The notation  $N(\mu, \sigma^2)$  indicates that the assigned prior was normal with mean  $\mu$  and variance  $\sigma^2$ ,  $U(l, u)$  indicates that the prior was uniform over the interval  $(l, u)$ , and  $N^+(\mu, \sigma^2)$  indicates that the prior was half-normal, following the shape of a normal distribution with mean  $\mu$  and variance  $\sigma^2$  but restricted to positive values.

To explore the posterior distributions, we employed JAGS, a program for Bayesian inference using Markov chain Monte Carlo (MCMC) sampling, controlled via an R script running under R 4.2.1 using package **rjags** 4.10 (Plummer 2022, R Core Team 2021). We ran the full model with four chains, for 25,000 iterations after an adaptive phase of 1,000 iterations and a burn-in of 10,000 iterations. Convergence was assessed with the Brooks–Gelman–Rubin potential scale reduction factor generated from the burn-in sample (Gelman and Rubin 1992, Brooks and Gelman 1998).

All environmental covariates were scaled prior to the model fitting process. The mean temperature, precipitation, and elevation data were standardized to have mean 0 and standard deviation 1. The land cover proportions were converted to percentages so that they fell between 0 and 100%. Pearson’s correlation coefficient was computed to assess potential collinearity among the covariates.

To interpret the results of our Bayesian model, we report the posterior mean as a point estimate and the 2.5% and 97.5%-iles as the bounds of 95% credible intervals (CIs) to summarize the posterior distribution, in the format of mean[95% CI]. For example, the

estimate of the factor  $\gamma_{\text{forest}}$  is  $-4.67[-5.24, -4.11]$ , which indicates strong evidence of a negative effect relative to that of grassland given that the range is well below zero and very narrow (indicating high precision for estimation). In contrast, the estimate of  $0.73[-1.50, 2.84]$  for  $\theta_{\text{urban}}$  provides no evidence of an effect given that the interval covers 0 and is very broad.

### **2.2.3 Model Comparison**

Along with our full model, we also fit two reduced models to subsets of the data in order to test specific hypotheses. The first assumes that the persistence probability is equal to one and removes the recolonization process so that each cell remains occupied in all years following its initial colonization (mathematically,  $P(Z_{i,t+1} = 1/Z_{i,t} = 1) = 1$ ). The second model removes the effect of directionality from the model (equivalent to setting  $\phi_2 = \gamma_2 = \theta_2 = 0$ ). We compared these models using the Watanabe-AIC or Widely Applicable Information Criterion (WAIC, Watanabe (2010), Vehtari et al. (2017)) computed via the **NIMBLE** package for R (de Valpine et al. 2017). However, the large number of latent variables meant that we were unable to compute the WAIC for the full data set. Instead, we fit the models to a subset of the data including only 5 years of observations from sites near the origin of the invasion in Florida. Further details and the results of these comparisons are provided in Section S1 of the Supplementary Materials.

### 3 Results

Evidence suggested convergence for the MCMC chains (point estimates of potential scale reduction factors  $\leq 1.06$  for all hyperparameters). The median of the Effective Sample Size (ESS) of the sampled parameters was 4391.26, with a minimum of 672.4 (temperature effect on initial colonization) and a maximum of 62,857.37 (intercept for recolonization in 1994). The Receiver Operating Characteristic (ROC) plot suggested that our full model produces very good discrimination of the occupied and unoccupied sites (AUC = 0.95, 95% CI = [0.93, 0.97]).

#### 3.1 Relationship between Parameters and Occupancy

The per survey detection probability of EUCO was estimated to be very low,  $\hat{p} = 0.0141$  (95% CI = [0.0141, 0.0142]), though the detection probability per cell remains relatively high because of the repeated surveys. Among the 2,208 cells considered in the study, 790 cells (95% CI = [789, 791]) were estimated to have a probability of occupancy greater than 0.50 in 2021, whereas the doves were only detected in 363 cells. Among the cells without EUCO detection records, doves were estimated to be present in 24 cells with at least one survey and 403 cells that were not surveyed at all. Failing to account for detectability would underestimate the number of occupied cells by 54.05% (95% CI = [53.98%, 54.12%]). This clearly demonstrates the need to account for missed detections (false negatives) when modelling a species' range.

Estimates of the covariate effects suggested that high probabilities of each stage of the colonization processes for EUCOs are associated with distinct environmental conditions (fig. 3). Specifically, the full model indicated that initial colonization probabilities were higher in

grassland habitats and increased with temperature, precipitation, and elevation. However, estimates indicated that persistence probabilities were higher in urban areas and increased with elevation and recolonization probabilities were higher in urban and grassland habitats.

Our full model provided strong evidence that the probabilities of the three dynamic processes in species occurrence – initial colonization, persistence, and recolonization – all increased in response to an increasing proportion of occupied neighbouring areas,  $L_1$  (fig. 4, left). Posterior means of these coefficients were all well above 0 relative to the width of the 95% credible intervals and the 95% credible intervals did not cover 0: initial colonization:  $\hat{\gamma}_1 = 3.34$  (95% CI = [2.54, 4.00]); persistence:  $\hat{\phi}_1 = 3.33$  (95% CI = [2.84, 3.82]); recolonization:  $\hat{\theta}_1 = 1.14$  (95% CI = [0.52, 1.78]).

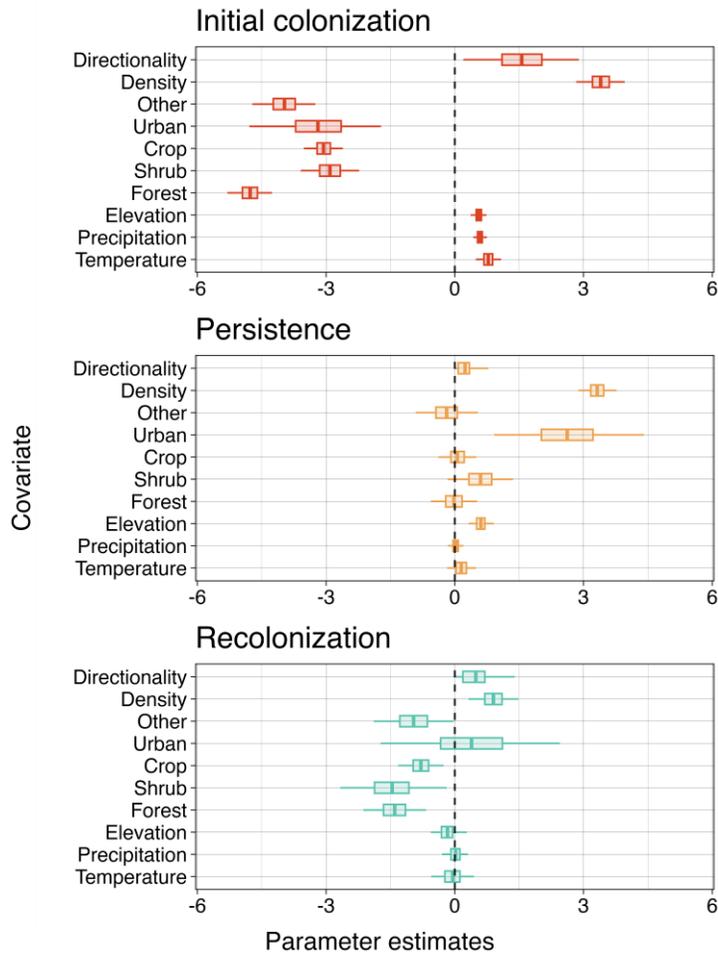


Figure 3: Estimated effects of parameters influencing the initial colonization, persistence and recolonization events for the dispersal of the Eurasian collared-dove (*Streptopelia decaocto*) in North America. Effects of the land cover variables (forest, shrub, crop, urban, and other) are estimated with respect to the effect of grassland. Within each box, horizontal lines denote posterior means; boxes represents 50% credible intervals and vertical extending lines denote 95% credible intervals. Note that the prior and posterior distributions of the directionality factors,  $\phi_2$ ,  $\gamma_2$ , and  $\theta_2$ , are truncated at zero.

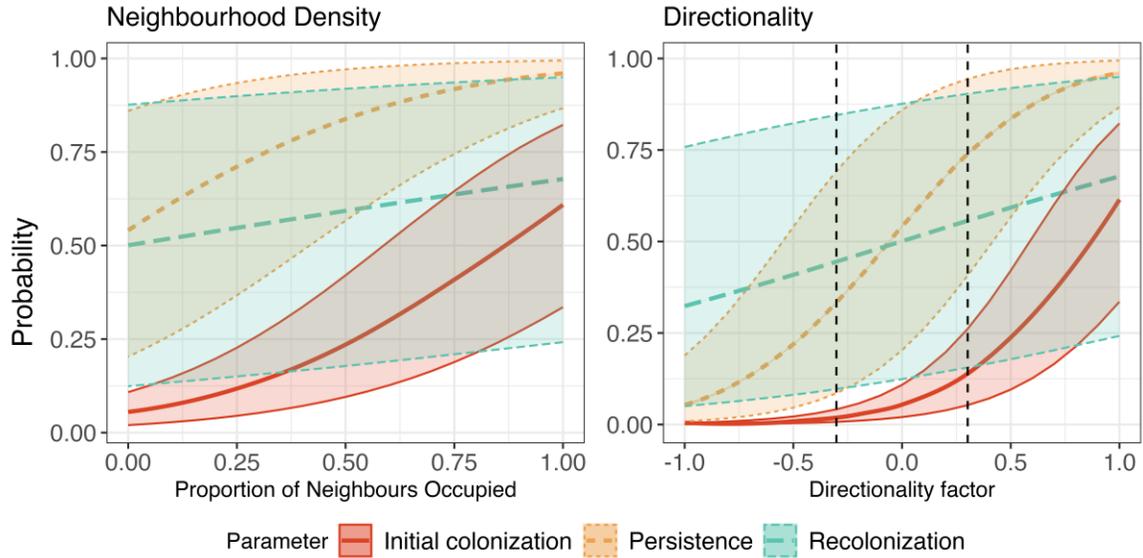


Figure 4: Estimates of the probabilities for each occupancy dynamic processes for the dispersion of the Eurasian collared-doves (*Streptopelia decaocto*) in North America as a function of the local density in neighbourhood (left) and the directionality factor (right), with 95% credible intervals. The habitat covariate was set to the reference (i.e., grassland) and all other covariates were set as mean. The vertical dashed lines indicate the bounds on the directionality factor for the majority of cells with 8 neighbours (75.95% of all cells).

Table 2: Parameter estimates for the full model, 95% credible intervals are shown in brackets. Attributes denoted with \* are considered to be significant as the 95% CIs for the corresponding parameters do not cover zero.

<b>Covariate</b>	<b>Initial Colonization(<math>\gamma</math>)</b>	<b>Persistence(<math>\phi</math>)</b>	<b>Recolonization(<math>\theta</math>)</b>
Temperature	0.78[ 0.49, 1.08]*	0.15[-0.18, 0.50]	-0.05[-0.55, 0.45]
Precipitation	0.59[ 0.43, 0.75]*	0.02[-0.16, 0.20]	0.02[-0.30, 0.31]
Elevation	0.56[ 0.37, 0.74]*	0.61[ 0.32, 0.91]*	-0.17[-0.56, 0.28]
Forest	-4.77[-5.30,-4.26]*	-0.02[-0.56, 0.52]	-1.40[-2.13,-0.67]*
Shrub	-2.91[-3.59,-2.23]*	0.60[-0.17, 1.36]	-1.46[-2.67,-0.19]*
Crop	-3.06[-3.52,-2.61]*	0.06[-0.38, 0.50]	-0.79[-1.32,-0.26]*
Urban	-3.19[-4.79,-1.72]*	2.62[ 0.92, 4.41]*	0.39[-1.73, 2.45]
Other	-3.97[-4.72,-3.25]*	-0.19[-0.91, 0.54]	-0.96[-1.89,-0.03]*
Neighbourhood Density	3.40[ 2.83, 3.96]*	3.32[ 2.88, 3.77]*	0.90[ 0.32, 1.48]*
Directional Preference	1.56[ 0.20, 2.89]*	0.24[ 0.01, 0.78]	0.49[ 0.02, 1.40]

All else being equal, the estimate of the initial colonization probability was generally lower than that of the other two dynamic occupancy processes, increasing above 50% only when all but one of the neighbouring cells were occupied for central cells with 8 available neighbours. The average persistence probability is globally high, indicating that doves tended to remain in areas once established. Although the estimated probability of recolonization increased with neighbouring occupancy, the precision for our estimate was very low. This occurs because there is little evidence of local extinction, which is necessary to obtain information about recolonization. Estimates from the full model suggest that the doves tend to persist once a cell has been occupied, and that any cases in which the doves are observed in a cell in one year, not observed for some years, and then observed again are likely the result of missed detection as opposed to temporary extirpation. This means that there is little chance for recolonization, so that

the data provide almost no information about this process and the factors on which it might depend. This conclusion is also supported by the comparisons of the full and reduced models which provide strong evidence in favour of the model without recolonization (see Supplementary Materials).

The probabilities of the three occupancy dynamics processes generally also increased with L2 (fig. 4, right). The largest effect was for the initial colonization probability, which had an estimated coefficient of 1.56 (95% CI = [0.02, 2.89]). This corresponds to a change in probability from  $< 0.01$  (95% CI = [0.00,  $< 0.01$ ]) to 0.61 (95% CI = [0.34, 0.82]) across the range (-1,1). However, these values apply only to the small number of cells with only 1 neighbour ( $< 0.02\%$ ). As noted in the Methods, the range of the directionality factor is restricted for the vast majority of cells which have 8 neighbours. For these cells, the estimated initial colonization probability ranges from 0.02 (95% CI = [0.01, 0.04]) to 0.14 (95% CI = [0.05, 0.26]) when land cover type is set to its reference value, grassland, and all other covariates are set to their means. In comparison, increasing the proportion of occupied cells within the neighbourhood from 0.25 to 0.75 when land cover type is set to grass and all other covariates are set to their means increases the probability of occupancy from 0.12 (95% CI = [0.05, 0.23]) to 0.41 (95% CI = [0.19, 0.65]). We take this as evidence that the density of EUCO within the neighbouring cells has a stronger effect on initial occupancy than the directionality. Note also that the 95% CI for the probabilities at the end of the range of the directionality factor for the majority of cells almost overlap, suggesting there is relatively weak evidence that directionality affects the initial occupancy probability. The change in the persistence probability across this range was 0.33 (95% CI = [0.09, 0.70]) to 0.74 (95% CI = [0.41, 0.94]) under the same conditions. In this case, the 95% credible intervals overlap considerably, providing little evidence that directionality affects the probability of persistence. Estimates of the effect of directionality on the recolonization probability were very uncertain ranging from 0.45 (95% CI = [0.10, 0.85]) to 0.56 (95% CI = [0.16, 0.90]), and provide no evidence of a true effect. We estimated the preferred direction of movement of EUCO to be N34.20°E, roughly pointing in the Northeast direction, with 95%

CI=[N19.70°W, S70.75°E]. However, our full model indicated that the effect of directionality was weaker than the neighbourhood effect.

## 4 Discussion

Understanding the factors driving dispersion dynamics is key in modelling invasive species range expansion. Specifically, our results highlight the importance of assessing the relationship between directional movement and environmental covariates when predicting the invasion process for EUCOs, and other invading species. Previous studies of the spread of EUCO have suggested that this species increases its range along a preference directional gradient (i.e., “north-westerly” hypothesis). However, the model of Bled et al. (2011) combines the effects of directionality and density within the neighbouring cells into a single term. Our models separate these effects and also incorporate the effects of other predictors. The results from the full model provide only weak evidence of directional preference on the initial colonization probability, and no evidence of directional preference for persistence or recolonization, once the occupancy of neighbouring cells and the environmental covariates have been taken into account. Moreover, the direction of preference indicated by our full model differs from that suggested in previous studies. The full model also provides strong evidence that the occupancy of cells by EUCO depends on the selected environmental covariates, such as a proportion of grassland during initial colonization and a propensity to persist in habitats with higher proportions of human activities and shrub coverage, suggesting support for the “climate matching” hypothesis. Further, the proportion of occupied areas surrounding an unoccupied habitat patch is an important factor to consider when modelling invasion dynamics. Our results do not support the northwestern preferential direction for the invasion spread suggested by Bled et al. (2011), which is rooted in the northwesterly dispersal hypothesis previously made by Coombs et al. (1981), Fujisaki et al. (2010) and Slager (2020). Specifically, our full model provides no evidence that the probabilities of the persistence and recolonization of EUCOs depend on the directionality

factor once the model accounts for the environmental covariates and once it disassociates the effect of neighbourhood occupancy density and directionality preference. We believe that the apparent orientation of EUCO spread observed by other researchers is at least partially governed by their tendency to colonize certain habitats, e.g. grassland, which happened to lie in a specific direction from their point of origin in North America. The full model did provide weak evidence of an effect of directionality on the initial colonization probability, but the effect is small in comparison to the effect of the neighbourhood effect. Moreover, the model identified the preferred direction of dispersal to point to the northeast instead of northwest. Discrepancies in directional finding between our results and Bled et al. (2011) are likely, in part, due to the inclusion of environmental covariates, differences in the dataset used for the analysis, and/or differences in model specification of the neighbourhood effect and directional propagation. These are discussed in turn, and in the context of the available literature on the invasive spread of EUCOs and previously described habitat selection based on habitat modelling outputs.

Our results do not support the northwestern preferential direction for the invasion spread suggested by Bled et al. (2011), which is rooted in the northwesterly dispersal hypothesis previously made by Coombs et al. (1981), Fujisaki et al. (2010) and Slager (2020). Specifically, our full model provides no evidence that the probabilities of the persistence and recolonization of EUCOs depend on the directionality factor once the model accounts for the environmental covariates and once it disassociates the effect of neighbourhood occupancy density and directionality preference. We believe that the apparent orientation of EUCO spread observed by other researchers is at least partially governed by their tendency to colonize certain habitats,

When moving into new habitats, we found a strong correlation between initial occupancy probability and habitats with a high proportion of grassland environments and highly occupied neighbouring areas. We also observed a moderate tendency of initial occupation of areas with higher temperature, precipitation and elevation. EUCOs are more likely to return to previously used habitats with human activities and shrub coverage. Their strong tendency

to occupy grasslands during initial colonization events may explain their early expansion of the range toward and occurrence in the midwestern grasslands of Wyoming, Colorado, and New Mexico before 2005, and the apparent avoidance of geographically closer states to Florida (the location of initial US occupation), such as North Carolina and Kentucky. EUCO's history of spreading to the Southwest US before dispersing to the North in the early years of their invasion may also correspond to their preference for higher temperatures and elevation. EUCOs strong tendency to return to urban areas might be one of the reasons behind their persistence around Los Angeles since 1999, making Southwestern California one of the earliest colonized spots where EUCOs still persist today.

Several smaller-scale studies have identified temperature and proportion of developed land as strong positive predictors for EUCO occurrence (Fujisaki et al. 2010, Scheidt and Hurlbert 2014, Bonter et al. 2010). For example, the study in Florida by Bonter et al. (2010) estimated an occupancy rate  $\geq 50\%$  for areas with  $\geq 40\%$  low-intensity development or  $\geq 20\%$  medium and high-intensity development. By separating the three stages of colonization, we show that these factors may contribute to higher occupancy by affecting different components of the colonization process. While higher temperature is preferable for initial colonization, urban land coverage is more attractive to EUCOs returning to a once-colonized habitat. This suggests that future researchers pay special attention to the colonization history in addition to environmental characteristics when evaluating the vulnerability of habitat to invasion. Results further show a slight positive effect of precipitation on the probability of initial colonization events, which is generally not characteristic of the grassland environments that EUCOs prefer. This result is also inconsistent with Fujisaki et al. (2010), who found a negative correlation between species abundance and annual precipitation. Finer-scaled studies, which assess the correlation between occupancy probability and precipitation are needed to clarify this relationship.

One possible explanation for the different conclusions regarding the direction of spread from our analysis and the work of Bled et al. (2011) is the difference in data selected for the

analysis. Specifically, the temporal scales of the input dataset could have partially contributed to our rejection of the northwesterly hypothesis. Our analytical timing window based on the available data from PFW covered the years 1994-2021, whereas Bled et al. (2011) used BBS data from 1986-2006. The earliest years of spread are therefore not captured in our models, making direct comparisons difficult. Further, even if direction spread was the main driver of initial colonization (which our results suggest is not the case), we would anticipate this effect to dampen with time as the species' movement is restricted by the Pacific Ocean to the west.

Covariates identified as important to the occupancy of EUCOs may also depend on the choice of input data. For example, using the PFW dataset, our full model found a slightly positive correlation between the initial colonization rate and higher precipitation. In contrast, a study conducted with the North American Breeding Bird Survey (BBS) dataset from the US reported the effect of precipitation on EUCO occupancy was inconclusive (Fujisaki et al. 2010). However, another study from the exact same spatial range found the correlation to be positive based on data from the BBS but negative based on data from the CBC (Fujisaki et al. 2010). The disparity between these findings could be due to differences in the sampling period, the data collection protocol, the analytical approach, or the geographic coverage of the input datasets. For example, both CBC and PFW are conducted from winter to spring (CBC: December–January; PFW: November–April), whereas BBS takes place during the summer breeding season (May–July). Even then, CBC counts occur early in the winter season, usually before severe winter storms, whereas the PFW volunteers survey throughout the winter. As EUCOs are reported as not being able to withstand extremely cold conditions, their winter mortality may be reflected by the variance between the CBC and PFW (Hengeveld 1988, Fujisaki et al. 2010). Differences in the count methodologies may also influence the modelling outcomes. Specifically, counts for both CBC and BBS are carried out within predefined areas which means these two datasets are more likely to reflect a balanced and broader range of environmental conditions. In contrast, the PFW protocol

allows participants to repeatedly report at a location of their choice throughout the winter. These locations typically provide supplemental food and occur at the volunteers' homes, backyards, nature centers, community areas, or other locations generally in areas with some human development. The resulting dataset may therefore suffer from under-representation of areas with lower human population densities and be biased toward specific habitat types. A major strength of our modelling approach is the ability to account for multiple factors affecting the range expansion of invasive species. Future research on the invasion dynamics of EUCOs would benefit from the inclusion of additional factors that we did not consider here. For example, research suggests that dispersal is likely to benefit from fragmented and small urban habitats (Eraud et al. 2011). Therefore, covariates that reflect the size and density of land cover clusters, in addition to their proportion within each cell, may be beneficial to include to achieve higher predictive accuracy. The density of free-ranging cats, the greatest predator of EUCOs, may also serve as an indicator of mortality affecting their ability to colonize and disperse (Eraud et al. 2011). While our models may have captured these effects in associated variables (e.g., the proportion of urban areas is likely a proxy for cat density (Flockhart et al. 2016)), refinement of covariates will likely improve the precision of these models and provide further information about the specific factors affecting dispersal. Combining various avian monitoring datasets using an integrative modelling approach would also provide researchers with data that covers the full annual cycle, improving our assessment of this invasive species' spread.

One limitation of our current models is that we have not explored variation in the detection probability. More specifically, we assume a binomial model for the number of detections within each grid cell in each year implying that surveys within each grid cell are independent and that the probability of observing doves is the same for all surveys. We made this decision to focus on developing models of the dynamic processes. However, it is possible, if not likely, that the probability of detection will vary both spatially and temporally. Detection is potentially affected by the density of doves within a cell, so that

doves are more likely to be observed where they are more abundant, and may also be affected by other predictors. For example, doves may visit feeders more regularly in areas where natural food resources are less abundant or where weather conditions are less conducive. Failing to account for this variation may impact the models of the dynamic processes if they are present, and we plan to develop further models to allow for variable detection in the future.

Our models extend the work of Bled et al. (2011) by individually accounting for the effect of neighbourhood occupancy and directional preference on initial colonization, persistence, and recolonization events. This approach allowed us to reconsider the assumption that assigns the same weight to neighbourhood effect and directional propagation. It also enables us to evaluate the effect of directional preference on invasion dynamics that are not accounted for by the environmental conditions of the recipient habitat. Our models highlight the importance of incorporating both neighbouring effects and environmental factors in the modelling and management of range-expanding species. Our results offer insights into the habitat preference and potential physiological tolerance of EUCCO, which may be helpful in understanding the mechanisms contributing to the success of this species and other avian invaders.

## **Acknowledgments**

We would like to thank the associate editor and two anonymous reviewers for their comments on a previous version of this manuscript. We are also grateful to Kerrie Wilcox, Manager of Project FeederWatch at Birds Canada, and all of the Project FeederWatch volunteers who have provided their time throughout the years. Simon Bonner is supported through a Discovery Grant through the National Science and Engineering Research Council (RGPIN-2016-04432) and through a Collaborative Research Team Award from the Canadian Statistical Sciences Institute.

## References

- Abellán, P., J. L. Tella, M. Carrete, L. Cardador, and J. D. Anadón (2017). Climate matching drives spread rate but not establishment success in recent unintentional bird introductions. *Proceedings of the National Academy of Sciences of the United States of America* 114:9385–9390.
- Banko, P. C., P. T. Oboyski, J. W. Slotterback, S. J. Dougill, D. M. Goltz, L. Johnson, M. E. Laut, and T. C. Murray (2002). Availability of food resources, distribution of invasive species, and conservation of a Hawaiian bird along a gradient of elevation. *Journal of Biogeography* 29:789–808.
- Bellard, C., P. Cassey, and T. M. Blackburn (2016). Alien species as a driver of recent extinctions. *Biology Letters* 12:20150623.
- Bled, F., J. A. Royle, and E. Cam (2011). Hierarchical modeling of an invasive spread: the Eurasian Collared-Dove *Streptopelia decaocto* in the United States. *Ecological Applications* 21:290–302.
- Bonter, D. N., and C. B. Cooper (2012). Data validation in citizen science: a case study from Project FeederWatch. *Frontiers in Ecology and the Environment* 10:305–307.
- Bonter, D. N., B. Zuckerberg, and J. L. Dickinson (2010). Invasive birds in a novel landscape: habitat associations and effects on established species. *Ecography* 33:494–502.
- Brooks, S. P., and A. Gelman (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* 7:434–455.
- Cohen, T., M. Hauber, T. Akriotis, P.-A. Crochet, G. Karris, A. Kirschel, F. Khoury, M. Menchetti, E. Mori, E. Per, L. Reino et al. (2022). Accelerated avian invasion into the Mediterranean region endangers biodiversity and mandates international collaboration.

- Journal of Applied Ecology* 59: 1440–1455.
- Coombs, C., A. Isaacson, R. Murton, R. Thearle, and N. Westwood (1981). Collared doves (*Streptopelia decaocto*) in urban habitats. *Journal of Applied Ecology* 18: 41–62.
- Davies, R. G., C. D. L. Orme, D. Storch, V. A. Olson, G. H. Thomas, S. G. Ross, T.-S. Ding, P. C. Rasmussen, P. M. Bennett, I. P. Owens, T. M. Blackburn, and K. J. Gaston (2007). Topography, energy and the global distribution of bird species richness. *Proceedings of the Royal Society B: Biological Sciences* 274:1189–1197.
- Davis, M. A. (2003). Biotic globalization: does competition from introduced species threaten biodiversity? *Bioscience* 53:481–489.
- de Valpine, P., D. Turek, C. J. Paciorek, C. Anderson-Bergman, D. T. Lang, and R. Bodik (2017). Programming with models: writing statistical algorithms for general model structures with NIMBLE. *Journal of Computational and Graphical Statistics* 26:403–413.
- Donati, M., K. Laroucau, M. Delogu, F. Vorimore, R. Aaziz, E. Cremonini, R. Biondi, C. Cotti, R. Baldelli, and A. Di Francesco (2015). Chlamydia psittaci in Eurasian collared doves (*Streptopelia decaocto*) in Italy. *Journal of Wildlife Diseases* 51:214–217.
- Eraud, C., J.-M. Boutin, D. Roux, and B. Faivr (2007). Spatial dynamics of an invasive bird species assessed using robust design occupancy analysis: the case of the Eurasian collared dove (*Streptopelia decaocto*) in France. *Journal of Biogeography* 34:1077–1086.
- Eraud, C., A. Jacquet, and P. Legagneux (2011). Post-fledging movements, home range, and survival of juvenile Eurasian Collared-Doves in Western France. *The Condor* 113:150–158.
- Flockhart, D., D. R. Norris, and J. B. Coe (2016). Predicting free-roaming cat population densities in urban areas. *Animal Conservation* 19:472–483.
- Fujisaki, I., E. V. Pearlstine, and F. J. Mazzotti (2010). The rapid spread of invasive

Eurasian Collared Doves *Streptopelia decaocto* in the continental USA follows human-altered habitats. *Ibis* 152:622–632.

- Gelman, A., and D. B. Rubin (1992). Inference from iterative simulation using multiple sequences. *Statistical Science* 7:457–472.
- Gurevitch, J., and D. K. Padilla (2004). Are invasive species a major cause of extinctions? *Trends in Ecology & Evolution* 19:470–474.
- Hagemeijer, W. J., and M. J. Blair (1997). The EBCC atlas of European breeding birds. Poyser, London 479.
- Hengeveld, R. (1988). Mechanisms of biological invasions. *Journal of Biogeography* 15:819–828.
- Hooten, M. B., and C. K. Wikle (2008). A hierarchical Bayesian non-linear spatio-temporal model for the spread of invasive species with application to the Eurasian Collared-Dove. *Environmental and Ecological Statistics* 15:59–70.
- Hooten, M. B., C. K. Wikle, R. M. Dorazio, and J. A. Royle (2007). Hierarchical spatiotemporal matrix models for characterizing invasions. *Biometrics* 63:558–567.
- Hudson, R. (1965). The spread of the Collared Dove in Britain and Ireland. *British Birds* 58:105–139.
- Keeley, J. E. (2006). Fire management impacts on invasive plants in the western United States. *Conservation Biology* 20:375–384.
- Koenig, W. D. (2020). What are the competitive effects of invasive species? Forty years of the Eurasian collared-dove in North America. *Biological Invasions* 22:3645–3652.
- MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, A. A. Royle, and C. A. Langtimm (2002). Estimating site occupancy rates when detection probabilities are less than one. *Ecology* 83:2248–2255.
- Meehan, T. D., N. L. Michel, and H. Rue (2019). Spatial modeling of Audubon Christmas

- Bird Counts reveals fine-scale patterns and drivers of relative abundance trends. *Ecosphere* 10:e02707.
- Morin, R. S., A. M. Liebhold, and K. W. Gottschalk (2009). Anisotropic spread of hemlock woolly adelgid in the eastern United States. *Biological Invasions* 11:2341–2350.
- Panella, N. A., G. Young, and N. Komar (2013). Experimental infection of Eurasian collared-dove (*Streptopelia decaocto*) with West Nile virus. *Journal of Vector Ecology* 38:210–214.
- Plummer, M. (2022). rjags: Bayesian Graphical Models using MCMC. URL <https://CRAN.R-project.org/package=rjags>.
- Poling, T. D., and S. E. Hayslette (2006). Dietary overlap and foraging competition between mourning doves and Eurasian collared-doves. *The Journal of Wildlife Management* 70:998–1004.
- R Core Team (2021). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Rappole, J. H., S. R. Derrickson, and Z. Hubalek (2000). Migratory birds and spread of West Nile virus in the Western Hemisphere. *Emerging Infectious Diseases* 6:319.
- Redding, D. W., A. L. Pigot, E. E. Dyer, Ç . H. Şekercioğlu, S. Kark, and T. M. Blackburn (2019). Location-level processes drive the establishment of alien bird populations worldwide. *Nature* 571:103–106.
- Romagosa, C., and T. McEneaney (1999). Eurasian collared-dove in North America and the Caribbean. *North American Birds* 53:348–353.
- Romagosa, C. M. (2002). Eurasian collared-dove: *Streptopelia decaocto*. Birds of North America, Incorporated.

- Romagosa, C. M., and R. F. Labisky (2000). Establishment and dispersal of the Eurasian Collared-Dove in Florida. *Journal of Field Ornithology* 71:159–166.
- Royle, J. A., and M. Kery (2007). A Bayesian state-space formulation of dynamic occupancy models. *Ecology* 88:1813–1823.
- Saâd, N., S. Hanane, M. D. El Hak Khemis, and K. Farhi (2021). Landscape composition governs the abundance patterns of native and invasive Columbidae species along an urban–rural gradient and contribute to their partitioning. *Biological Invasions* 23:2077–2091.
- Sala, O. E., F. S. Chapin, J. J. Armesto, E. Berlow, J. Bloomfield, R. Dirzo, E. Huber-Sanwald, L. F. Huenneke, R. B. Jackson, A. Kinzig, R. Leemans, et al. (2000). Global biodiversity scenarios for the year 2100. *Science* 287:1770–1774.
- Sax, D. F., and J. H. Brown (2000). Global Ecology and Biogeography - The paradox of invasion. *Global Ecology and Biogeography* 9:363–371.
- Scheidt, S. N., and A. H. Hurlbert (2014). Range expansion and population dynamics of an invasive species: The Eurasian collared-dove (*Streptopelia decaocto*). *PloS one* 9:e111510.
- Slager, D. L. (2020). Seasonal and directional dispersal behavior in an ongoing dove invasion. *Journal of Avian Biology* 51.
- Stimmelmayer, R., L. Stefani, M. Thrall, K. Landers, F. Revan, A. Miller, R. Beckstead, and R. Gerhold (2012). Trichomonosis in free-ranging Eurasian collared doves (*Streptopelia decaocto*) and African collared dove hybrids (*Streptopelia risoria*) in the Caribbean and description of ITS-1 region genotypes. *Avian Diseases* 56:441–445.

- Vehtari, A., A. Gelman, and J. Gabry (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing* 27:1413–1432.
- Watanabe, S (2010). Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory. *Journal of Machine Learning Research* 11:3571–3594.
- Wells, J. V., K. V. Rosenberg, E. H. Dunn, D. L. Tessaglia-Hymes, and A. A. Dhondt (1998). Feeder Counts as Indicators of Spatial and Temporal Variation in Winter Abundance of Resident Birds (Contajes en Comederos Como Indicadores de Variaciones Espaciales y Temporales en la Abundancia Invernal de Aves Residentes). *Journal of Field Ornithology* 69: 577–586.