Hierarchical modeling of an invasive spread: the Eurasian Collared-Dove *Streptopelia decaocto* in the United States

FLORENT BLED,[1,3] J. ANDREW ROYLE,[2] AND EMMANUELLE CAM[1]

[1]*Laboratoire Évolution et Diversité Biologique, Université Paul Sabatier, 118 Route de Narbonne, 31062 Toulouse, Cedex 9 France*
[2]*U.S. Geological Survey, Patuxent Wildlife Research Center, Laurel, Maryland 20708 USA*

Abstract. Invasive species are regularly claimed as the second threat to biodiversity. To apply a relevant response to the potential consequences associated with invasions (e.g., emphasize management efforts to prevent new colonization or to eradicate the species in places where it has already settled), it is essential to understand invasion mechanisms and dynamics. Quantifying and understanding what influences rates of spatial spread is a key research area for invasion theory. In this paper, we develop a model to account for occupancy dynamics of an invasive species. Our model extends existing models to accommodate several elements of invasive processes; we chose the framework of hierarchical modeling to assess site occupancy status during an invasion. First, we explicitly accounted for spatial structure and how distance among sites and position relative to one another affect the invasion spread. In particular, we accounted for the possibility of directional propagation and provided a way of estimating the direction of this possible spread. Second, we considered the influence of local density on site occupancy. Third, we decided to split the colonization process into two subprocesses, initial colonization and recolonization, which may be ground-breaking because these subprocesses may exhibit different relationships with environmental variations (such as density variation) or colonization history (e.g., initial colonization might facilitate further colonization events). Finally, our model incorporates imperfection in detection, which might be a source of substantial bias in estimating population parameters.

We focused on the case of the Eurasian Collared-Dove (*Streptopelia decaocto*) and its invasion of the United States since its introduction in the early 1980s, using data from the North American BBS (Breeding Bird Survey). The Eurasian Collared-Dove is one of the most successful invasive species, at least among terrestrial vertebrates. Our model provided estimation of the spread direction consistent with empirical observations. Site persistence probability exhibits a quadratic response to density. We also succeeded at detecting differences in the relationship between density and initial colonization vs. recolonization probabilities. We provide a map of sites that may be colonized in the future as an example of possible practical application of our work.

Key words: Breeding Bird Survey, BBS; colonization; detectability; Eurasian Collared-Dove; hierarchical modeling; initial colonization; invasive species; recolonization; site occupancy; Streptopelia decaocto.

INTRODUCTION

Regularly claimed as the second threat to biodiversity and threatened species after habitat destruction (Glowka et al. 1994 [cited in Williamson (1999)], Wilcove et al. 1998), invasive species are defined as species not naturally present in a geographic area that have been introduced by man, and that have succeeded in establishing and colonizing this area. Even if only a small fraction of transported species become established and, of these, generally only about 1% become pests (Williamson 1996), events of invasion are not uncommon. Invasive species are present in a wide range of

taxa, from microorganisms to vertebrates, plants, and invertebrates as compiled by the IUCN Invasive Species Specialist Group in the Global Invasive Species Database (*available online*).[4] Over the last 500 years, invasive species have been estimated to have come to dominate 3% of the Earth's ice-free surface (Mack [1985] cited in Mooney and Cleland 2001). Moreover, in Europe alone, 10 670 species (fauna and flora) have been considered as "invasive exotic species" in the framework of the European Union DAISIE program (Delivering Alien Invasive Species Inventories for Europe; *available online*).[5]

Although invasive species are generally presented as a strong threat to indigenous species, many studies on this

[4] ⟨http://www.issg.org/database⟩
[5] ⟨http://www.europe-aliens.org/⟩

subject are just correlative. They cannot conclusively determine if invasive species are responsible for the loss of biodiversity or simply respond to what has caused this loss, e.g., habitat alteration (Gurevitch and Padilla 2004). Invasive exotic species are causing changes in many ecological systems worldwide, and are altering many communities and ecosystems (Gurevitch and Padilla 2004). However, because these species are not the only element affecting biodiversity and because they usually co-occur with other threats, it is essential to understand invasion mechanisms and dynamics in a wider context of global change to develop a relevant response to the potential consequences associated with invasions: whether to emphasize management efforts to prevent new colonization (i.e., try to contain the invasion) or to eradicate the species in places where it has already settled. Quantifying and understanding what influences rates of spatial spread is a key research area for invasion theory (e.g., Skellam 1951, Okubo 1980, Andow et al. 1990, Kot et al. 1996, Neubert and Caswell 2000, Wikle 2003). Invasive spread may exhibit important features such as the presence of a preferential direction for the spread (see Wikle 2003, Hastings et al. 2005, Morin et al. 2009), and may depend on the distance between "suitable" locations. Determining these characteristics is essential to take relevant management decisions.

The Eurasian Collared-Dove (*Streptopelia decaocto*) is one of the most successful invasive birds in North America (Romagosa and Labisky 2000). This species was introduced in North America during the early 1980s through Florida. It has been hypothesized that invasion started in the late 1970s from the Bahamas, where Eurasian Collared-Doves escaped from captivity, established a wild population, and then reached Florida (Smith 1987). This species has high colonization capacities. It invaded Europe in less than 30 years. In fact, in the case of the United States, it took less than 25 years for the Eurasian Collared-Dove population to reach the west coast (Dunn and Alderfer 2006, Hooten and Wikle 2008). Even if this species has not been proven to be a direct threat to other species or ecosystems, it is still logically considered as a potential threat because it is an invasive species (Hengeveld 1993) and therefore might compete with other species such as Mourning Doves (*Zenaida macroura*), White-winged Doves (*Zenaida asiatica*), or Common Ground-Doves (*Columbina passerine*) (Romagosa and McEneaney 2000, Romagosa 2002). It could also be a disease vector (Romagosa and Labisky 2000).

Hierarchical modeling has been developed and used to estimate site occupancy (Royle and Kéry 2007). In this framework, hierarchical models are typically based on three components. The first component corresponds to the data, the observed quantity. This component is defined conditional on a second component, the state variable (e.g., true occupancy status); the relationship between these components is accounted for by parameters (e.g., detection probability). A typical recent hierarchical model in population ecology would constitute an ecological process underlying occupancy (e.g., balance between extinction and colonization) and a level corresponding to the observation process (which depends on detectability; Royle and Dorazio 2008). Because non-detection is not equal to absence, this class of models (i.e., including an observation process) is essential when dealing with detection–non-detection data (usually improperly named presence–absence data). Indeed, not accounting for detection issues may lead to substantial bias in estimating population parameters (MacKenzie et al. 2002). This type of data is typically the one available for invasive species, where detectability might be an issue, especially during the beginning of the colonization. Hierarchical models are powerful and flexible and are used in many problems, with many applications to public health and ecology (e.g., Banerjee 2003, Waller and Gotway 2004, Lawson 2006, Clark 2007, Gelman and Hill 2007, Lawson 2008, Ntzoufras 2009). However, to our knowledge, they have not been used consistently to assess invasive species dynamics (even if some examples exist; Hooten et al. 2007, Hooten and Wikle 2008). Importantly, detectability issues (MacKenzie et al. 2006) have seldom been taken into account when addressing hypotheses about vertebrate species distribution or the invasive dynamics of vertebrate species (e.g., Ibarra et al. 2005, Ficetola et al. 2007, Leprieur et al. 2008). This might be an extremely important issue, as ignoring it may lead to underestimation of the actual colonized area and provide erroneous information about the key locations requiring a special regulatory effort (e.g., locations where the invasion is starting and where settlement has not yet happened).

Hierarchical modeling provides a convenient means of incorporating biological hypotheses of population dynamics in an explicit way. With a hierarchical approach, it is straightforward to express population dynamics parameters (colonization and persistence probabilities) as functions of variables such as density or reproductive success. Nowadays, many studies use phenomenological models (i.e., models accounting for spatiotemporal patterns of species detection without incorporating specific hypotheses about ecological processes governing species distribution) and emphasize simple descriptions or patterns in data. Such models may lead to satisfactory descriptions of data but are not necessarily easy to interpret biologically. Fewer studies have focused on the development of mechanistic approaches (Bennett et al. 2001), i.e., models accounting for species presence and detection, with presence expressed according to explicit ecological and biological hypotheses about the dynamics of species distribution. Recent improvements have been made to develop hierarchical models that account for uncertainty by encompassing both spatial and time dimensions (Hooten et al. 2007, Royle and Kéry 2007, Hooten and Wikle 2008), and that directly include

scientific insight in model processes (e.g., reaction–diffusion motivation; Wikle 2003).

In this paper, we develop a new model based on that of Royle and Kéry (2007). In the framework of Bayesian modeling, we extend this model to account more accurately for invasive colonization processes: we consider an explicit spatial structure in a dynamic model. This ecological process accounts for the density of occupied sites in the neighborhood. It also considers the influence of distance among sites. Indeed, we expect the occupancy status of close neighboring sites to have a stronger influence on site persistence (or colonization) probability of a given site than remote sites. The ecological process also explicitly accounts for the possibility of a directional spread and allows detection of the direction of this spread, if any. To our knowledge, this is the first time that spatial structure has been included in a spatiotemporal occupancy model with such an explicit structure formulation. The manner in which the potential anisotropy or directional spread is integrated in the model is a breakthrough, especially in the framework of occupancy dynamics modeling. We draw a distinction between initial colonization and recolonization, because an initial colonization might facilitate further colonization events. Our model also includes an observation process, conditional on the underlying ecological process, to deal with detectability issues. Importantly, we consider how previous detections may influence, and potentially improve, detectability. Eurasian Collared-Dove data were collected in the framework of the North American Breeding Bird Survey (Robbins et al. 1986) with the help of volunteer observers. In this case, the presence of an exotic species, easily mistaken for other dove species, might not be properly detected the first time it appears. However, with repeated detection events over consecutive years and the accumulation of external confirmation of this "unusual" species in the region, observers might improve their ability to detect this species, through better identification and more likely because of knowledge that this species is present in the area. This is the reason why we incorporate a potential "learning effect" in the observational process, and the resulting gain in detectability.

MATERIALS AND METHODS

Eurasian Collared-Dove data are based on the Breeding Bird Survey (BBS; Robbins et al. 1986). This program has been monitoring avian populations in North America since 1966. Observers are assigned to a number of routes, where they stop 50 times. "Each survey route is 24.5 miles long with stops at 0.5-mile intervals. At each stop, a 3-minute point count is conducted. During the count, every bird seen within a 0.25-mile radius or heard is recorded.... Over 4100 surveyed routes are located across the continental United States and Canada" (BBS web site).[6] For each

[6] ⟨http://www.pwrc.usgs.gov/bbsapps/⟩

survey route, these raw data include the number of stops where individuals of a given species have been detected, and the total number of individuals detected. Because we are interested in occupancy status (not in abundance), we decided to perform our analyses on the number of stops where the Eurasian Collared-Dove has been detected. Joseph et al. (2006) have shown that abundance methods lead to a larger variation in estimations than detection–absence methods. Moreover, they have shown that presence–absence surveys were more optimal for low budget and low detectability. These points and the fact that detection–non-detection data are often easier to obtain led us to choose this type of data (depending on conditions, one can decide to use abundance, as in Wikle [2003]).

We used data from 1986 (first detection of the European Collared-Dove in the United States in the BBS data set) to 2006. We developed a grid over the BBS map region, merging data from all routes contained in the same grid cell. This grid goes from the point of spatial latitude and longitude coordinates (24°; −129°) to the point (57°; −51°). The side length of a cell is equal to 1°. We did not take into account grid cells that only correspond to ocean. We did not consider grid cells that did not have at least one neighbor cell; therefore we considered a total of 1259 cells.

We define the neighborhood of a cell i as the first (N_1) and second (N_2) layers that surround this cell. Although cells in the first layer N_1 share a border with cell i, cells in the second layer N_2 are separated from cell i by one cell, as shown in Fig. 1. We expect the influence of one site (i.e., cell) occupancy status on another to decrease with increasing distance between the two sites. This means that the occupancy status of a close site should have a stronger influence on the probability of occupancy of a given site that sites that are located farther away, as in a diffusive process.

THE MODEL

Occupancy state model

We considered occupancy data obtained by repeated sampling of $i = 1, 2, \ldots, M$ spatial units (i.e., patches, or "sites," depending on the context; here cells), over $t = 1, 2, \ldots, T$ periods of time. Usually, these periods of time refer to significant biological seasons, depending on the species of interest. For reference, all parameters used in our model are summarized in Table 1.

The dynamics of the occupancy status will be accounted for by two parameters: persistence ϕ or its complement, local extinction: $(1 - \phi)$, and colonization γ. Both can be indexed by time and/or site depending on the question of interest. For example, we may consider that site persistence (i.e., a cell staying occupied) varies over time depending on the growth rate in a population (stable, growing, or declining). If the population is declining, the persistence of sites will decrease as the number of individuals decreases. On the other hand,
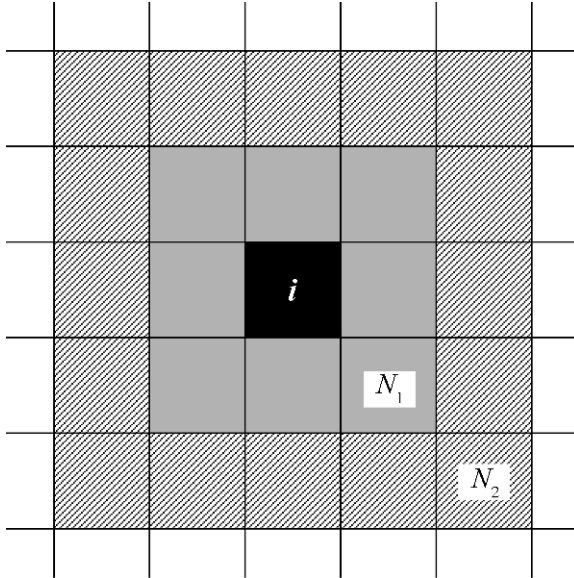
FIG. 1. Representation of the two layers of neighboring cells, used in our model to estimate local density. Cell "i" (in black) is the cell where occupancy state depends on the proportion of occupied cells in both the first layer (N_1, gray cells) and second layer (N_2, hatched cells) of cells that surround cell i and constitute its neighborhood.

colonization probability may vary among sites with different characteristics.

Let $Z_{i,t}$ be the occupancy state of cell i in year t. If the cell i is occupied at time t, then $z_{i,t} = 1$; otherwise, $z_{i,t} = 0$. We are interested in the probability of site occupancy $\mu_{i,t} = \Pr(Z_{i,t} = 1 | z_{i,t-1})$ (the probability that a cell is occupied is conditional on the cell's occupancy state in the previous year). As in population demographic

processes of survival and recruitment, here, local extinction and colonization can be used as parameters to model changes in occupancy over time.

Let's define ϕ_t the probability that an occupied site "survives" from time t to $t + 1$, that is, given that it was occupied at time t, the probability that it is occupied again at time $t + 1$, i.e., $\phi_t = \Pr(Z_{i,t+1} = 1 | z_{i,t} = 1)$. Where MacKenzie et al. (2003) used local extinction probability $(1 - \phi_t)$, we prefer to use its complement, persistence probability. In addition, let γ_t stand for the local colonization probability, from time t to $t + 1$, i.e., γ_t $(=\Pr(Z_{i,t+1} = 1 | z_{i,t} = 0)$. In this model, as in metapopulation systems, local colonization can be viewed as the analog of the recruitment process of individuals in populations. $Z_{i,t}$ is a Bernoulli variable with expected value $\mu_{i,t}$:

$$Z_{i,t} | z_{i,t-1} \sim \text{Bern}(\mu_{i,t}) \qquad (1)$$

where

$$
\begin{aligned}
\mu_{i,t} &= \Pr(Z_{i,t} = 1 | Z_{i,t-1} = z_{i,t-1}) \\
&= z_{i,t-1} \phi_{t-1} + (1 - z_{i,t-1})\gamma_{t-1}. \qquad (2)
\end{aligned}
$$

In this model, the occupancy status at time t depends on previous occupancy status. This model can be easily extended. In particular, we consider two different stages in the colonization process, and we structure the dynamics parameters to take into account the spatial structure of a point process.

Recolonization reparametrization

Here, all cells that are not occupied have the same probability of being colonized, i.e., site colonization is random and does not depend on the previous status of the cell, whether it has been occupied and then deserted,

TABLE 1. Parameters used in the model for occupancy dynamics of an invasive species, the Eurasian Collared-Dove (*Streptopelia decaocto*).

Parameter	Description	Algorithm notation
$Z_{i,t}$	occupancy state of cell i in year t	z[i, t − 1]
$\mu_{i,t}$	site occupancy probability	muZ[i, t]
$\phi_{i,t}$	persistence parameter	phi[i, t − 1]
$\gamma_{i,t}$	initial colonization parameter	gamma[i, t − 1]
$\theta_{i,t}$	recolonization parameter	theta[i, t − 1]
$A_{i,t}$	availability of a site for initial colonization	A[i, t − 1]
$D_{i,t}$	local density	D[i, t − 1]
w_{ij}	weight reflecting importance of occupancy status of cell j on future occupancy status of site i	computed in the algorithm
a_t, b_t, c_t	intrinsic persistence, initial colonization, and recolonization parameters, respectively	lphi0[t], lgamma0[t], ltheta0[t]
$\phi_1, \gamma_1, \theta_1$	linear factors of persistence, initial colonization, and recolonization parameters to local density, respectively	lphi1, lgamma1, ltheta1
$\phi_2, \gamma_2, \theta_2$	quadratic factors of persistence, initial colonization, and recolonization parameters to local density, respectively	lphi2, lgamma2, ltheta2
$K_{i,t}$	number of replications in the cell i at time t	K[i, t]
p	detection probability	p[i, t]
α, β	relative contributions of layers 1 and 2 density to local density, respectively	alpha, beta
δ_1	direction of propagation	dirSpread
$A'_{i,t}$	availability for a first detection	Aprim[i, t]
λ_1	first detection probability	ObsInit
λ_2	gain in detectability after a previous detection	DeltaObs

or has never been colonized at all. However, it is interesting to draw a distinction between two components in the colonization process. We can distinguish between a process of "first colonization" (i.e., the site has never been occupied before; this could be considered as the creation of the site in the data set), and "recolonization" (i.e., colonization of a site after a previous "extinction"). This is especially relevant for invasive species, in which initial colonization by some individuals might facilitate further colonization events. For example, in the case of invasive plants, seeds can be left by initial colonizers and can germinate long after initial individuals have disappeared, leading to a new and facilitated colonization event (Harrod and Reichard 2001, Keeley 2006). Although this example is not directly valid for animals, the model proposed in this paper can be applied to plant species with minor modifications and therefore accounts for this particular situation. Mechanisms for first colonization and recolonization can also differ; e.g., first colonization could be related to a diffusion type of process, whereas recolonization could be related to the intrinsic dynamics of the local population. In habitat selection theory, one hypothesis is that sites are expected to be chosen based upon their quality; in this case, "better" sites should be occupied (and reoccupied if abandoned) first, leading to a distinction between first and recolonization probabilities. These two subprocesses of colonization are modeled by splitting the colonization parameter γ_t into a first colonization parameter that will keep the name γ_t, and a recolonization parameter θ_t. We then have to define a new auto-covariate $A_{i,t}$ that will express the availability of a site for first colonization. $A_{i,t} = 1$ if the site has never been occupied or colonized before (i.e., for the period T, $A_{i,T} = 1$ if $\Sigma_{t=1}^{T} z_{i,t} = 0$), $A_{i,t} = 0$ otherwise. Therefore, we can formally express $A_{i,t}$ as the indicator function $A_{i,t} = \prod_{k=1}^{t-1}(1 - z_{i,k})$, ($k \in \{1, 2, \dots, t-1\}$). As a consequence, sites will have different colonization probabilities depending on whether $A_{i,t} = 1$ or $A_{i,t} = 0$: respectively, γ_t and θ_t.

Consequently, our model becomes

$$Z_{i,t+1} \mid z_{i,t-1} \sim \text{Bern}(\mu_{i,t+1}) \qquad (3)$$

with

$$\mu_{i,t+1} = \phi_t z_{i,t} + \gamma_t(1 - z_{i,t})A_{i,t} + \theta_t(1 - z_{i,t})(1 - A_{i,t}). \quad (4)$$

With this model, site colonization can be assessed at two different levels depending on whether it has already been colonized or not. We have a dynamic model, but we wish to consider the spatial structure by including space in the dynamic parameters.

Spatial structuration

In the model as it stands, the occupancy status of each cell is independent of its environment (i.e., the location of occupied cells and the spatial structure of occupancy). However, in a large number of cases, we may want to consider the possible impact of the spatial structure of occupancy of several sites (i.e., in a given area) on the fate of individual sites. We may be interested in the influence of variation in climate on a species' spatial distribution, or the influence of the spatial structure of the landscape and the changes that may occur in environmental factors due to human activities. Despite the obvious implications of such processes for both basic and applied ecology, the statistical framework for modeling of spatiotemporal occupancy systems is not well developed, despite some recent efforts, as in Hooten and Wikle (2008) or Zhu et al. (2005). Here, we express the spatial structure over the dynamic parameters of persistence, first colonization, and recolonization.

Let N_i represent the set of cells that are neighbors of the cell i, and let n_i be the number of neighbors of cell i (i.e., cardinality of N_i). Then we can define a spatiotemporal autocovariate $D_{i,t}$ as

$$D_{i,t} = \sum_{j \in N_i} z_{j,t} \, w_{ij} \qquad (5)$$

where w_{ij} is a weight that can be used to specify a difference in the influence of a neighbor site j on the fate of site i according to the locations of sites. It can be based on a simple connection net, where $w_{ij} = 1$ if site j is connected to site i (for example, if the distance between i and j is under the defined/determined threshold "influence" distance), and $w_{ij} = 0$ otherwise (i.e., if site i is not in the influence area of j). We can also set w_{ij} to weight $D_{i,t}$ by the inverse distance of sites j to i, or in any other way, depending on the decisions made to define the connections network in a relevant manner according to the studied topic. Although $D_{i,t}$ is a measurement of local occupancy, it is important to note that it can be viewed as a surrogate for local density in the vicinity of cell i at time t (Royle and Nichols 2003).

We included this spatiotemporal autocovariate in the model by allowing our dynamics parameters to depend on the autocovariate. Let's take the example of the persistence parameter ϕ. It can simply be expressed as a function of $D_{i,t}$ as follows:

$$\text{logit}(\phi_{i,t}) = a_t + \phi_1 D_{i,t} + \phi_2 D_{i,t}^2. \qquad (6)$$

It should be noted that ϕ is now indexed by both time period t and site i. We used a quadratic function of $D_{i,t}$ to estimate $\phi_{i,t}$; this function allows ϕ to vary in a way that may lead to a peak at intermediate values of $D_{i,t}$. Such a pattern may account for density dependence or the "Allee effect" (e.g., Courchamp et al. 1999, Keyser et al. 2005). Moreover, this formulation permitted us to detect a difference between "random" colonization, as might be expected in a stable metapopulation, and a dynamic of diffusive spread such as might happen in an expanding population. Indeed, a_t will represent the "intrinsic persistence parameter," i.e., the parameter describing what is happening with no neighbor "effect," or when site i does not have any occupied neighbor. We

decided to allow this "intrinsic persistence parameter" to vary over time, because it can be hypothesized that a_t differs when the population is growing, stable, or decreasing. In contrast, ϕ_1 and ϕ_2 represent the influence of occupancy of neighbors on persistence probability, and therefore they embody diffusive or dynamic spread due to gradients in local density or occupancy.

To summarize, our model can be described as follows:

$$Z_{i,t+1} \mid z_{i,t} \sim \text{Bern}(\mu_{i,t+1}) \qquad (7)$$

with

$$\mu_{i,t+1} = \phi_{i,t} z_{i,t} + \gamma_{i,t}(1 - z_{i,t})A_{i,t} + \theta_{i,t}(1 - z_{i,t})(1 - A_{i,t}) \qquad (8)$$

where

$$\begin{cases} \text{logit}(\phi_{i,t}) = a_t + \phi_1 D_{i,t} + \phi_2 D_{i,t}^2 \\ \text{logit}(\gamma_{i,t}) = b_t + \gamma_1 D_{i,t} + \gamma_2 D_{i,t}^2 \\ \text{logit}(\theta_{i,t}) = c_t + \theta_1 D_{i,t} + \theta_2 D_{i,t}^2 \end{cases} \qquad (9)$$

depending on

$$D_{i,t} = \sum_{j \in N_i} z_{j,t} w_{ij}. \qquad (10)$$

The ecological component of our model now has both time and spatial dimensions. although the time dimension is supported by the link between $z_{i,t}$ and $Z_{i,t+1}$, the spatial dimension is encompassed in the weights matrix $\mathbf{W}$ (where w_{ij} is the influence of site j on site i). In the following section, we consider how observations are related to the ecological process component of the model and we develop an explicit model for imperfect observation of this process.

Observation model

One of the principal sources of uncertainty in monitoring data is that due to imperfect detection (or "detectability") of species. That is, a species might be present at some point in space and time and yet go undetected. Many modeling approaches and strategies have been devised for dealing with this issue (Williams et al. 2002, MacKenzie et al. 2006).

The simplest way of incorporating detectability (i.e., the probability of contacting an individual of the species given that the species is present in the sites/area sampled) is to consider sampling protocol where a cell i has been visited repeatedly. Then $K_{i,t}$ defines the number of visits or replications in the cell i at time t, and p is the probability that a species or individual is detected during one stop on a road if it is present. $K_{i,t}$ corresponds to the number of routes nested in cell i times the number of stops per route (i.e., 50). $Y_{i,t}$ is the total number of stops at which a species or individual has been detected in cell i at time t during the $K_{i,t}$ visits. Then, observation $Y_{i,t}$ follows a binomial with $K_{i,t}$ trials and a probability $pz_{i,t}$:

$$y_{i,t} \sim \text{Bin}(K_{i,t}, pz_{i,t}). \qquad (11)$$

This means that if the cell is unoccupied, then $Y_{i,t}$ is equal to zero. Otherwise, the binomial probability is equal to the detection probability. $K_{i,t}$ can be viewed as analogous to the monitoring effort. When the number of visits, $K_{i,t}$, increases, so does the global detection probability in the corresponding cell. Of course, p can be indexed by time and/or site if needed, depending on the question of interest. We have now a full time–space hierarchical model with both ecological and observation processes.

Model adaptations

The model developed previously is a general model that can be easily modified and adapted to a large set of problems, each adaptation having to match the specific questions addressed. Considering the specificities of our data and our questions of interest (detection of invasive spread characteristics, improvement of detectability after a first detection), we made some adjustments to the core model previously described.

Spatial structure.—One of the main objectives of our work is to provide a model with temporal and a spatial dimensions. To do this, it is important to have a clear and logical definition of the spatial structure. Here, we were interested in the impact of distance between sites on occupancy.

We decided to use the proportion of occupied neighbors in the first and second layers, respectively, $D_{1i,t}$ and $D_{2i,t}$, as estimators of local density for the cell i.

The proportion of occupied cells in the first layer is just the mean of the number of occupied neighbor sites j among the n_i sites neighboring the cell i:

$$D_{1i,t} = \frac{\sum_{j \in N_{1i}} w_{ij} z_{j,t}}{n_{1i}}. \qquad (12)$$

We defined density in the second layer in the same way:

$$D_{2i,t} = \frac{\sum_{j \in N_{2i}} w_{ij} z_{j,t}}{n_{2i}}. \qquad (13)$$

Then, local density $D_{i,t}$ of the site i at time t is a weighted sum of relative densities in the first and the second layers: $D_{i,t} = \alpha D_{1i,t} + \beta D_{2i,t}$.

Therefore, α and β correspond to the relative contribution of each layer of neighbors. They are estimated by MCMC (Markov chain Monte Carlo), like the dynamic parameters. At this stage in model formulation, the weight w_{ij} is equal to 1 because we consider each site of a layer to be equivalent to the others. According to our hypothesis that close sites should be more influential than distant ones, we expect α to be higher than β.

Anisotropy or directional spread.—Anisotropy is the property of being directionally dependent. Invasion dynamics can be expected to exhibit such a property. The propagation of a species may be governed by a
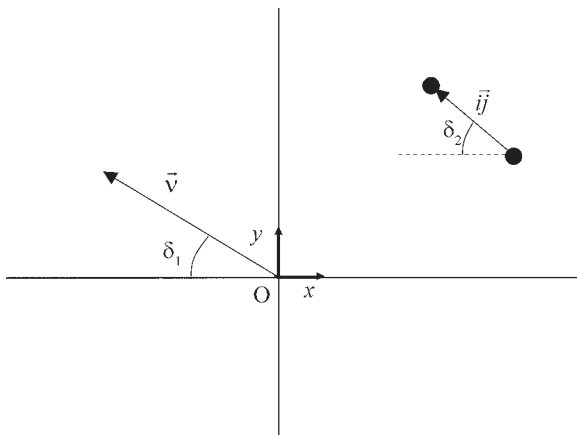
FIG. 2. Diagrammatic representation of points and vectors used in anisotropy modeling, where δ_1 is the angle made by the vector of invasion propagation $\vec{v}$ and the abscissa and δ_2 is the angle made by vector $\overrightarrow{ij}$ (going from site i to site j) and the abscissa; O represents the origin.

specific environmental gradient (e.g., density, temperature, humidity), leading to an oriented spread instead of a simple diffusive expansion. We added this element to our model. Let's consider a spread going from north to south. In this case, the occupancy status of a site j located north of a site i should be more influential than a site located east or west, and even more than a site located south.

Set a coordinate plane with standard basis, (O, x, y) (Fig. 2), and define δ_1 as the angle made by vector $\vec{v}$ (vector indicating the direction of propagation) and the abscissa, and δ_2 as the angle made by the vector of $\overrightarrow{ij}$ (going from site i to site j) and the abscissa.

As previously explained, the weights w_{ij} can be used to specify the spatial structure. We used this property to estimate the impact of the direction of propagation on dynamics parameters. We express w_{ij} depending on the direction of the spread and the relative position of site j compared to i, which means that we have to express w_{ij} as a function of angle Δ made by vectors $\vec{v}$ and $\overrightarrow{ij}$, as shown in Fig. 2. It is straightforward to show that $\Delta = \delta_1 - (\pi + \delta_2)$. Then, we find that

$$w_j \propto 1 + \cos(\delta_1 - \pi - \delta_2) \tag{14}$$

and

$$w_j \propto 1 + \cos(\delta_1 - \pi - \delta_2)$$
$$\propto 1 - \cos(\delta_1)\cos(\delta_2) - \sin(\delta_1)\sin(\delta_2).$$

We calculate δ_2 from coordinates of sites i and j, respectively, (x_i, y_i) and (x_j, y_j). We then have

$$\cos(\delta_2) = \frac{x_j - x_i}{\sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}} \tag{15}$$

and

$$\sin(\delta_2) = \frac{y_j - y_i}{\sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}}. \tag{16}$$

Therefore, the only unknown parameter is δ_1 (indicating the angle of propagation), which will be estimated by MCMC simulations.

We now have a hierarchical time–space model, with an ecological process accounting for the impact of distance among sites through α and β, and the direction of a spread and the relative position of site j to i, through δ_1.

Observation process.—Concerning the observation process, we dealt with a problem that may occur due to the way in which raw data are collected in the BBS. These data are collected by volunteers and are subject to various types of uncertainty, including observer error (Sauer et al. 1994). We think that it is interesting to modify the observation process because of possible misidentifications between Eurasian Collared-Dove and Ringed Turtle-Dove, for example, and the propensity of the observer to consider unlikely the presence of Eurasian Collared-Dove because it is an invasive species (i.e., non-endemic species). In fact, we think that a previous confirmed observation may lead to larger subsequent detection probability. In the framework of invasive species and volunteers, the underlying idea is that people collecting data may not identify a new species that is not supposed to be there (nonindigenous), and may easily confound it with another. However, as soon as this species has been detected without doubt, detection skills of observers may improve, and they may no longer consider the presence of this exotic species unlikely. We note that this is analogous to a "behavioral response" in classical capture–recapture modeling (Otis et al. 1978, Williams et al. 2002: Chapter 14).

We consider the previous observation process

$$y_{i,t} \sim \text{Bin}(K_{i,t}, p_{i,t} z_{i,t}) \tag{17}$$

but this time we allow detection probability p to vary with time and site. We can accommodate this behavioral response by including an effect on detection probability of prior detection of the species:

$$\text{logit}(p_{i,t}) = \lambda_1 + \lambda_2(1 - A'_{i,t}). \tag{18}$$

Here $A'_{i,t}$ corresponds to the availability for a first detection. If the species has never been previously detected in patch i before time t, then $A'_{i,t} = 1$. Otherwise, $A'_{i,t} = 0$. Then λ_1 is the probability of first detection (the species has never been detected before), and λ_2 is the gain in detectability due to a previous detection.

Our model now incorporates an observation process that considers improvement in detectability. Consequently, we can estimate a part of the error due to observer inexperience.

Bayesian analysis and implementation in WinBUGS

Because of the conditional specification of this time–space dynamic site occupancy model, the model is well

suited to Bayesian analysis via Markov chain Monte Carlo. The model parameters were estimated using Gibbs sampling (Casella and George 1992), which is based on drawing samples of each unknown quantity from their "full-conditional" distributions, i.e., the distribution of a parameter conditional on all other unknown quantities and the data (Royle and Kéry 2007). In particular, implementing our model with a software program such as WinBUGS is straightforward. For the analysis presented in this publication, models were implemented under the free software package WinBUGS 1.4.3 (Lunn et al. 2000) called from R 2.9.0 (R Development Core Team 2008) using the R add-on library R2WinBUGS (Sturtz et al. 2005). We ran three chains and based our inference on 100 000 samples from the posterior distribution of parameters, after 20 000 discarded iterations. The code for this model is provided in the Supplement. All estimates are presented with standard deviation.

To validate our model, we compared probabilities of detecting at least one individual in a cell for each year from 1986 to 2005 to the detection data (i.e., the corresponding cell has been detected or not to be occupied) of the following year (i.e., from 1987 to 2006) using the R package "ROCR" by Sing et al. (2009). We provide the result of the area under the curve (AUC) for the corresponding ROC diagrams for the last 15 years.

RESULTS

With a mean AUC for the last 15 years equal to 0.72 $\pm$ 0.14, reaching 0.76 $\pm$ 0.07 for the last 8 years, our model can be considered to be a fair estimation of the invasion process.

Spatial structure

Distance.—With respective values of $\alpha = 0.79 \pm 0.12$ and $\beta = 0.87 \pm 0.09$, weights for layers 1 and 2 do not reveal a clear prevalence of one layer over the other. Surprisingly, and in opposition to our predictions, this means that we failed to detect a clear influence of distance on dynamics parameters, at least at the scale considered here.

Invasion spread direction.—We detected a preferential direction for the invasion spread with $\delta_1 = 140.3° \pm 14.9°$. This approximately corresponds to a spread going from southeast to northwest. This matches empirical observations that indicate a first phase of slow colonization toward the north, followed by a second phase of rapid expansion toward the west, as shown in Fig. 3. The invasion progresses mainly from east to west, with a component south to north. Starting in Florida, this invasion could not have proceeded other than northerly. After that, a limitation due to latitude, with temperatures not being suitable for this species over the northern limit of the invasion front, is likely to appear.

In Fig. 3, we provide an example that permits us to compare our estimations for sites that have a high probability of being colonized (the hatched area indicates the probability of being colonized is $p > 0.5$) in 2006 based on 2005 data, and actual occupancy observed in 2006 (gray cells). We see that estimations are accurate and globally correspond to the actual observed occupancy. This type of prediction map could be used for management decisions to contain the propagation of an invasive species. Here, using the estimation of spread direction and the impact of density, we illustrate the western invasion process front and the likely recolonization of some abandoned sites.

Density.—Site persistence probability (black line, Fig. 4) increases at low density and remains high at intermediate densities before dropping at the highest densities. Initial colonization probability (dashed line, Fig. 4) is globally low but exhibits a quadratic relationship with density, with virtually no initial colonization at low and high density, and a slightly higher value at a short range of medium densities. Recolonization probability (dotted line, Fig. 4) is quite high at low to intermediate densities, and then decreases at higher densities.

Detectability

The initial detection probability estimate at the stop level is equal to 0.03 ($\lambda_1 = -3.44 \pm 0.28$). Surprisingly, after a previous detection, detection probability drops around 0.01 ($\lambda_2 = -1.633 \pm 0.5395$). With such a low detection probability, not accounting for this issue would have led to underestimation of occupancy probabilities and therefore the area occupied.

DISCUSSION

We have developed a time–space hierarchical model accounting for an invasion process via estimation of site persistence probability, initial colonization, and recolonization, and their relationship with density. We were able to detect the direction of an invasive spread, but we have not detected any influence of distance between occupied sites on dynamics parameters. Detection probability estimates were low and decreased after the first detection of the presence of the species in a site, which is counterintuitive.

Our model provides an estimate of the direction of spread direction consistent with empirical observations: globally, the spread direction is toward the northwest. The invasion started in the early 1980s through Florida. Colonization really started in direction of the north in the 1990s, before expanding toward the west in the late 1990s. If the spread had not followed a specific direction, and had been globally the same in every spatial direction, the estimate for standard error for δ_1 would have been large. Compared to Wikle (2003), the spatiotemporal model that we propose accounts for the possibility of a directional diffusion. Moreover, here the spatial structure is explicitly specified, which may be interesting, especially if structures are known (or expected) to facilitate or hamper population or individual movements. The explicit formulation of the spatial
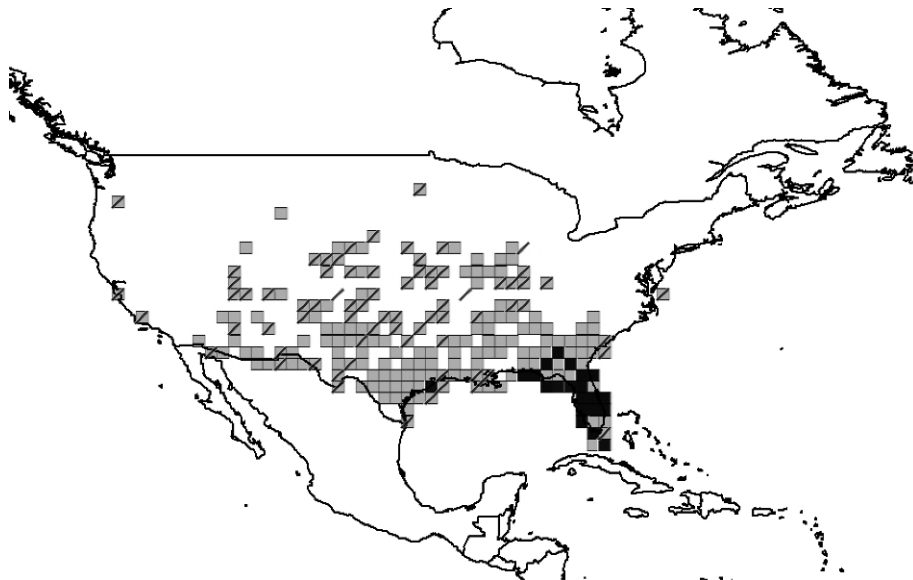
FIG. 3. Eurasian Collared-Dove (*Streptopelia decaocto*) site occupancy observations for 1996 (black squares) and 2006 (gray squares), and our corresponding estimations of probability of being colonized in 2006 (hatched squares; probability of being colonized > 0.5).

structure may be used to reflect a particular underlying geographical structure that may affect the spread of the invasion. In this study, we have considered the simplest case with no environmental or topographical barriers, but this could be supported by the weight matrix. If we had considered these elements, estimation of the direction of the spread, and impact of the distance might have been different. For example, we might expect the Great Plains to facilitate the invasion process, whereas the Rocky Mountains might stop or slow down this propagation. On the other hand, human activities such as farming or even bird feeding in cities might help the Collared-Dove to settle in areas in the north that would not be particularly suitable due to lower temperature.

A development of this model would be to estimate the spread direction for each year to have a more accurate and detailed description of the colonization process and of the local anisotropic conditions, but this would require tremendous computing time and calculus capacities. This year-specific approach may permit us to evaluate if propagation of the invasive is facilitated in the south once the species has reached the Mexican border, because of the suitable climatic conditions. Moody and Mack (1988) and Hajek et al. (1996) have discussed the importance of targeting control efforts at the leading edge of invading populations, especially at recently colonized sites ahead of the main body of the invasion. If these sites are colonized, this spread can speed the overall invasion progress. Accurate estimation of the spread direction and colonization probabilities can greatly improve how and where control efforts should be devoted. For this purpose, estimates provided by the type of model that we developed permit the

construction of predictive maps for the invasion progress of a particular species, which should help to target management efforts to locations where they would be more useful and pertinent.

The relationship between site persistence probability and density is quadratic, which may reflect an Allee effect at low densities. The high persistence probability at intermediate to high densities indicates that once the species has started to spread and has settled in a location, it tends to stay there, which leads to a durable invasion. The slight observed decrease at the highest densities might be due to extrapolation beyond the observed densities. The relatively high recolonization
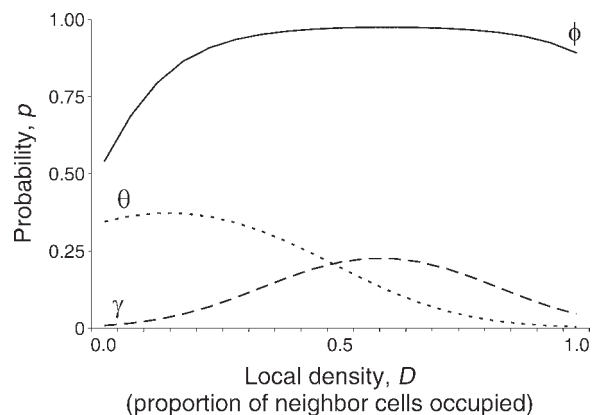

FIG. 4. Estimates of dynamics parameters as a function of local density, D. ($D = 1.0$ corresponds to all neighboring sites to be occupied.) The solid line is the site persistence probability, ϕ. The dashed line is the initial colonization probability, γ, and the dotted line is the recolonization probability, θ.

probability at low density indicates that even if a site is abandoned, it will be reoccupied. The decrease of recolonization at higher densities is probably due to the fact that there is no available site and they all stay occupied, considering the high persistence probability at these densities. The low initial colonization probability at low densities indicates that there is no initial colonization event in isolated areas. The higher initial colonization probability at intermediate density shows that colonization events occur at the leading edge of the invading population (i.e., the margin of geographical distribution). The distinction between initial colonization and recolonization may be useful to deal with a wide range of topics in which these two processes (or "subprocesses") are influenced by different factors. For example, in the framework of the selection of a breeding site, we may expect higher-quality sites to be chosen first (Møller 1982, Newton and Marquiss 1982), and when they get deserted for any reason by individuals (e.g., death, dispersal, ...), we may expect these abandoned sites to be reoccupied first, before new breeding sites are "created." Here, we have seen that initial colonization and recolonization are influenced differently by density. For management purposes, it may be useful to distinguish these two subprocesses because they might not be equally affected by control measures; the latter may even lead to opposite effects in each subprocess, and to an unwanted response.

We expected that weights attributed to site occupancy status decrease as the distance between sites increases. It is logical to think that sites close to already occupied sites have higher persistence and colonization probabilities than distant sites (especially in the case of an avian invasive species). For example, the proximity of a source of dispersing individuals has been shown to increase the likelihood of a given site being colonized (Kolar and Lodge 2001, Lockwood et al. 2005). However, our results do not corroborate this hypothesis: we have not detected any relationship between distance among sites at the scale considered and the dynamics parameters. The most likely explanation for this non-detection of a distance effect is the scale at which we are working (i.e., the size of the cell). At a smaller scale, this effect might have been detected. This may depend on the speed of the invasion: indeed, we might expect that the site influence area increases in size as the invasion speed increases. In other words, if a species colonizes an area rapidly, distant sites are as likely to be quickly colonized as close sites. At a small scale, variation in distance might not allow detection of an impact of the distance because neighboring sites might all have a high colonization probability if the species is mobile. On the other hand, if the scale were too large, all sites would have a low colonization probability, no matter their distance from an occupied cell, because even directly neighboring cells might be over the influence area of the occupied point within the cell. Therefore we would not detect an impact of distance on colonization probabilities if the scale we

are working at did not allow us to draw a distinction between significant classes of distance (for the colonization process). Especially in the case of avian invasive species, individuals can be expected to have high dispersive capacities. Therefore, the distance effect would be small and hard to detect. It would be interesting to try to rerun this model using a grid including a larger number of cells (i.e., reduce the size of cells) and a more important number of layers to incorporate more distance classes, but once again, such an advanced decomposition requires an important amount of data and tremendous computing capacities.

Although detection probability estimates seem very low, it should be noted that they correspond to the detection probability at the stop level; therefore, global detection probability for the cell is significantly higher. A simple calculation shows that when the species is present in a cell, the probability of detecting the species at least once is $1 - (1 - p)^K$. For example, even if we have $p = 0.02$, the probability of detecting at least once the species in a cell with only one route (i.e., $K = 50$, because there are 50 stops per route) will be $1 - 0.98^{50} = 0.64$. This probability is equal to 0.87 if there are two routes in the cell ($K = 100$), and increases up to 0.95 for three routes. Furthermore, the deterioration of detectability after a first detection is probably due to the type of data that we used and the scale considered. We merged BBS occupancy data from several routes when variability among routes can sometimes be substantial. Consequently, there can be a very large intrinsic variance in patterns of detections within a cell. The artificial homogenization resulting from the merging of several routes into a cell may diminish our ability to detect sensitive effects in the data. We suspect that the combination of the scale considered and the grouping of data from several routes within cells might have led to a poor estimation of detection probability. Analyses at a smaller scale with abundance data may provide more reliable detectability estimates and may allow assessment of the "learning effect" due to a previous detection, if any. Additionally, the BBS data are subject to multiple sources of uncertainty, especially within-site variability (Link et al. 1994, Sauer et al. 1994), which should be considered seriously if such data were used to fit abundance models.

The model developed in this paper is based on a simple, and yet powerful (in terms of versatility and quality of estimations), hierarchical formulation aimed at assessing the occupancy status of sites as functions of persistence and colonization probabilities. The basic model has been described and used in MacKenzie et al. (2003), who gave the basic likelihood formulation, and Royle and Kéry (2007), who provided the hierarchical Bayesian form. Conditional on the presence of the species or individuals in a site, data (i.e., observations) are then expressed as a function of detection probability. It is important to realize that this class of models can be fitted at very different ecological scales. In fact, as

described in Royle and Dorazio (2008), if you consider a "static" system, the ecological scale (e.g., population of individuals, metapopulations) is expressed by a size parameter (e.g., number of individuals or species). Dynamic systems include parameters such as survival and recruitment probabilities for populations, analogous to extinction and colonization parameters in metapopulations and community systems. Therefore, such models can be used to address a large variety of topics. The model shown in this paper has been developed to be fitted at scales as different as colonization of the United States by the Eurasian Collared-Dove, or nest selection inside a cliff by the Black-legged Kittiwake *Rissa tridactyla*.

Our model presents some components specific to invasive species that can easily be modified. For example, the quadratic expression of dynamics parameters as a function of density allows us to account for peaks, and therefore is useful to consider a priori biological assumptions. However, this is an arbitrary formulation. It is possible to develop an individual-based model to account for a more accurate relationship between dynamics parameters and density (in terms of biological and/or behavioral mechanisms). Moreover, in the current state of the model, ϕ_1 and ϕ_2 do not depend on time. We made this decision because we assumed that the way the neighborhood influences a site's fate does not depend on the growth status of the population. This assumption can be easily relaxed. Moreover, we modeled $\gamma_{i,t}$ and $\theta_{i,t}$ as $\phi_{i,t}$, but this can be changed depending on whether investigators want to have all of the dynamics parameters to be influenced by the spatial structure of occupancy, or not. The spatial spread of an invasive species follows successful establishment, which is a step driven by colonization and extinction processes (Eraud et al. 2007). Several factors associated with variation in colonization probabilities were highlighted by recent theoretical work on invasion and metapopulation theory. Environmental heterogeneity and distance between patches of suitable habitats (Hastings et al. 2005), for example, have an influence on persistence and colonization probabilities (With 2002) and can be modeled differently for each dynamics parameter.

Modeling of invasive species population dynamics is often based on population density and growth rate estimation (Veit and Lewis 1996, Taylor and Hastings 2004, Le Maitre et al. 2008), and efficient ways of estimating occupancy are emerging via the development of hierarchical models. However, these models are not yet extensively used to deal with invasive species. Several mathematical formulations have been developed to define, estimate, and ultimately model the spatial expansion of spreading population (e.g., Van den Bosch et al. 1992, Neubert and Caswell 2000, Ortega-Cejas et al. 2004). But these classical models, despite the interesting advances they represent, suffer some limitations in terms of assessment and understanding of the

expansion process (Eraud et al. 2007); e.g., lack of estimates of relevant components such as occupancy rate and local colonization probability. Furthermore, existing models are usually fitted to distribution maps derived from observational counts of organisms, assuming perfect detectability of individuals or species (Eraud et al. 2007). Models such as the one developed in this paper, through their high modularity and consideration of detectability issues, can expand and complete observations given by these classical models.

Invasive species have become a major problem, not only in North America but also throughout the world (Cronk and Fuller 1995, Williamson 1996, 1999, Mooney and Cleland 2001, Molnar et al. 2008). Investigations of the spatiotemporal dynamics of invasive species will permit predictions about future spread as well as about the likely efficiency of management actions designed to control such spread (e.g., Wikle 2003). Ultimately, management and control of invasive species must be conducted in an integrative framework where ecological, statistical, and dynamical approaches are to be coupled with genetic studies. As a matter of fact, because of possible hybridization with sympatric species, it is necessary to couple occupancy observations with potential genetic consequences of this invasion on other close species to fully realize the impact of invasive species on local ecosystems.

The main aim of this work was to develop a hierarchical model that encompasses time and space dimensions in a convenient and flexible way. Moreover, the distinct relationships of initial and recolonization probabilities with density confirm that it is essential to consider these two subprocesses to understand the global colonization process, to fully understand how an invasion progresses, and ultimately to be able to make suitable management decisions.

Literature Cited

Andow, D. A., P. M. Kareiva, S. A. Levin, and A. Okubo. 1990. Spread of invading organisms. Landscape Ecology 4:177–188.

Banerjee, S. 2003. Hierarchical modeling and analysis for spatial data. Chapman and Hall/CRC, Boca Raton, Florida, USA.

Bennett, R. E., J. D. Nichols, J.-D. Lebreton, R. Pradel, J. E. Hines, and W. M. Kitchen. 2001. Methods for estimating dispersal probabilities and related parameters using marked animals. Pages 3–17 in J. Clobert, E. Danchin, A. A. Dhondt, and J. D. Nichols, editors. Dispersal. Oxford University Press, Oxford, UK.

Casella, G., and E. I. George. 1992. Explaining the Gibbs sampler. American Statistician 46:167–174.

Clark, J. S. 2007. Models for ecological data: an introduction. Princeton University Press, Princeton, New Jersey, USA.

Courchamp, F., T. Clutton-Brock, and B. Grenfell. 1999. Inverse density dependence and the Allee effect. Trends in Ecology and Evolution 14:405–410.

Cronk, Q. C. B., and J. L. Fuller. 1995. Plant invaders. The threat to natural ecosystems. Chapman and Hall, London, UK.

Dunn, J. L., and J. Alderfer. 2006. National Geographic field guide to the birds of North America. Fifth edition. National Geographic Society, Washington, D.C., USA.

Eraud, C., J.-M. Boutin, D. Roux, and B. Faivre. 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.

Ficetola, F. G., C. Coïc, M. Detaint, M. Berroneau, O. Lorvelec, and C. Miaud. 2007. Pattern of distribution of the American bullfrog Rana catesbeiana in Europe. Biological Invasions 9:767–772.

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

Glowka, L., F. Burhenne-Guilmin, and H. Synge. 1994. A guide to the convention on biological diversity. IUCN (International Union for the Conservation of Nature), Gland, Switzerland.

Gurevitch, J., and D. K. Padilla. 2004. Are invasive species a major cause of extinctions? Trends in Ecology and Evolution 19:470–474.

Hajek, A. E., J. S. Elkinton, and J. J. Witcosky. 1996. Introduction and spread of the fungal pathogen Entomophaga maimaga (Zygomycetes: Entomophthorales) along the leading edge of gypsy moth (Lepidoptera: Lymantriidae) spread. Environmental Entomology 25:1235–1247.

Harrod, R. J., and S. Reichard. 2001. Fire and invasive species within the temperate and boreal coniferous forests of western North America. Pages 95–101 in K. E. M. Galley and T. P. Wilson, editors. Proceedings of the Invasive Species Workshop: the role of fire in the control and spread of invasive species. Miscellaneous publication 11. Tall Timbers Research Station, Tallahassee, Florida, USA.

Hastings, A., et al. 2005. The spatial spread of invasions: new developments in theory and evidence. Ecology Letters 8:91–101.

Hengeveld, R. 1993. What to do about the North American invasion by the collared dove? Journal of Field Ornithology 64:477–489.

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.

Ibarra, A. A., Y. S. Park, S. Brosse, Y. Reyjol, P. Lim, and S. Lek. 2005. Nested patterns of spatial diversity revealed for fish assemblages in west European river. Ecology of Freshwater Fish 14:233–242.

Joseph, L. N., S. A. Field, C. Wilcox, and H. P. Possingham. 2006. Presence–absence versus abundance data for monitoring threatened species. Conservation Biology 20:1679–1687.

Keeley, J. E. 2006. Fire management impacts on invasive plants in the western United States. Conservation Biology 20:375–384.

Keyser, P. D., D. C. Guynn, Jr., and H. S. Hill, Jr. 2005. Density-dependent recruitment patterns in white-tailed deer. Wildlife Society Bulletin 33:222–232.

Kolar, C. S., and D. M. Lodge. 2001. Progress in invasion biology: predicting invaders. Trends in Ecology and Evolution 16:199–204.

Kot, M., M. A. Lewis, and P. Van Den Driessche. 1996. Dispersal data and the spread of invading organisms. Ecology 77:2027–2042.

Lawson, A. B. 2006. Statistical methods in spatial epidemiology. Second edition. Wiley Series in Probability and Statistics. John Wiley, New York, New York, USA.

Lawson, A. B. 2008. Bayesian disease mapping: hierarchical modeling in spatial epidemiology. Chapman and Hall/CRC, Boca Raton, Florida, USA.

Le Maitre, D. C., R. M. Krug, J. H. Hoffmann, A. J. Gordon, and T. N. Mgidi. 2008. Hakea sericea: development of a model of the impacts of biological control on population dynamics and rates of spread of an invasive species. Ecological Modelling 212:342–358.

Leprieur, F., O. Beauchard, S. Blanchet, T. Oberdorff, and S. Brosse. 2008. Fish invasions in the world's river systems: when natural processes are blurred by human activities. PLoS 6:e28.

Link, W. A., R. J. Barker, and J. R. Sauer. 1994. Within-site variability in surveys of wildlife populations. Ecology 75:1097–1108.

Lockwood, J. L., P. Cassey, and T. Blackburn. 2005. The role of propagule pressure in explaining species invasions. Trends in Ecology and Evolution 20:223–228.

Lunn, D. J., A. Thomas, N. Best, and D. Spiegelhalter. 2000. WinBUGS: a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and Computing 10:325–337.

Mack, R. N. 1985. Invading plants: their potential contribution to population biology. I. Pages 127–142 in J. White, editor. Studies in plant demography. Academic Press, London, UK.

MacKenzie, D. I., J. D. Nichols, J. E. Hines, M. G. Knutson, and A. D. Franklin. 2003. Estimating site occupancy colonization and local extinction when a species is detected imperfectly. Ecology 84:2200–2207.

MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. A. Royle, and A. C. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83:2248–2255.

MacKenzie, D. I., J. D. Nichols, J. A. Royle, K. H. Pollock, L. L. Bailey, and J. E. Hines. 2006. Occupancy estimation and modeling. Academic Press, San Diego, California, USA.

Møller, A. P. 1982. Characteristics of Magpie Pica pica territories of varying duration. Ornis Scandinavica 13:94–100.

Molnar, J. L., R. L. Gamboa, C. Revenga, and M. D. Spalding. 2008. Assessing the global threat of invasive species to marine biodiversity. Frontiers in Ecology and the Environment 6:485–492.

Moody, M. E., and R. N. Mack. 1988. Controlling the spread of plant invasions: the importance of nascent foci. Journal of Applied Ecology 25:1009–1021.

Mooney, H. A., and E. E. Cleland. 2001. The evolutionary impact of invasive species. Proceedings of the National Academy of Sciences USA 98:5446–5451.

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.

Neubert, M. G., and H. Caswell. 2000. Demography and dispersal: calculation and sensitivity analysis of invasion speed for structured populations. Ecology 81:1613–1628.

Newton, M., and I. Marquiss. 1982. Habitat preference in male and female sparrowhawks Accipiter nisus. Ibis 124:324–328.

Ntzoufras, I. 2009. Bayesian modeling using WinBUGS. John Wiley, New York, New York, USA.

Okubo, A. 1980. Diffusion and ecological problems: mathematical models. Springer-Verlag, Berlin, Germany.

Ortega-Cejas, V., J. Fort, and V. Méndez. 2004. The role of the delay time in the modeling of biological range expansions. Ecology 85:258–264.

Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. 1978. Statistical inference from capture data on closed animal populations. Wildlife Monographs 62:1–135.

R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ⟨http://www.R-project.org⟩

Robbins, C. S., D. Bystrak, and P. H. Geissler. 1986. The Breeding Bird Survey: its first fifteen years, 1965–1979. U.S. Department of the Interior, U.S. Fish and Wildlife Service, Washington, D.C., USA.

Romagosa, C. M. 2002. Eurasian Collared-Dove (*Streptopelia decaocto*). Number 630 *in* A. Poole and F. Gill, editors. The birds of North America. Birds of North America, Philadelphia, Pennsylvania, USA.

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.

Romagosa, C. M., and T. McEneaney. 2000. The Eurasian Collared-Dove in North America and the Caribbean. North American Birds 53:348–353.

Royle, J. A., and R. M. Dorazio. 2008. Hierarchical modeling and inference in ecology. The analysis of data from populations, metapopulations, and communities. Academic Press, New York, New York, USA.

Royle, J. A., and M. Kéry. 2007. A Bayesian state-space formulation of dynamic occupancy models. Ecology 88:1813–1823.

Royle, J. A., and J. D. Nichols. 2003. Estimating abundance from repeated presence–absence data or point counts. Ecology 84:777–790.

Sauer, J. R., B. G. Peterjohn, and W. A. Link. 1994. Observer differences in the North American Breeding Bird Survey. Auk 111:50–62.

Sing, T., O. Sander, N. Beerenwinkel, and T. Lengauer. 2009. ROCR: visualizing the performance of scoring classifiers. R package version 1.0-4. ⟨http://CRAN.R-project.org/package=ROCR⟩

Skellam, J. G. 1951. Random dispersal in theoretical populations. Biometrika 38:196–218.

Smith, P. W. 1987. The Eurasian Collared-Dove arrives in the Americas. American Birds 41:1370–1379.

Sturtz, S., U. Ligges, and A. Gelman. 2005. R2WinBUGS: a package for running WinBUGS from R. Journal of Statistical Software 12:1–16.

Taylor, C. M., and A. Hastings. 2004. Finding optimal control strategies for invasive species: a density-structured model for *Spartina alterniflora*. Journal of Applied Ecology 41:1049–1057.

Van den Bosch, F., R. Hengeveld, and J. A. J. Metz. 1992. Analysing the velocity of animal range expansion. Journal of Biogeography 19:135–150.

Veit, R. R., and M. A. Lewis. 1996. Dispersal, population growth, and the Allee effect: dynamics of the house finch invasion of eastern North America. American Naturalist 148:255–274.

Waller, L. A., and C. A. Gotway. 2004. Applied spatial statistics for public health data. Wiley Series in Probability and Statistics. John Wiley, New York, New York, USA.

Wikle, C. K. 2003. Hierarchical Bayesian models for predicting the spread of ecological processes. Ecology 84:1382–1394.

Wilcove, D. S., D. Rothstein, J. Dubow, A. Phillips, and E. Losos. 1998. Quantifying threats to imperiled species in the United States. BioScience 48:607–615.

Williams, B. K., J. D. Nichols, and M. J. Conroy. 2002. Analysis and management of animal populations. Academic Press, San Diego, California, USA.

Williamson, M. 1996. Biological invasions. Chapman and Hall, London, UK.

Williamson, M. 1999. Invasions. Ecography 22:5–12.

With, K. A. 2002. The landscape ecology of invasive spread. Conservation Biology 16:1192–1203.

Zhu, H., H.-C. Huang, and C.-T. Wu. 2005. Modeling spatial-temporal binary data using Markov random fields. Journal of Agricultural, Biological, and Environmental Statistics 10:212–225.

SUPPLEMENT

WinBUGS code for the model of invasion process (*Ecological Archives* A021-014-S1).