

## ENVIRONMENTAL HETEROGENEITY, BIRD-MEDIATED DIRECTED DISPERSAL, AND OAK WOODLAND DYNAMICS IN MEDITERRANEAN SPAIN

DREW W. PURVES,<sup>1,3</sup> MIGUEL A. ZAVALA,<sup>2</sup> KIONA OGLE,<sup>1</sup> FERNANDO PRIETO,<sup>2</sup> AND JOSÉ M. REY BENAYAS<sup>2</sup>

<sup>1</sup>*Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey 08544 USA*

<sup>2</sup>*Departamento de Ecología, Edificio de Ciencias, Universidad de Alcalá, E-28871, Alcalá (Madrid), Spain*

**Abstract.** Vegetation dynamics in complex landscapes depend on interactions among environmental heterogeneity, disturbance, habitat fragmentation, and seed dispersal processes. We explore how these features combine to affect the regional abundances and distributions of three *Quercus* (oak) species in central Spain: *Q. faginea* (deciduous tree), *Q. ilex* (evergreen tree), and *Q. coccifera* (evergreen shrub). We develop and parameterize a stochastic patch occupancy model (SPOM) that, unlike previous SPOMs, includes environmentally driven variation in disturbance and establishment. Dispersal in the model is directed toward local (nearby) suitable habitat patches, following the observed seed-caching behavior of the European Jay. Model parameters were estimated using Bayesian methods and survey data from 12 047 plots. Model simulations were conducted to explore the importance of different dispersal modes (local directed, global directed, local random, global random). The SPOM with local directed dispersal gave a much better fit to the data and reproduced observed regional abundance, abundance–environment correlations, and spatial autocorrelation in abundance for all three species. Model simulations suggest that jay-mediated directed dispersal increases regional abundance and alters species–environment correlations. Local dispersal is estimated to reduce regional abundances, amplify species–environment correlations, and amplify spatial autocorrelation.

Parameter estimates and model simulations reveal important species-specific differences in sensitivity to environmental perturbations and dispersal mode. The dominant species *Q. ilex* is estimated to be highly fecund, but on the edge of its climatic tolerance. Therefore *Q. ilex* gains little from directed dispersal, suffers little from local dispersal, and is relatively insensitive to changes in habitat cover or disturbance rate; but *Q. ilex* is highly sensitive to altered drought length. In contrast, the rarest species, *Q. coccifera*, is well adapted to the climate and soils but has low fecundity; thus, it is highly sensitive to changes in dispersal, habitat cover, and disturbance but insensitive to altered drought length. Finally, *Q. faginea* is estimated to be both at the edge of its climatic tolerance and to have low fecundity, making it sensitive to all perturbations. Apparently, co-occurring species can exhibit very different interactions among dispersal, environmental characteristics, and physiological tolerances, calling for increased attention to species-specific dynamics in determining regional vegetation responses to anthropogenic perturbations.

**Key words:** *biogeography; climate change; climate envelope; corvids; dispersal limitation; Garrulus; holm oak; Iberian peninsular; incidence function; patch model; species migration; zoochory.*

### INTRODUCTION

It has long been recognized that the structure, composition, and distribution of plant communities is highly correlated with the physical environment, at scales from global (Walter 1973, 2002, Archibold 1995) to regional (Whittaker 1956, 1960) to local (Cowles 1899). Yet, the development of a quantitative understanding of these relationships has remained a significant challenge, because they arise from complex and often nonlinear interactions between species-specific ecophysiological traits, competitive and facilitative interactions within

and among species, and large-scale population processes including dispersal limitation. Moreover, all of these processes operate within heterogeneous landscapes that are subject to natural disturbance and anthropogenic habitat fragmentation (Crawley 1997). And yet a quantitative understanding of these processes is needed for predicting vegetation responses to climate change and other natural and anthropogenic perturbations (Pacala and Hurtt 1993, Lawton 2000).

A common approach to studying environmental forcing in plant communities is regression analysis of the occurrence or abundance of focal species vs. abiotic variables (e.g., Austin et al. 1990, Franklin 1998, Leathwick and Whitehead 2001, Heegaard 2002). We refer to this as “gradient analysis”; for climate-specific

Manuscript received 6 December 2005; accepted 27 March 2006; final version received 10 July 2006. Corresponding Editor: S. W. Seagle.

<sup>3</sup> E-mail: [dpurves@princeton.edu](mailto:dpurves@princeton.edu)

applications, it has also been referred to as “bioclimate envelope” or “climate envelope” modeling (Davis et al. 1998a, b, Pearson and Dawson 2003). Gradient analysis has successfully reproduced observed distributions of species (e.g., Franklin 1998), and it has been used for predicting how distributions might be altered by climate change (e.g., Iversen and Prasad 1998). However, gradient analysis is not well suited for inferring underlying ecological processes, or for predicting future patterns, because it lacks population dynamic processes (Pacala and Hurtt 1993, Davis et al. 1998a, b). Hence, gradient analysis inherently assumes that current species–environment relationships will hold in the future, and it cannot predict the time course of responses to perturbations. Moreover, while changes in physical and climatic conditions may be critical in determining the distributions and abundance of some plant species in the future, other factors are likely to be at least as important, including habitat loss, fragmentation, and changes in disturbance rates (Primack 1998, Davies et al. 2001).

These issues call for quantitative methods that explicitly link population dynamics with environmental forcing (Pacala and Hurtt 1993, Lawton 2000). One direction is the development of individual-based models rich in mechanistic functions describing growth, mortality, competition, and dispersal processes (e.g., Desanker 1996, Pacala et al. 1996). Although attractive because of their attention to biological detail, such models are difficult to implement because they are computationally demanding, require large quantities of data collected at a variety of scales, and are too complex to be parameterized top-down from survey data on the abundance and distribution of species. There are promising mathematical techniques for estimating the aggregate behavior of individual-based models, which reduces computing time and may enable top-down parameterization (Bolker et al. 2000, Iwasa 2000, Law and Dieckmann 2000, Moorcroft et al. 2001, Law et al. 2003), but these techniques remain challenging to implement.

Another potential route, which lies between the extremes of gradient analysis and individual-based models, is metapopulation modeling (Hanski and Gilpin 1997, Hanski and Gaggiotti 2004). The classic Levins metapopulation model (Levins 1969) reduces the regional population of a species to a collection of patches, each of which is occupied or empty. Patch occupancy, and thus the frequency and distribution of species, arises from a balance between local extinctions and colonizations (see Hanski 1997). Models of this type are known as stochastic patch occupancy models (SPOM; Hanski and Simberloff 1997, Etienne et al. 2004). Importantly, SPOMs exclude all within-patch dynamics, and so are quick to implement and include a much smaller set of parameters than individual-based models. Hence, SPOMs can be parameterized from survey data and implemented at large spatial scales (Etienne et al. 2004). In addition, the SPOM framework is especially well suited for addressing the consequences of habitat loss,

fragmentation, and changes in disturbance rates (e.g., Hanski and Simberloff 1997). However, the use of metapopulation modeling is obviously only suited to populations that appear to be subject to a metapopulation structure (see Freckleton and Watkinson 2002).

However, despite their simplicity and applicability to regional-scale population dynamics, the simple SPOMs that form the core of applied metapopulation modeling in animal populations have apparently not been implemented for studying vegetation dynamics (reviewed by Etienne et al. 2004). Also, although they have great potential for assessing species–environment relationships and for predicting plant population responses to changes in abiotic factors, we are not aware of any studies that directly incorporate continuous variation in environmental forcing (drought, temperature, altitude, and so on) into the SPOM framework. Rather, in common with metapopulation ecology in general, work on SPOMs has tended to assume a homogenous environment, with fundamental rates of local extinction, colonization and dispersal being constant across the patch network (though see theoretical treatments by Hiebler 2000 and Lopez and Pfister 2001). And yet, site-specific variation in vital rates can be incorporated into SPOMs readily, as demonstrated by studies that have included extinction and colonization rates that depend on patch size (e.g., Hanski 1997, Vos et al. 2000).

Here, we extend the theoretical and empirical framework that has been established for SPOMs (see Etienne et al. 2004) by making the key population processes—establishment and disturbance—functions of local physical and climatic conditions. This approach gives a parsimonious modeling framework that is capable of capturing interactions between environmental forcing and population dynamics; provides a means for identifying key biotic–abiotic linkages; and yields a framework for predicting potential effects of environmental change on vegetation dynamics, including, for example, climate change, habitat loss, and fragmentation. We show how the SPOM can be parameterized from survey data, and then used to address basic ecological questions, and to provide estimates for the sensitivities of vegetation to changes in climate, disturbance regimes, and land use.

We apply the SPOM to three co-occurring oak species (*Quercus faginea* Lam., *Q. ilex* L. ssp. *ballota*, and *Q. coccifera* L.) that are a major component of woodlands in central Spain and other Mediterranean ecosystems (Grove and Rackham 2001) and for which excellent survey data are available (Inventario Forestal Nacional 1995). The distribution and abundance of these oak species are likely to be governed by complex metapopulation–environment interactions because (1) the dominant plant community types are highly fragmented (de Miguel 1999, Arianoutsou and Papanastasis 2004; and see Fig. 1b, c); (2) stand-leveling disturbances, principally fire, are common (Vázquez et al. 2002); (3) the

distribution of these species is thought to depend on the ability of seedlings to establish under different environmental conditions (Rey Benayas 1998, Retana et al. 1999); and (4) the region is characterized by pronounced small-scale heterogeneity in physiography, soils, and climate (e.g., Gavilán and Fernández-González 1997).

An important process affecting these woodlands is large-scale dispersal, which is primarily due to acorn movement and caching by the European Jay (*Garrulus glandarius*), whose range includes the entire Iberian Peninsula (see Plate 1). Jays move seeds over sufficient distances to colonize empty patches, although dispersal is nevertheless local at the multi-kilometer scales examined here (Gomez 2003). Also, the jays preferentially cache acorns in habitats suitable for seedling establishment (Gomez 2003). Directed dispersal of this kind is expected to have fundamental effects on the response of metapopulations to changes in habitat cover and disturbance (e.g., Etienne 2000, Purves and Dushoff 2005).

We use simulations of the SPOM to explore the sensitivity of the distribution and frequency of the oak species to changes in some key biotic and abiotic factors. First, we evaluate the contribution of two different aspects of jay-mediated seed dispersal (spatially local, and directed toward suitable habitat) to the current frequencies and distributions. Second, climate models predict that Spanish summers will become drier over the next century (Intergovernmental Panel on Climate Change 2001: Figs. 3–5), so we investigate sensitivity to changes in drought length. Third, fire frequency could change in the future over and above any changes in climate, although the direction of change is uncertain due to the influence of human-induced ignitions, fire management, and land-use policy (Vázquez et al. 2002). Thus, we implement simulations under different disturbance frequencies. Finally, this region is composed of natural and seminatural landscapes interspersed with agricultural and urban areas, so the impacts of changes in land use on the region's vegetation dynamics are of particular interest (Blondel and Aronson 1995, Gomez-Limón and Fernandez 1999). Therefore, we conduct simulations under different levels of habitat cover and fragmentation.

#### STUDY REGION AND SPECIES

The study area consists of the provinces of Madrid and Castile La-Mancha, located in the center of the Iberian Peninsula (Fig. 1). The region spans 38.0° N to 41.3° N and 0.8° W to 5.5° W, encompassing a large altitudinal gradient (300–2000 m). The climate in this area is quite variable, with mean annual precipitation ranging from 300 to 1900 mm, warm summers (average July temperature: 24° to 36°C) and fairly cold winters (average January temperature: –5° to 2°C). The landscape is a mosaic of seminatural forests, savannas, shrublands, grasslands, intensive agricultural fields, and urban areas. The Second Spanish Forest Inventory (Inventario Forestal Nacional 1995) sampled this region

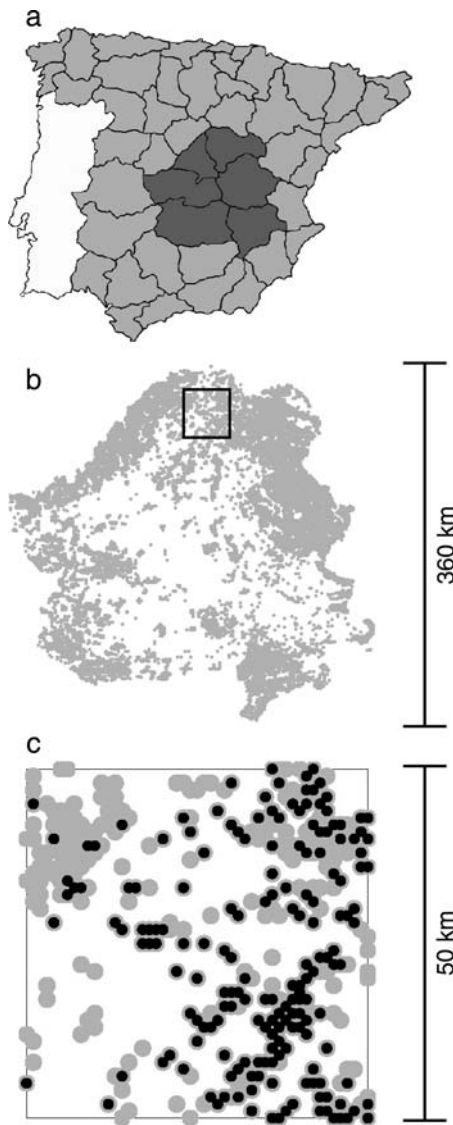


FIG. 1. (a) Map of the Iberian Peninsula, showing Portugal (unshaded) and Spain (light gray), including the study region (Madrid and Castile La-Mancha, dark gray). (b) Distribution of the 12 047 forest inventory plots within the study region. The plot density follows the density of forest, woodland, and shrubland. (c) Distribution of inventory plots (gray), and inventory plots in which *Quercus faginea* was recorded as present (black), within the 50 × 50 km region outlined in (b).

from a network of survey plots superimposed onto wooded areas with a density of approximately one plot per 4 km<sup>2</sup> (Villanueva 1993), giving a total of 12 047 plots. Each plot was censused for several attributes, including the presence or absence of several species of trees and shrubs, stem diameter data for some tree species, slope, aspect, and soil conditions. For the purpose of this study, we extracted the presence–absence data for three *Quercus* species (where presence or absence means presence or absence of a living tree or shrub of the species in a 25 m radius survey plot).

TABLE 1. Explanation of notation used in the stochastic patch occupancy model (SPOM) and in parameter estimation.

Symbol	Description
$j, i, q, t$	Indices for species ( $j$ ), model site ( $i$ ), survey plot ( $q$ ), time ( $t$ )
$Z_{j,i,q}$	Presence (1) or absence (0) of species $j$ , in model site $i$ ( $Z_{j,i}$ ) or observed in survey plot $q$ ( $Z_{j,q}^{\text{obs}}$ )
$\mathbf{Z}_j, \mathbf{Z}_j^{\text{obs}}$	Vectors of all $Z_{j,i}$ and $Z_{j,q}^{\text{obs}}$ for all sites and/or survey plots
$q(i), i(q)$	Nearest $q$ to $i$ , and nearest $i$ to $q$ , respectively
<b>SPOM</b>	
$\phi_i, \alpha_{j,i}$	Annual probability of extinction (for site $i$ ); and probability of seed establishment (for a seed of species $j$ at site $i$ )
$S_{j,i}(t), s(j, i, i')$	Seed rain ( $\text{yr}^{-1}$ ): total into site $i$ , and from site $i'$ to $i$ , respectively
$\omega_j$	Total seed output ( $\text{yr}^{-1}$ ) from a patch occupied by species $j$
$g(i, i')$	Dispersal kernel for movement of seeds from $i'$ to $i$
$a, b$	Parameters for local dispersal kernel (Eq. 4)
$\mu_{1,j} \cdots \mu_{3,j}, \sigma_{1,j} \cdots \sigma_{3,j}, \kappa_{3,j}$	Parameters defining the relationship between $\alpha_{j,i}$ and physical variables at site $i$ , $V_{1,j}(i) \cdots V_{3,j}(i)$
<b>Parameter estimation</b>	
$l\{\mathbf{Z}_j^{\text{obs}}   \theta\}$	Log-likelihood of data $\mathbf{Z}_j^{\text{obs}}$ given parameter set $\theta$
$P_{j,q,0}, P_{j,i,0}$	Probability that species $j$ will be found in plot $q$ or site $i$ , respectively
$f\{\theta   \mathbf{Z}_j^{\text{obs}}\}$	Posterior probability of parameter set $\theta$ given $\mathbf{Z}_j^{\text{obs}}$
$f\{\theta\}$	Prior probability of parameter set $\theta$

Inventario Forestal Nacional data from repeated surveys were not available at time of publication.

We focus on three oak species common in this region: *Quercus faginea* Lam. (Portuguese oak), *Q. ilex* L. spp. *ballota* (holm-oak), and *Q. coccifera* L (prickly-oak). Although these oaks are codominant in this region and co-occur in many areas, they exhibit quite different life history and ecophysiological strategies (Villar-Salvador et al. 1997, Fotelli et al. 2000, Grove and Rackham 2001, Corcuera et al. 2002, Valladares et al. 2002, Rey Benayas et al. 2005). *Quercus faginea* is a moderate-sprouting deciduous tree that is common in mesic calcareous locations. Conversely, *Q. coccifera* and *Q. ilex* are sclerophyllous evergreen Mediterranean specialists that are well-adapted to dry habitats, and both can resprout vigorously after cutting or browsing (Grove and Rackham 2001). Genets of all three species tend to be killed by fire (López-Soria and Castell 1992). In this region, *Q. ilex* grows as a shrub or tree, but *Q. coccifera* is found only as a shrub.

#### MODEL DESCRIPTION

The SPOM that we developed consists of a grid of sites  $i = 1, 2, \dots, N$  covering the study region. To match the spatial resolution of the inventory data, all model sites were set to  $2 \times 2$  km. Each site  $i$  is classified as suitable or unsuitable, where suitable means that the species can occur at the site and unsuitable means that they cannot (e.g., cropland, urban). The Inventario Forestal Nacional survey plots were only placed in locations with cover of woody plants, so we used the spatial pattern of the survey plots to determine the suitability of each site  $i$ . That is, all model sites  $i$  occurring within 2.0 km of a survey plot  $q$  were classified as suitable, yielding a total of 11 039 suitable sites and 11 901 unsuitable sites. The grid of model sites and the

network of survey plots exhibit a nearly one-to-one correspondence, but the overlap is not perfect because some survey plots are irregularly spaced.

There are several substantial areas of contiguous suitable habitat, and it is also common to find a group of two or more suitable sites nested within unsuitable habitat (see Fig. 1b, c). Within each of these contiguous areas of suitable habitat, the division into discrete patches is arbitrary. Therefore, the selection of patches in our model differs from previous SPOMs, which have been applied to habitats that are fundamentally discrete, such as ponds (Vos et al. 2000) or small clumps of host plants for insects (Hanski 1994). The SPOM that we developed for the Spanish oak system is described in detail below, and Table 1 summarizes the notation used in the model.

#### Disturbance and colonization

At any time  $t$ , each site  $i$  is either occupied by species  $j$  (i.e.,  $Z_{j,i}(t) = 1$ ) or unoccupied by species  $j$  (i.e.,  $Z_{j,i}(t) = 0$ ). Unsuitable sites are never occupied. By definition, the SPOM does not include any representation of within-site population dynamics. Therefore the state of the metapopulation at time  $t$  is given by the vector  $\mathbf{Z}_j(t)$ . The metapopulation dynamics are driven by two events: local extinction and colonization of suitable sites. At time  $t$ , and for each model site  $i$ , we calculate a probability that it will be disturbed during this iteration, denoted by  $\phi_i$  ( $\text{yr}^{-1}$ ). When site  $i$  is disturbed, all species present are set to absent ( $Z_{j,i}(t) = 0$  for all  $j$ ). Thus, we do not consider differential survival, or resprouting ability, but the available evidence suggest that fires tend to kill genets of all three species (López-Soria and Castell 1992). The value of  $\phi_i$  varies from site to site because of spatial variation in fire frequency in this region (see Appendix C). We also calculate a probability that

species  $j$  will colonize site  $i$  during this iteration, equal to  $1 - (1 - \alpha_{j,i})^{S_{j,i}(t)}$ .

These rules define a set of transition probabilities for each site  $i$ , which together constitute a complete description of the model:

$$P[z_{j,i}(t+1)|z_{j,i}(t)] = \begin{cases} \phi_i & \text{if } z_{j,i}(t) = 1 \text{ and } z_{j,i}(t+1) = 0 \\ 1 - \phi_i & \text{if } z_{j,i}(t) = 1 \text{ and } z_{j,i}(t+1) = 1 \\ 1 - (1 - \alpha_{j,i})^{S_{j,i}(t)} & \text{if } z_{j,i}(t) = 0 \text{ and } z_{j,i}(t+1) = 1 \\ (1 - \alpha_{j,i})^{S_{j,i}(t)} & \text{if } z_{j,i}(t) = 0 \text{ and } z_{j,i}(t+1) = 0 \end{cases} \quad (1)$$

where  $S_{j,i}(t)$  is the seed rain, i.e., the number of seeds of species  $j$  arriving at site  $i$  ( $\text{yr}^{-1}$ ); and  $\alpha_{j,i}$  is the probability that a seed of species  $j$  arriving at site  $i$  will become established as an adult (Eq. 1 refers to suitable sites  $i$  only; unsuitable sites are set to zero for all  $t$ ). Eq. 1 is derived by assuming that the probability that a new population takes hold at site  $i$  is equal to one minus the probability that none of the arriving seeds become established. Note that establishment by species  $j$  is not affected by the presence or absence of species other than  $j$ , and multiple species can coexist within a site.

#### Seed dispersal

The seed rain of species  $j$  at site  $i$  is calculated as the sum of the seed arriving from each occupied site  $i'$ ,  $s(j, i, i')$  where  $i' \neq i$ :

$$S_{j,i}(t) = \sum_{i' \in \{Z_{j,i'}(t)=1\}} s(j, i, i'). \quad (2)$$

The set  $\{Z_{j,i'}(t)=1\}$  contains all suitable sites  $i$  occupied by species  $j$  at time  $t$ , and

$$s(j, i, i') = \omega_j \cdot g(i, i') \quad (3)$$

where  $\omega_j$  is the seed output of an extant population of species  $j$  (number of seeds produced per year per patch). The fraction of seeds from site  $i'$  that land in site  $i$  is described by the dispersal kernel  $g(i, i')$ , which we assume is the same for all three species.

*Local vs. global dispersal.*—We explore two dispersal kernels that differ with respect to the distances over which seeds move. The first gives local dispersal, where seeds are more likely to arrive at nearby sites, and follows the generalized exponential function

$$g(i, i') = \frac{\exp[-a \cdot (d_{i,i'})^b]}{\sum_{i \in H} \exp[-a \cdot (d_{i,i'})^b]} \quad (4)$$

where  $d_{i,i'}$  is the Euclidean distance between the centers of sites  $i$  and  $i'$  (km);  $H$  is the set of all sites that can potentially receive seeds from site  $i'$ ; and parameters  $a$  and  $b$  set the rate at which seed dispersal declines with distance and the curvature of this decline, respectively.

An alternative formulation gives global dispersal, where seeds are distributed to near and far sites with

equal probability:

$$g(i, i') = \frac{1}{|H|} \quad (5)$$

where  $|H|$  is the size of set  $H$  (i.e., number of sites receiving seeds). There are a finite number of model sites, thus the two dispersal kernels are probability mass functions and  $\sum_i g(i, i') = 1$ .

*Directed vs. random dispersal.*—The denominators in Eqs. 4 and 5 contain the set  $H$ , and the definition of  $H$  determines whether dispersal is random or directed. To reflect the seed caching behavior of the European Jay, we assume that seeds are moved toward suitable habitat, and we refer to this as directed dispersal. “Directed” here implies that the seeds are moved nonrandomly with respect to patch type, which in this case refers to suitable (woodland) or not. In this case, the fraction of seeds that site  $i$  receives from site  $i'$  necessarily increases as the cover of suitable habitat decreases, because there are fewer patches for the jays to visit. The alternative, null model is random dispersal; in this case, seeds are distributed to both suitable and unsuitable habitat with equal probability, thus the fraction of seeds that arrive at site  $i$  from  $i'$  is independent of the number of suitable sites. Hence,  $H$  either contains all suitable sites (directed dispersal) or all suitable and unsuitable sites (random dispersal).

In total, Eqs. 3–5 give four dispersal modes: local directed (LD), global directed (GD), local random (LR), and global random (GR), where LD dispersal best describes the observations of jay-mediated seed dispersal (Gomez 2003).

#### Seed establishment

There are several different ways to incorporate environmental forcing into a SPOM model. We have chosen to restrict the influence of the environment to fire frequency ( $\phi_i$ ) and seed establishment ( $\alpha_{i,j}$ ). Here, we model  $Y_{j,i} = \text{logit}(\alpha_{i,j})$  as a function of three site-specific physical and/or climatic variables,  $V_{1,j}(i) \cdots V_{3,j}(i)$ , such that

$$\alpha_{j,i} = 1/[1 + \exp(-Y_{j,i})] \quad (6a)$$

$$Y_{j,i} = \kappa_j + \sigma_{1,j}[V_{1,j}(i) - \mu_{1,j}]^2 + \sigma_{2,j}[V_{2,j}(i) - \mu_{2,j}]^2 + \sigma_{3,j}[V_{3,j}(i) - \mu_{3,j}]^2. \quad (6b)$$

Eq. 6 provides for a very flexible relationship between the predictor variables  $V_{1,j}(i) \cdots V_{3,j}(i)$  and the probability of seed establishment. The parameter  $\kappa_j$  sets the value of  $\alpha_{j,i}$  when  $V_{1,j}(i) = \mu_{1,j}$ ,  $V_{2,j}(i) = \mu_{2,j}$ , and  $V_{3,j}(i) = \mu_{3,j}$ . The parameters  $\sigma_{1,j} \cdots \sigma_{3,j}$  set the direction and sensitivity of  $\alpha_{j,i}$  to variation in  $V_{1,j}(i) \cdots V_{3,j}(i)$  away from  $\mu_{1,j} \cdots \mu_{3,j}$ . When  $\sigma_{1,j} \cdots \sigma_{3,j}$  are negative,  $\alpha_{j,i}$  takes its maximum when  $V_{1,j}(i) = \mu_{1,j}$ ,  $V_{2,j}(i) = \mu_{2,j}$ , and  $V_{3,j}(i) = \mu_{3,j}$ , and thus the parameters  $\mu_{1,j} \cdots \mu_{3,j}$  essentially describe the environmental conditions for which establishment is

maximal. We also allowed positive  $\sigma$  values in the parameter estimation, but most  $\sigma$  values were estimated to be significantly less than zero, and so this interpretation of Eq. 6 holds for most species–variable combinations. Importantly, we also allowed  $\mu_{x,j}$  to lie outside the observed range of  $V_{x,j}$ . In this case,  $\alpha_{j,i}$  is either monotonically decreasing or increasing over the  $V_{x,j}$  range, so Eq. 6 does not constrain the relationship between physical variables and seed establishment to have internal maxima or minima.

The set of predictor variables  $V_{1,j}(i) \cdots V_{3,j}(i)$  for each species  $j$  was selected according to a preliminary gradient analysis that identified the three factors most strongly correlated with the observed presence/absence of each species (Appendix A and Table 3). The variables selected differed between the species (see Table 3). Drought length (number of months when PET exceeds precipitation) was included for all three species. In addition, the deciduous *Q. faginea* depended on seasonality (CV) of precipitation, and annual average potential evapotranspiration (PET) which is a combined measure of temperature and humidity (high PET signifies high temperatures and low humidity). The evergreens depended on altitude and PET (*Q. ilex*), and on altitude and annual mean temperature (*Q. coccifera*).

For each site and species the values of  $V_{1,j}(i) \cdots V_{3,j}(i)$  were generated by spatial interpolation of data from weather stations (Appendix A), yielding values for each survey plot  $q$ . To set the environmental conditions at the model sites, we used the same rules as for assigning observed presence or absence and thus

$$V_{x,i} = V_{x,\hat{q}(i)} \quad (7)$$

where  $\hat{q}(i)$  is the nearest survey plot to model site  $i$ .

#### PARAMETER ESTIMATION

There are four groups of parameters that reflect various components of the SPOM model, including disturbance, dispersal characteristics, seed rain, and seed establishment. We use fixed values, based on the literature and supplementary data, for the disturbance and dispersal parameters. The seed output and establishment parameters are estimated from the survey data using a Bayesian approach.

##### Fixed parameters

The most important stochastic disturbance in the study region is wildfires, which, in the current landscape, have an average return interval of 10–100 years (Naveh 1990). We used a statistical model to predict fire frequency for each location from the drought length, annual precipitation, and altitude at that location (see Appendix C). This model gives the disturbance probability  $\phi_i$  for each site  $i$  based on the conditions at site  $i$ . The fire model was parameterized separately from additional sources of data (see Appendix C), and its parameters were fixed upon incorporating it into the SPOM model. The parameter values for the local

dispersal kernel were also fixed ( $a = 0.11$ ,  $b = 0.60$ ), and were chosen to reproduce the jay-mediated acorn movements documented by Gomez (2003). Other types of dispersal are likely to operate within sites (Bossema 1979), but these are not captured here because the SPOM describes between patch dynamics only. The seed movements reported by Gomez (2003) were typically less than 500 m. Hence, in the SPOM, local dispersal results in seeds being primarily moved between neighboring cells. However, the parameter values and the functional form of the local dispersal kernel yield a “long tail,” thus some acorn movement occurs over several to tens of kilometers.

##### Free parameters

The SPOM for species  $j$  contains eight free parameters:  $\kappa_j$ ,  $\mu_{1,j} \cdots \mu_{3,j}$ ,  $\sigma_{1,j} \cdots \sigma_{3,j}$ , and  $\omega_j$ . These parameters are estimated by a Bayesian analysis of the Inventario Forestal Nacional presence–absence data. A necessary part of the analysis is an analytical description of the likelihood function, which quantifies the likelihood of the data, conditional on the model constraints and parameters. To implement the analysis efficiently, we made some key simplifying assumptions, generating a pseudo-likelihood function that can be calculated rapidly, but which is an approximation to the true likelihood.

*Simplifying assumptions.*—We denote the observed presence/absence data for species  $j$  in survey plot  $q$  as  $Z_{j,q}^{\text{obs}}$ . The total data set for species  $j$  is the vector  $\mathbf{Z}_j^{\text{obs}}$  of 12,047 ones and zeros for presence/absence in each survey plot. Conditional on the SPOM assumptions and parameter set  $\theta = (\phi, a, b, c, \kappa_j, \mu_{1,j} \cdots \mu_{3,j}, \sigma_{1,j} \cdots \sigma_{3,j}, \omega_j)$ , each  $Z_{j,q}^{\text{obs}}$  is assumed to arise from an independent Bernoulli process. Thus, the log-likelihood function for  $\mathbf{Z}_j^{\text{obs}}$  is

$$\ell\{\mathbf{Z}_j^{\text{obs}}|\theta\} = \sum_q [Z_{j,q}^{\text{obs}} \ln(P_{j,q,\theta}) + (1 - Z_{j,q}^{\text{obs}}) \ln(1 - P_{j,q,\theta})] \quad (8)$$

where  $P_{j,q,\theta}$  is the probability that species  $j$  will be found in plot  $q$ , or equivalently, the fraction of time that plot  $q$  is expected to be occupied by species  $j$  (see Hanski 1994, Etienne et al. 2004). The most direct way to estimate the  $P_{j,q,\theta}$  is to implement the SPOM with parameter set  $\theta$ , and run the model for a sufficient number of iterations to reach equilibrium and give an accurate assessment of  $P_{j,q,\theta}$ . However, a new model simulation must be implemented for each new  $\theta$ , and the necessary computing time prohibited this approach in this study. Instead, we calculate  $P_{j,i,\theta}$  by noting that, at equilibrium, the probability of extinction is equal to the probability of colonization:

$$\phi_i P_{j,i,\theta} = [1 - (1 - \alpha_{j,i}) \bar{S}_{j,i}^i] (1 - P_{j,i,\theta}) \quad (9)$$

where  $\bar{S}_{j,i}$  is the average annual seed rain of species  $j$  into model site  $i$ , and the solution for  $P_{j,i,\theta}$  is

$$P_{j,i,\theta} = \left[ 1 + \frac{\Phi_i}{1 - (1 - \alpha_{j,i})^{\bar{S}_{j,i}}} \right]^{-1}. \quad (10)$$

The expected probabilities for each survey site are obtained by assuming

$$P_{j,q,\theta} = P_{j,\hat{i}(q),\theta} \quad (11)$$

where  $\hat{i}(q)$  is the nearest model site  $i$  to survey plot  $q$ .

Here, a problem arises because  $S_{j,i}$  and  $P_{j,i,\theta}$  are both functions of each other. There is no exact solution to this problem. Methods are becoming available to deal with this issue, but these require repeated surveys, the use of repeated model simulations, or the use of latent variables; the latter two options would make the analysis much more complex than the one presented here (O'Hara et al. 2002, ter Braak and Etienne 2003, Etienne et al. 2004). In this study, surveys were not repeated and extra simulations were unfeasible. Instead, we reduce the computational requirements by initializing the SPOM with observed presence/absence data:

$$Z_{j,i}^{(\text{obs})} = Z_{j,\hat{i}(i)}^{(\text{obs})}. \quad (12)$$

From this pattern of presence/absences in the model, we calculate  $S_{j,i}$  using Eqs. 2–5, which we denote  $S_{j,i}^{(\text{obs})}$ . Finally we set

$$\bar{S}_{j,i} = S_{j,i}^{(\text{obs})}. \quad (13)$$

Thus, the average seed rain into each model site is assumed to be equal to the seed rain calculated from the model when initialized with the observed data. Because of this simplification, this method yields an estimate of the log-likelihood, not the true value. If the species distributions are close to their equilibrium state, this approach is not thought to introduce significant errors (Hanski 1994), a view which is supported by the close fit to observations obtained here.

*Bayesian analysis.*—We implemented a Bayesian analysis because it conveniently gives point estimates for the free parameters, and it explicitly quantifies parameter uncertainty. Although both Bayesian (e.g., Gelman et al. 2004) and maximum likelihood (ML) methods (e.g., Hilborn and Mangel 1997) share the same likelihood function and often yield similar results, the Bayesian approach differs from ML in a few key points. Most important, the free parameters in  $\theta$  are treated as random quantities, so uncertainty in these parameters can be directly incorporated into simulation results. The product of a Bayesian analysis is a joint posterior distribution for the parameters given the data  $f(\theta_j | Z_j^{\text{obs}})$ , and this posterior is proportional to the likelihood times the prior such that  $f(\theta_j | Z_j^{\text{obs}}) \propto \exp[\ell(Z_j^{\text{obs}} | \theta_j)] \cdot f(\theta_j)$  (e.g., Gelman et al. 2004). Thus, the priors allow the incorporation of previous information about likely parameter values, but they can also be chosen to be non-informative, as here (i.e., we chose flat priors such that  $f(\theta_j) \propto C$ , where  $C$  is a constant). We subscript  $\theta$  by  $j$

to explicitly indicate that each species is associated with its own set of parameter values.

The SPOM returns a likelihood function that is highly nonlinear, and analytical solutions for the joint posterior density cannot be derived easily. Therefore, we implemented a Metropolis-Hastings (M-H) Markov chain Monte Carlo (e.g., Chib and Greenberg 1995, Robert and Casella 1999) numerical algorithm for sampling from  $f(\theta_j | Z_j^{\text{obs}})$ . From these samples, we can calculate measures of parameter centrality (e.g., mean, median, mode) and spread (e.g., credible intervals, which are much like confidence intervals). To define  $f(\theta_j)$  we note that  $\mu_{1,j} \cdots \mu_{3,j}$  can take any value between  $\pm\infty$ , but the other parameters ( $\sigma_{1,j} \cdots \sigma_{3,j}$ ,  $\kappa_j$ ,  $\omega_j$ ) are all positive valued. Thus, we chose flat priors for  $\mu_{1,j} \cdots \mu_{3,j}$ ,  $\ln(\sigma_{1,j}) \cdots \ln(\sigma_{3,j})$ ,  $\ln(\kappa_j)$ , and  $\ln(\omega_j)$ .

Separate M-H simulations were conducted for each of the three oak species and for each dispersal mode (i.e., LD, GD, LR, GR). For each of these 12 SPOMs, an initial burn-in of 10 000 iterations was required to eliminate the effects of starting conditions. After the M-H samples converged to the posterior distribution, an additional 10 000 iterations were run and every 50th parameter set was stored, providing an independent sample of 200 from  $f(\theta_j | Z_j^{\text{obs}})$ . We use these samples to conduct simulation experiments with the SPOM, thereby accounting for parameter uncertainty, preserving the covariance between parameters.

#### MODEL SIMULATIONS AND ANALYSIS

To assess uncertainty in model predictions and to estimate the effects of dispersal, we conducted a Monte Carlo simulation using the SPOM. A comparison of the four dispersal modes indicated that the local directed (LD) dispersal mode gave a much better fit to the survey data than did the other dispersal modes (see Table 2). Therefore, for each of  $m = 1, \dots, 40$  Monte Carlo simulations,  $\theta$  was drawn at random from the empirical posterior density generated for LD dispersal. For each parameter combination, we ran separate simulations for each of the four dispersal modes. Thus, parameters governing seed output and seed establishment were held the same regardless of dispersal mode, allowing us to infer the direct effects of dispersal mode alone on the distribution of the three oak species.

Each of the 40 model simulations began with an initial random state that was generated by assuming a 1% chance that each site is occupied. The SPOM for each simulation was run for 3000 iterations to ensure that equilibrium abundances and spatial distributions were reached. Within a simulation, the parameters were held constant for all iterations. This resulted in equilibrium species distributions that were comparable to observed patterns. We also ran simulations beginning with a 50% and 90% chance of occupancy. This had no measurable effect on the model equilibria, indicating that the equilibrium was not dependent on starting conditions. For each simulation  $m$  and species  $j$ , we recorded the

TABLE 2. Model fits for the stochastic patch occupancy model (SPOM) for three *Quercus* species under each of four dispersal modes: local directed (LD), global directed (GD), local random (LR), and global random (GR).

Species	Dispersal mode			
	LD	LR	GD	GR
<i>Q. faginea</i>	<b>8739.7</b>	8895.8	11 824.9	11 824.2
<i>Q. ilex</i>	<b>9538.7</b>	12 420.2	12 089.6	13 612.4
<i>Q. coccifera</i>	<b>5419.9</b>	5547.0	6543.0	6414.3

Notes: Table entries are  $D$  (deviance) values. The mean posterior deviance ( $\bar{D}$ ), is calculated for each model by averaging the log-likelihood over all  $j = 1, \dots, 200$  posterior samples of  $\theta$ , such that  $D = -2 \times (1/200) \sum_{k=1}^{200} \ell(Z_j^{\text{obs}} | \theta_j^{(k)})$ . Lower  $D$  values indicate a superior model fit (Gelman et al. 2004). This index is appropriate for model comparison here because the dimensionality of the parameter set is the same for each model (for models with unequal numbers of parameters, see Spiegelhalter et al. [2002]). For each species, the model with the best fit is shown in boldface. The smallest difference between the best model and the next best model is 114 units (*Q. faginea*, LD vs. LR dispersal). This difference corresponds to  $\sim 57$  log-likelihood units, which is a highly significant difference in fit, given that the number of parameters is the same for all models (Hilborn and Mangel 1997). The differences for *Q. faginea* and *Q. ilex* are larger than 114 units.

average species frequency  $F_{j,m}$  at the end of the 3000 model iterations:

$$F_{j,m} = (1/n) \sum_{i \in R} Z_{j,i} \quad (14)$$

where the set  $R$  contains all suitable sites  $i$ , and  $n = 11\,039$  is the total number of suitable sites. The values of  $F_{j,m}$  for each dispersal mode are denoted by  $F_{j,m}^{\text{LD}} \dots F_{j,m}^{\text{GR}}$ . We also recorded the distribution as a function of altitude  $F_{j,m,\text{alt}}$  and drought length  $F_{j,m,\text{dl}}$ :

$$F_{j,m,\text{alt}}(x) = [1/n(x)] \sum_{i \in R(x)} Z_{j,i} \quad (15)$$

$$F_{j,m,\text{dl}}(w) = [1/n(w)] \sum_{i \in R(w)} Z_{j,i} \quad (16)$$

where  $x$  refers to altitude (m) and  $w$  to drought length (months); the sets  $R(x)$  and  $R(w)$  contain all suitable sites  $i$  located with altitude between  $x - 50$  and  $x + 50$  m or with drought length between  $w - 0.25$  and  $w + 0.25$  months, respectively; and  $n(x)$  and  $n(w)$  are the numbers of sites in  $R(x)$  and  $R(w)$ . The values specific to each dispersal mode are denoted by  $F_{j,m,\text{alt}}^{\text{LD}} \dots F_{j,m,\text{alt}}^{\text{GR}}(x)$ , and  $F_{j,m,\text{dl}}^{\text{LD}} \dots F_{j,m,\text{dl}}^{\text{GR}}(w)$ .

To calculate the effect of dispersal mode on the mean frequencies  $\bar{F}_j$  and distributions  $F_{j,\text{alt}}$  and  $F_{j,\text{dl}}$ , we compared the values given by GD and LR dispersal with those given by LD dispersal for each  $m$ :

$$\Delta F_{j,m}^y = F_{j,m}^{\text{LD}} - F_{j,m}^y \quad (17)$$

$$\Delta F_{j,m,b}^y(x) = F_{j,m,b}^{\text{LD}}(x) - F_{j,m,b}^y(x) \quad (18)$$

where  $\Delta F_{j,m}^y$  is the difference in the average frequency given by LD dispersal mode relative to dispersal mode  $y$  (=LR, GD, or GR). Similarly,  $\Delta F_{j,m,b}^y(x)$  is the difference in frequencies at altitude or drought length  $x$  ( $b = \text{alt}$  or  $b = \text{dl}$ , respectively). This pairwise procedure was conducted to standardize for the effects of parameter uncertainty (e.g., comparing frequencies after they have been averaged across all  $m$  simulations would have underestimated the significance of the effect of dispersal mode; see Fig. 2, a vs. b). The effect of local (or directed) dispersal is given by the  $\Delta F$  values comparing simulations with and without this effect. For example, LD minus GD measures the effect of local dispersal, and LD minus LR measures the effect of directed dispersal. Observed values for  $F_j$ ,  $F_{j,m,\text{alt}}$ , and  $F_{j,m,\text{dl}}$  were also calculated from the survey data.

### Spatial statistics

Spatial structure in the observed and simulated distributions was quantified using a modification of the spatial covariance functions given in Purves and Law (2002). These statistics give a value for the auto-covariance of species  $j$  at a spatial lag  $r$ ,  $C_j(r)$ , which we compare to the expected correlation under spatial randomness  $E\{C_j(r)\}$ . The ratio of these two quantities gives a dimensionless measure of departure from spatial randomness  $\Omega_j(r)$  (Condit et al. 2000). To correct for apparent differences in spatial structure due to differences in average frequency, we divided  $\Omega_j(r)$  by the maximum possible value given the current frequency, to give  $\Omega_j^{(n)}(r)$ :

$$\Omega_j^{(n)}(r) = [\Omega_j(r) - 1] / (F_j^2 - 1) \quad (19)$$

where  $F_j$  is the fraction of suitable sites occupied by species  $j$  (observed, or in a given model simulation). A value of  $\Omega_j^{(n)}(r) > 0$  indicates aggregation of species  $j$  at spatial lag  $r$ ,  $\Omega_j^{(n)}(r) < 0$  indicates segregation, and  $\Omega_j^{(n)}(r) \cong 0$  indicates spatial randomness. Pairwise comparisons of  $\Omega$  from model simulations with different dispersal modes, similar to those described in Eqs. 17–18, were used to calculate the effect of dispersal on spatial structure. The  $\Omega$  statistic was used because it is simpler than some alternatives (e.g., semivariance, Ripley's  $k$  [Ripley 1981]), but it is likely to have yielded similar results to them. The estimates of  $\Omega$  do not depend on the arrangement of survey plots.

### Perturbation experiments

We also evaluated the sensitivity of the regional population abundance of the species to altered drought length, disturbance rate, and habitat cover (Appendix B). All three of these features were varied relative to their observed values. Drought length was manipulated by applying the following transformation:

$$V_{\text{dl},i} = V_{\text{dl},\hat{q}(i)} + \Delta_{\text{dl}} \quad (20)$$

where  $V_{\text{dl},i}$  is constrained to lie between 0 and 12



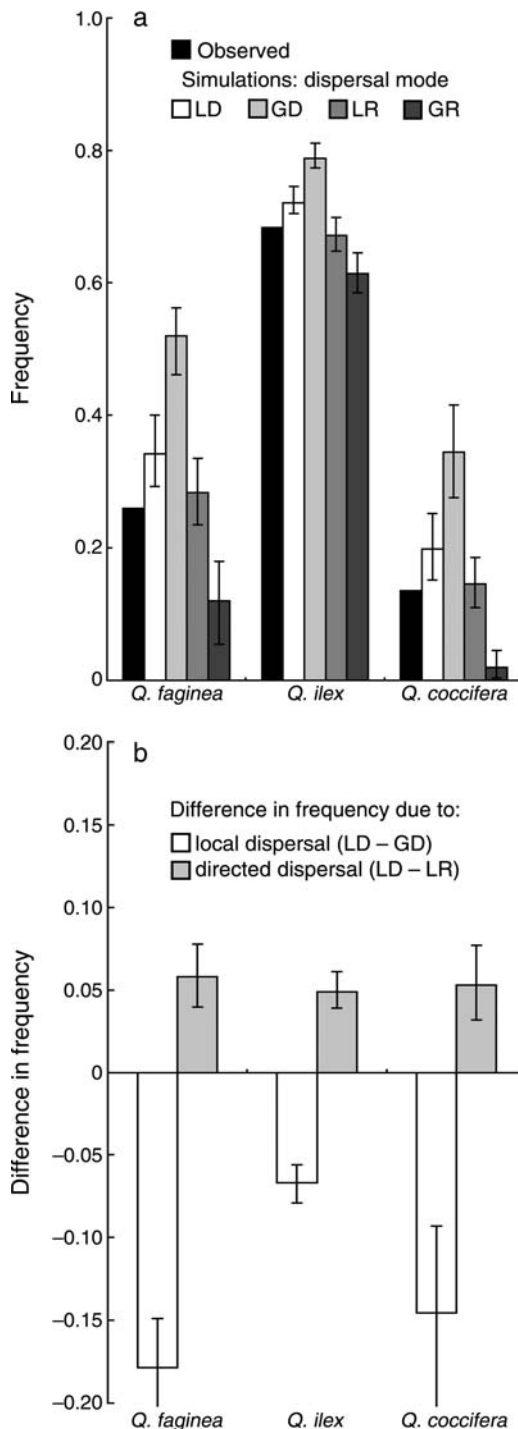


FIG. 2. Mean frequency of three *Quercus* species. (a) Observed frequency and model simulations with different dispersal modes: local directed (LD), global directed (GD), local random (LR), and global random (GR). Simulations differed in dispersal mode only: all simulations used parameters estimated using LD dispersal, so the apparent superior fit of LR dispersal is not relevant (for all three species, the parameters generated using LR dispersal lead to an inferior fit for all three species; Table 2). Error bars are envelopes containing 95% of the simulation results. Variation in model output reflects both

months, and the change in drought length was varied between  $\Delta_{di} = -4$  to  $+5$  months. Disturbance rate  $\phi_i$  was manipulated by applying the transformation

$$\theta_i = \Delta_{fire} \theta_{i(t)} \quad (21)$$

subject to the constraint  $0 < \theta_i < 1$ ; with  $\Delta_{fire} = 1/3, 2/3, 1, 2,$  or  $3$ . For each perturbation experiment, 10 simulations were implemented, with parameters drawn from the posterior density each time.

Habitat cover, which refers to the fraction of all model sites  $i$  designated as suitable, was varied from 5% to 95%, with the observed value, estimated from survey plots, being 48%. The habitat cover changes were implemented using two different methods, both of which began with the observed spatial arrangement of suitable sites (Appendix B: Fig. B1). The first method added or removed sites at random until the target cover was obtained. This tended to give a highly fragmented pattern of suitable habitat. The second method made the addition or removal of sites much more likely at suitable-unsuitable interfaces. This tended to give a highly aggregated pattern of suitable habitat, which more accurately reflects the observed pattern (Fig. 1).

## RESULTS

### *Fits to survey data*

*Model comparisons.*—We compared how well the SPOM with the four different dispersal modes fit the survey data for each *Quercus* species (Table 2). For all three species, local dispersal modes (LD and LR) gave a much better fit to the data than global dispersal modes (GD and GR; Table 2). And overall, local directed (LD) dispersal was far superior to the other three models (Table 2). Because the LD model provided the best fit to the survey data and because LD dispersal is most consistent with field observations of jay-mediated acorn movements (Gomez 2003), we consider the parameter estimates for the SPOM with LD dispersal to be those that correspond to reality most closely. Thus, we use the posterior distributions of the LD parameters in all simulation experiments, and parameter estimates generated using other dispersal modes are not considered further.

*Parameter estimates for LD model.*—Parameter estimates for the SPOM with LD dispersal are given in Table 3 for the three oak species. In general, the parameters were well constrained by the data (Table 3), as indicated by fairly narrow credible intervals. Across the three species, estimated seed output ranged from approximately 45 to 1900 seeds per occupied site (see  $\omega$ ; Table 3). These values may appear small given the size of

inherent model stochasticity and parameter uncertainty. (b) Effect of dispersal mode on mean frequency, estimated by calculating pairwise differences between simulations with identical parameter values but different dispersal modes (see *Methods*).

TABLE 3. Bayesian estimates of model parameters associated with the SPOM model, with local directed (LD) dispersal.

Estimate, by species				Drought length (months)†		Annual PET (mm)‡		Annual mean temp. (°C)§	
	$\omega$	$\kappa$	$\omega \exp(\kappa)$	$\mu_1$	$\sigma_1$	$\mu_2$	$\sigma_1$	$\mu_2$	$\sigma_2$
<i>Q. faginea</i>									
Mean	44.5	-8.76	0.0070	5.68 (15)	-0.209	-1106.0 (83)	$-1.05 \times 10^{-5}$		
Lower	23.5	-8.28	0.0105	5.15	-0.254	-1018.4	$-1.65 \times 10^{-5}$		
Upper	79.9	-9.59	0.0033	6.19	-0.142	-1214.4	$-3.16 \times 10^{-6}$		
<i>Q. ilex</i>									
Mean	356.8	-9.51	0.0264	5.70 (16)	-0.244	1063.2 (71)	$-1.88 \times 10^{-5}$		
Lower	226.3	-10.08	0.0172	5.48	-0.261	1020.7	$-2.72 \times 10^{-5}$		
Upper	527.4	-9.00	0.0419	5.97	-0.218	1108.6	$-1.18 \times 10^{-5}$		
<i>Q. coccifera</i>									
Mean	1898.5	-12.32	0.0068	7.65 (66)	-0.136			12.9 (59)	-0.102
Lower	270.2	-13.37	0.0055	6.95	-0.212			11.7	-0.154
Upper	4790.6	-10.65	0.0086	8.20	-0.0612			14.0	-0.0293

Notes: For each parameter, the posterior mean and 95% credible interval limits (defined by the lower and upper 2.5th percentiles of the posterior samples) are given. The numbers in parentheses given with mean  $\mu$  values are the percentages of plots with a  $V$  value below the corresponding  $\mu$  value. Therefore (50) indicates that the  $\mu$  value is close to the median, and (0) and (100) indicate that the  $\mu$  value is outside the range of  $V$  values reported in the data, implying a simple increasing or decreasing relationship between  $\sigma_{j,i}$  and  $V$ . The quantity  $\omega \exp(\kappa)$  is a measure of maximum fecundity. A value of  $\omega \exp(\kappa)$  was calculated from each of the samples from the posterior distribution provided by the Bayesian analysis, and the mean and credible intervals were taken from this list of values.

† Range 0.5–11.5 months.

‡ Potential evapotranspiration; range 548.5–1655.7 mm.

§ Range 6.14–23.9.

¶ Range 28.18–79.56%.

|| Range 300–2100 m.

the patches. For example 527 seeds per 4 km<sup>2</sup>, which is the upper 95th percentile of  $\omega$  value for *Q. ilex*, implies an output of 0.00013 seeds·m<sup>-2</sup>·yr<sup>-1</sup>. However, the SPOM model only follows seeds that enter into inter-patch movements, whereas the great majority of seeds are expected to remain within their native patch. Seed output was estimated to be in the order *Q. faginea* < *Q. ilex* << *Q. coccifera*, with the differences being statistically significant (Table 3, credible intervals for  $\omega$  for one species do not contain the mean  $\omega$  for other species). However, the maximum probability of seed establishment, which is given by  $\exp(\kappa)$ , differed substantially in the opposite order: *Q. coccifera* < *Q. ilex* < *Q. faginea* (Table 3). A measure of maximum effective fecundity is given by the product  $\omega \exp(\kappa)$ , which varied in the order *Q. coccifera*  $\approx$  *Q. faginea* < *Q. ilex* (see Table 3).

The parameter estimates also suggest that the three species are segregated along environmental gradients with respect to requirements for seedling establishment. For example, the drought length that maximizes seed establishment is significantly greater for *Q. coccifera* than for *Q. faginea* or *Q. ilex*, as indicated by the fact that the  $\mu_1$  credible intervals for *Q. coccifera* do not contain the posterior mean of the other species, and vice versa (Table 3). According to these estimates, the species rank in the following order, from least to most drought adapted: *Q. ilex*  $\approx$  *Q. faginea* < *Q. coccifera* (Table 3). Moreover, seed establishment in *Q. coccifera* was estimated to be less sensitive to drought length than the other two species (Table 3,  $\sigma_1$  smaller for *Q. coccifera*). In addition, the evergreens *Q. ilex* and *Q. coccifera* appear to be segregated with respect to

altitude, with establishment greatest at high elevations for *Q. ilex* (Table 3,  $\mu_3 = 1455.5$ ) and at intermediate elevations for *Q. coccifera* ( $\mu_3 = 913.4$ ).

*Reproducing observed patterns.*—For all three species, the equilibrium average frequency given by the SPOM with LD dispersal was slightly greater than observed (see LD bars; Fig. 2a). This was unexpected because the parameter estimates used in the simulations were obtained by fitting the LD model to the survey data. However, this bias is possible because of an approximation of the likelihood function used to generate the posterior distributions of the parameters (see Eqs. 8–13). Except for this general overestimation, the spatial distributions given by simulations of the SPOM corresponded closely to the observed distributions (Fig. 3).

With LD dispersal, the SPOM reproduced the distribution of the species along gradients of altitude and drought length quite accurately, with the majority of observations falling within the range predicted by the model (Fig. 4). Discrepancies can be attributed to the positive bias discussed above, but this effect was approximately constant across the gradients in altitude and drought length.

The model also successfully reproduced the spatial structure of each species' distribution (Fig. 4), capturing differences in the nature, intensity and scale of spatial structure between the species. The model slightly overestimated the intensity of spatial structure for *Q. coccifera*, but otherwise the model and observations showed excellent agreement. Importantly, for *Q. faginea* and *Q. ilex* the SPOM reproduced the spatial structure

TABLE 3. Extended.

CV precipitation (%)¶		Altitude (m)	
$\mu_3$	$\sigma_3$	$\mu_3$	$\sigma_3$
79.47 (99)	$5.7 \times 10^{-4}$		
57.51	$3.1 \times 10^{-4}$		
106.00	$9.5 \times 10^{-4}$		
		1455.5 (93)	$1.72 \times 10^{-6}$
		1203.4	$9.31 \times 10^{-7}$
		1754.2	$2.58 \times 10^{-6}$
		913.4 (56)	$-1.43 \times 10^{-6}$
		159.7	$-3.69 \times 10^{-6}$
		1415.0	$7.80 \times 10^{-7}$

much more accurately than did a gradient model with the same predictor variables (Fig. A1 in Appendix A). This indicates that their spatial structure could not be explained by the effects of environmental heterogeneity alone (gradient model), but could be explained as an

interaction between environmental heterogeneity and metapopulation dynamics (SPOM).

#### Simulation experiments

*Alternative dispersal modes.*—The Monte Carlo simulations conducted with the SPOM for the four different dispersal modes, using the parameter estimates from the LD fits, are summarized in Fig. 2. For mean abundance, the LD, LR, and GD model predictions gave abundances close to, or greater than the observed value for each species (Fig. 2a). The most accurate prediction of abundance was given using LR dispersal, which might appear to contradict the finding that LD dispersal gave a superior fit to data (Table 1). However, the model fit for LR dispersal (Table 1) refers to parameter estimates generated assuming LR dispersal, whereas these simulations (Fig. 2a) used parameters generated by assuming LD dispersal (see methods). The explanation of the apparent superior fit of the LR model in this case is that the positive bias has led to an overestimate of abundance by the best model (LD), which has then

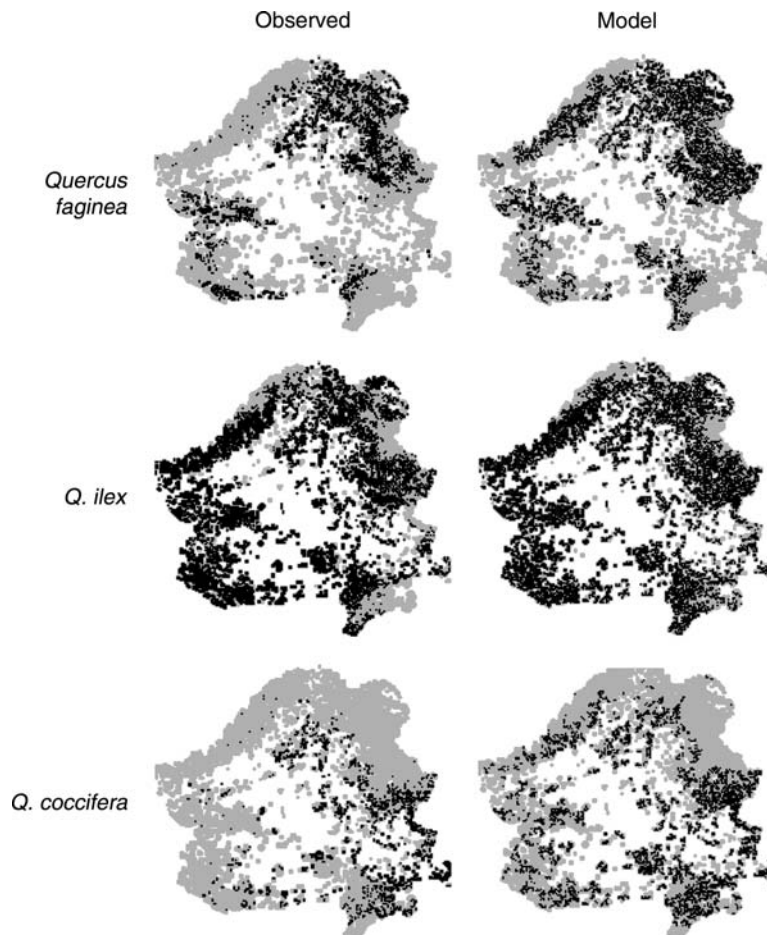


FIG. 3. Comparison of the observed spatial distribution of three *Quercus* species with output from the stochastic patch occupancy model (SPOM) with local directed (LD) dispersal, using the mean of the Bayesian posterior distribution for each parameter value. The output is a snapshot from a single realization of the model, after sufficient iterations to reach quasi-equilibrium. Parameters were estimated using the survey data, so this comparison is not an independent model validation.

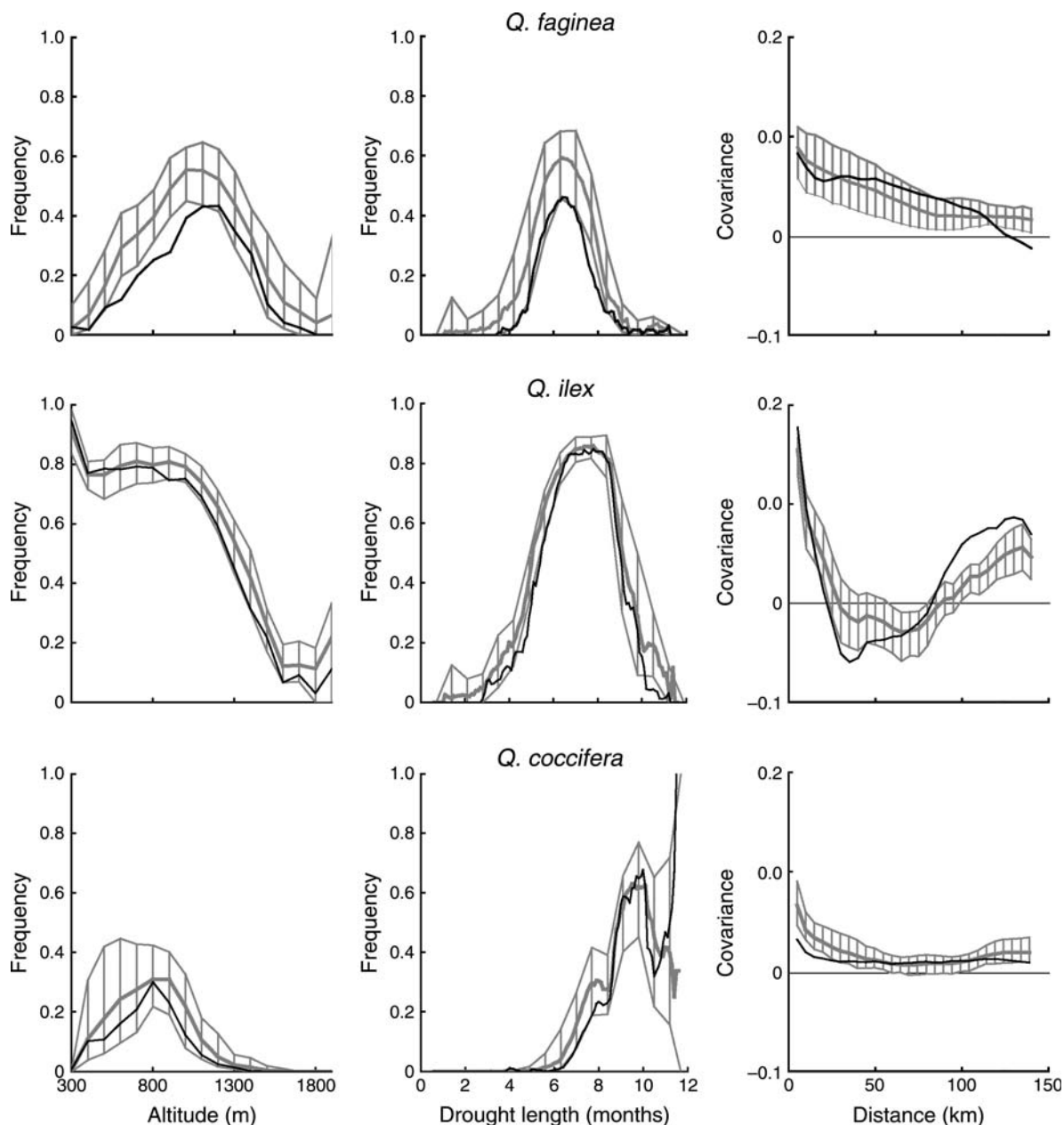


FIG. 4. Distributions along environmental gradients (left, middle) and spatial autocorrelation (right): observed (black) and model simulations with LD dispersal (gray). For simulation results, the heavy gray line is the mean over repeated simulations, and the lighter gray lines define an envelope containing 95% of the simulation results.

mostly been compensated for by the reduction in abundance caused by changing dispersal to random rather than directed.

We evaluated the importance of local vs. global dispersal (given directed dispersal), and directed vs. random dispersal (given local dispersal) using paired simulations (Eqs. 17–18). For all three species, LD dispersal gave a lower frequency than GD dispersal (Fig. 2b). These differences were statistically significant (95% intervals did not include zero). In relative terms, this

effect was much greater for *Q. faginea* and *Q. coccifera* (reduction in frequency 28% and 42% respectively; Fig. 2a) than it was for *Q. ilex* (7% reduction). Conversely, LD dispersal resulted in significantly greater abundances than LR dispersal for all three species (Fig. 2). Again, the relative magnitude of this effect was larger for *Q. faginea* (+22%) and *Q. coccifera* (+37%) than for *Q. ilex* (+7%), although the absolute effect was very similar for all three species (Fig. 2). The simulations also identified a significant interaction between local and directed

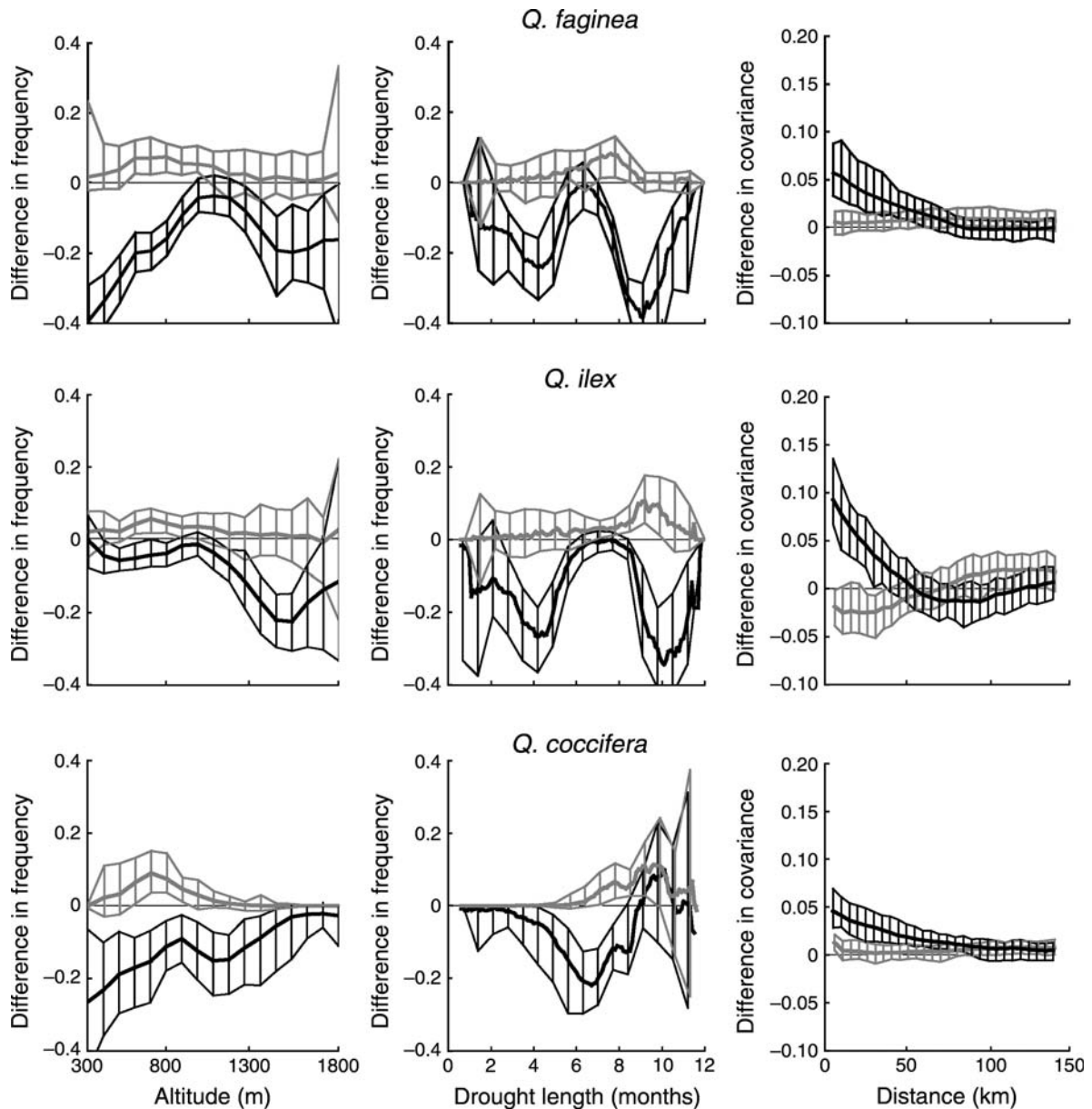


FIG. 5. Estimated effect of dispersal mode on distributions along environmental gradients (left, middle), and spatial autocorrelation (right). The effect of local dispersal (LD minus GD) is shown in black; the effect of directed dispersal (LD minus LR) is shown in gray. In both cases, the heavy line is the mean over repeated simulations, and the lighter lines define an envelope containing 95% of the pairwise differences (see Eqs. 17–18 and accompanying text). Envelopes that do not cross zero indicate a significant effect of dispersal mode.

dispersal: when dispersal was directed toward suitable patches, local dispersal reduced the regional frequency of all three species (LD vs. GD in Fig. 2a). Conversely, when dispersal was random (i.e., occurred independent of patch suitability), local dispersal increased regional frequency (LR vs. GR).

The simulations also indicated that both local and directed dispersal alter the apparent responses of the species to environmental gradients (Fig. 5). Local dispersal acts to reduce frequencies in marginal sites

(i.e., areas where altitude or drought length differ greatly from the values that maximize seed establishment), beyond the low levels predicted by global dispersal. On the other hand, local dispersal either increases (*Q. coccifera*) or does not affect (*Q. faginea*, *Q. ilex*) abundances in favorable sites (Fig. 5). In this way, local dispersal amplifies the apparent responses of the species to environmental gradients. The magnitude of this effect was large in some cases. For example, the simulations estimate that local dispersal, resulting in dispersal

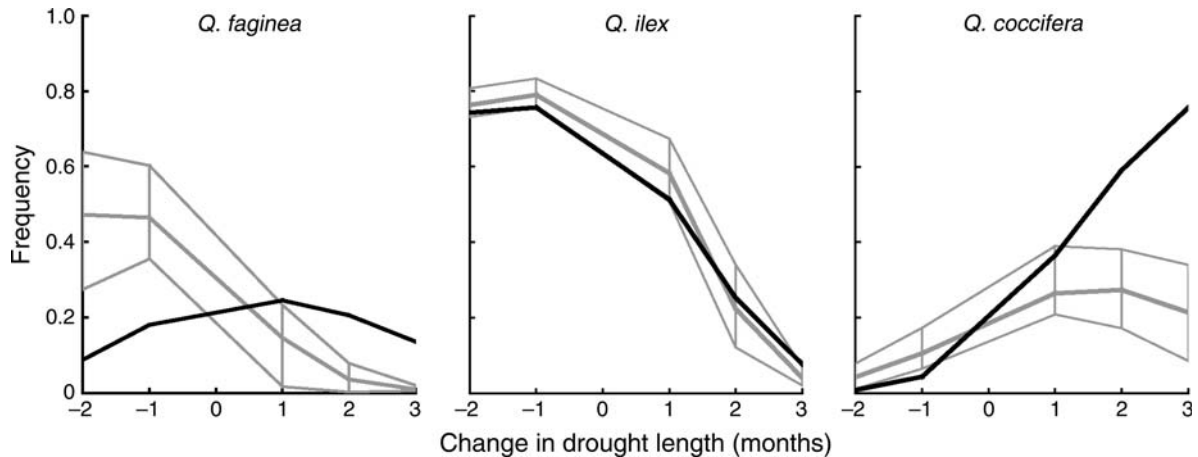


FIG. 6. Equilibrium average frequency in simulations with altered drought length, given by the SPOM with LD dispersal (gray) and by the gradient model (black). For simulation results, heavy lines are the mean over 10 simulations, and lighter lines give the range. The simulations used LD dispersal, with parameters estimated with respect to LD dispersal. Analogous simulation results using different dispersal modes are given in Appendix B (Fig. B2).

limitation, reduces the frequency of *Q. faginea* in marginal sites from around 0.4 (GD) to around 0.04 (LD; Fig. 5). Directed dispersal increases frequency disproportionately at altitudes close to 800 m and drought lengths of 8–10 months (Fig. 5).

Dispersal mode also had significant effects on regional-scale spatial structure (Fig. 5). Local dispersal is estimated to cause a substantial increase in the intensity of spatial aggregation of all three species (Fig. 5). Directed dispersal had mixed effects on short-scale spatial structure. For example, the intensity of aggregation was decreased for *Q. ilex*, slightly increased for *Q. coccifera*, and unchanged for *Q. faginea* (Fig. 5); but for all three species, directed dispersal increased the intensity of aggregation at larger distances (>75 km).

*Perturbation experiments.*—Simulations of the SPOM estimated that the frequencies of all three species are sensitive to changes in drought length (Fig. 6). A one-month increase in drought length reduced the average predicted frequency of *Q. faginea* and *Q. ilex* by 47% and 24% respectively, and increased *Q. coccifera* by 31% (Fig. 6). With drought length extended by three months, *Q. faginea* and *Q. ilex* were predicted to fall to near extinction (–96% reduction for both species; Fig. 6) and *Q. coccifera* was predicted to attain a lower frequency than it would under a one-month increase (+21% compared to current frequency; Fig. 6). Importantly, for two of the species (*Q. faginea* and *Q. coccifera*) the predicted responses to drought length from the SPOM did not match the predictions of the gradient model in either nature or magnitude (Fig. 6). For *Q. ilex*, the gradient and SPOM models were in very close agreement, with both predicting near extinction with a drought length change of +3 months (Fig. 6).

In model simulations, all three species were sensitive to changes in disturbance rate (fire frequency), but they differed in the degree to which their populations were

affected (Fig. 7). *Quercus faginea* and *Q. coccifera* were estimated to suffer the most from more frequent disturbance: compared to no change, a doubling of disturbance rate reduced their frequencies by over 70%, and a tripling of disturbance rate sent them both extinct (Fig. 7). In contrast, *Q. ilex* was predicted to be relatively robust: a doubling reduced frequency by only 17%, and a fivefold change by only 70% (Fig. 7). The simulated responses to changes in drought and fire frequency were affected by dispersal mode, with GR dispersal giving a more sensitive response to change than the other dispersal modes, although this effect was relatively small in magnitude (Fig. B2 in Appendix B).

Model simulations predicted varying sensitivity of the species to changes in habitat cover (Fig. 6 and Appendix B). Regardless of whether habitat loss was accompanied by fragmentation, *Quercus ilex* was predicted to be relatively insensitive to change, with *Q. coccifera* being very sensitive: under 5% habitat cover *Q. ilex* maintained a within-habitat frequency of at least 0.40, whereas the frequency of *Q. coccifera* declined to close to zero (Fig. 7). In contrast, the predicted sensitivity of *Q. faginea* depended critically on the nature of habitat loss (Fig. 7): where habitat loss was accompanied by fragmentation, the frequency declined to zero at 5% habitat cover (Fig. 7, middle panel), but with habitat loss restricted to edges, such that the remaining habitat was highly aggregated frequency showed no trend with habitat cover (Fig. 7, right panel).

Additional simulations revealed substantial interactions between the nature of habitat loss and dispersal mode (Appendix B). Where habitat loss was spatially random, all three species were highly sensitive to habitat loss under LR or GR dispersal; intermediate in sensitivity under LD dispersal; and highly robust under GD dispersal (Appendix B: Fig. B3 [top panels]). In contrast, where habitat loss was concentrated at edges,

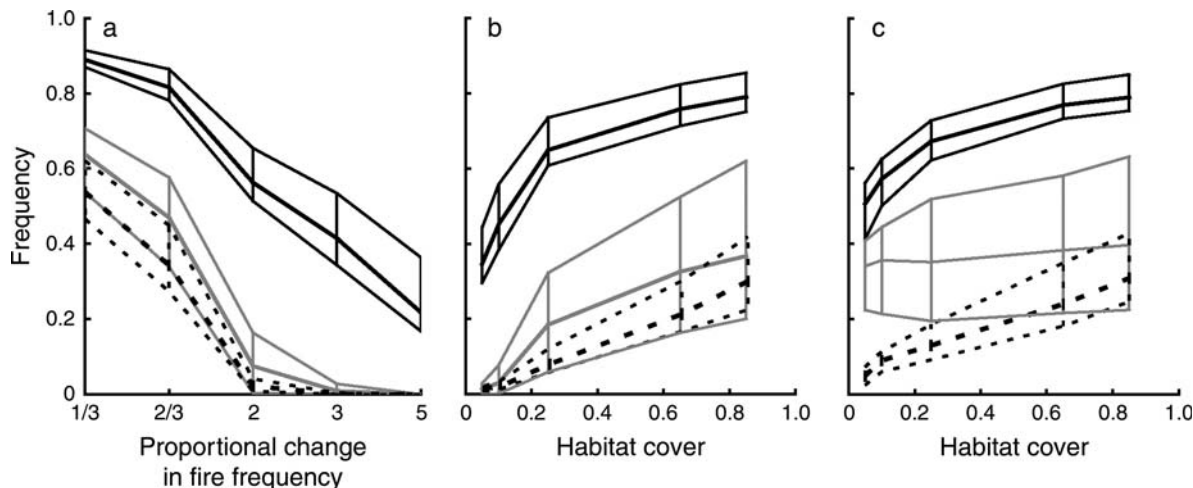


FIG. 7. Equilibrium average frequency of *Quercus faginea* (gray line), *Q. ilex* (solid black line), and *Q. coccifera* (dashed black line), given by simulations implemented with either altered disturbance (fire frequency) or habitat cover. All simulations used LD dispersal, with parameters estimated with respect to LD dispersal. Heavy lines are the mean over 10 simulations, and lighter lines give the range of results. Proportional change in fire frequency refers to change compared to current values, e.g., 2 means a doubling. To produce different habitat covers, losses and gains of habitat were implemented either (b) randomly or (c) concentrated at habitat edges. The observed habitat cover is 0.48. Frequency means frequency within suitable habitat. Analogous simulation results using different dispersal modes are given in Appendix B (Fig. B3).

all three species were highly robust under any dispersal mode except GR (Appendix B: Fig. B3 [bottom panels]).

#### DISCUSSION

##### *Oak woodland dynamics in central Spain*

An understanding of how abiotic controls on population processes translate into differences in community composition is essential to developing a quantitative understanding of vegetation dynamics in general, and the dynamics of Mediterranean woodlands in particular. Like much of the Mediterranean, the landscape of central Spain is exceptionally heterogeneous with respect to land cover (Fig. 1) and physical and climatic factors (see Table 1 and Rey 1999). By linking environmental heterogeneity to both seed establishment and variable disturbance, the SPOM was able to provide a population dynamic model that reproduced the observed variation in abundances of the focal oak species, *Quercus faginea*, *Q. ilex*, and *Q. coccifera* at several scales, including total regional abundances (Fig. 2); spatial structure within the region (Figs. 3–4); and small-scale variability associated with gradients in drought length, summer precipitation, temperature, and altitude (Fig. 4). The three oak species are dominant in this region and occur throughout much of the Mediterranean (Terradas and Savé 1992, Costa et al. 1998, Grove and Rackham 2001), and the SPOM presented here helps to understand the principal factors that regulate their distribution and abundance.

The SPOM parameter estimates suggest that these oaks differ in several important respects (Table 3), leading to contrasting abundances, spatial distributions, and responses to various perturbations (Figs. 2–7).

However, the results point clearly to a simple classification of the species that helps to explain most of the simulation results. Specifically, *Q. ilex* is highly fecund, but is relatively badly adapted to the current climate in terms of establishment, *Q. coccifera* has low fecundity but is very well adapted to the current climate, and *Q. faginea* has low fecundity and is relatively badly adapted to the current climate (Table 3). Thus, *Q. ilex* can be considered to be predominantly environment (or resource) limited, *Q. coccifera* dispersal limited, and *Q. faginea* limited by both. The simulation results agree with this observation. When comparing results for the three species, *Q. ilex* suffers least from local dispersal and gains least from directed dispersal (Fig. 2), *Q. ilex* is most robust to changes in habitat cover and disturbance rates (Figs. 6–7), and *Q. coccifera* is least affected by increasing drought length (Fig. 6). According to this classification, the dominance of *Q. ilex* in this highly fragmented, highly disturbed region is linked to its high fecundity, an explanation which is in agreement with recent empirical observations on *Q. ilex* seed production (Pulido and Díaz 2005).

It is not yet clear how to extrapolate these results to other regions, especially given the lack of data from other regions and the uncertainties inherent to the modeling approach. However, a simple extrapolation suggests some generalities about the factors that might determine the relative dominance of these species across the Mediterranean region, both at present, and in response to future perturbations. For example, the relative dominance of *Q. ilex* and its subspecies would be expected to be increased by habitat fragmentation or increases in disturbance rates, whereas increasing



PLATE 1. European Jay (*Garrulus glandarius*) swallowing acorns. Photo credit: Peter Preece.

drought length would be expected to favor *Q. coccifera*. A similar extrapolation would suggest that deciduous oaks (several of which occur in the Mediterranean region) would be favored in regions with a more mesic climate, lower disturbance rates, and less fragmented habitat. However, it is important to note that the analysis also points to the importance of other differences between the species, including contrasting correlations between altitude and seed establishment. Overall, the results suggest that species differences in the degree of environment-limitation vs. dispersal-limitation are critical to regional-scale vegetation dynamics, but additional empirical and theoretical studies are necessary to test the generality and validity of this prediction.

#### *Dispersal and environmental heterogeneity*

Dispersal processes can fundamentally alter species–environment interactions, thereby playing a vital role in landscape-level woodland dynamics. Simulations with the SPOM suggest that local directed dispersal most accurately captures the observed distribution of the three species (Table 2), and that this dispersal mode may be crucial in maintaining regional abundances of the species, especially *Q. faginea* and *Q. coccifera* (Fig. 2b). The European Jay is the primary agent that moves acorns sufficient distances from parent trees to areas suitable for seedling establishment (Gomez 2003). The jays bury acorns upon caching, which reduces the chance of seed predation and enhances germination rates and seedling survival (Gomez 2003). Therefore, although the simulation only altered the nature of dispersal (e.g., directed vs. random), in reality any reduction in the abundance of jays would also reduce the number of seeds exchanged between patches (i.e., reduce

inter-patch dispersal), increasing the deleterious effect of the reduction in the bird population. Together, field observations and model simulations point to the European Jay as a keystone species in oak woodlands (sensu Paine 1966). This in turn implies that management of these woodlands must consider the vegetation and the jays as a tightly coupled, and therefore potentially fragile, system.

The importance of local directed dispersal, reflecting animal-mediated seed movement, to regional vegetation dynamics is not unique to these three *Quercus* species or these woodlands. Many forests and woodlands around the world appear intimately tied to associated animal dispersers. For example, significant areas of the terrestrial biosphere are dominated by tree species that depend on corvid birds for long-distance dispersal (Powell and Zimmermann 2004). Over 20 *Pinus* species are associated with corvids, including Siberian, Japanese, and Korean stone pine communities, which stretch from the Ural to the Bering Sea; pinyon pine, which covers 75 000 square miles (46 500 km<sup>2</sup>) of the southwestern United States; and whitebark pine, which covers much of the U.S. Sierra Nevada and northern Rockies (Lanner 1996). Like the Spanish oak woodlands, all of these communities are characterized by stand-destroying fires, followed by bird-mediated recolonization, and the seed-caching behavior of the birds tends to be distinctly nonrandom (i.e., directed; Vander Wall and Balda 1977, Woolfenden and Fitzpatrick 1984, Marzluff and Balda 1992, Lanner 1996, Johnson et al. 1997, Gomez 2003). However, to our knowledge, this is the first time that these effects have been included in a population dynamic model to provide a quantitative estimate of their importance for the dependent tree species (but see



Purves and Dushoff 2005 for a related study of an aquatic perennial).

Theoretical explorations also suggest that details of dispersal that can have important effects on the abundance and distribution of plant species (reviewed by Levine and Murrell 2003). In this study, we used the SPOM to assess the interaction between local and directed dispersal in a highly heterogeneous environment. In many respects, the work here confirmed theoretical expectations. For example, in most simulations local dispersal (whether random or directed) decreased the predicted frequency of the species (Fig. 2b). This is expected because local dispersal preferentially delivers propagules to patches that are already occupied by the species where, by definition, the propagules do not lead to a new local population (Tilman et al. 1997). Local dispersal also amplified the species responses to environmental heterogeneity (Fig. 5), implying a narrowing of the realized environmental niche (Pacala and Hurtt 1993). Also, directed dispersal increased average frequency in all simulations (Fig. 2b; Appendix B: Figs. B2 and B3), which is expected in fragmented landscapes because dispersal prevents seeds from being lost to unsuitable habitat (Gomez 2003, Vander Wall and Balda 1997).

However, unlike previous theoretical work, the SPOM allowed an assessment of the interaction between local dispersal, directed dispersal, habitat fragmentation, and environmental heterogeneity (though, for related studies, see Hiebler 2000, King and With 2002, Purves and Dushoff 2005; and see Svenning 2001, Svenning and Skov 2002 for related empirical studies). For example, simulations suggested that either local dispersal, or directed dispersal, but not both, are needed to sustain viable regional populations of these species (Fig. 2a). This interaction is particularly prominent in these Spanish oak woodlands because of the fragmented but aggregated pattern of suitable habitat (Fig. 1). Thus, local dispersal tends to deliver seeds to suitable habitat even if dispersal is not actively directed toward suitable patches (Fig. 1b, c). This also explains why the predicted response of the species to changes in habitat cover depended on the combination of dispersal mode and nature of habitat loss. Either local or directed dispersal conferred robustness to edge-based habitat loss, because in this case the remaining habitat is aggregated, in which case local dispersal delivers seeds to suitable habitat. But under random habitat loss, only global, directed dispersal conferred robustness, because in this case random dispersal always wastes most seed to unsuitable habitat, and local dispersal does not move seeds over sufficient distances to reach from one patch of suitable habitat to the next. A similar effect is evident in the estimated loss of frequency due to local dispersal in the current landscape (Fig. 2b; LD minus GD dispersal). The reduction was smaller than might be expected, given that local dispersal results in nearly all seeds moving to closest-neighbor patches. This is explained by the

pattern of variation in the environmental conditions within suitable patches. This variation is spatially autocorrelated, such that local dispersal tends to deliver seeds to sites where local environmental conditions are favorable for seed establishment (Bolker 2003, Snyder and Chesson 2003). This positive effect of local dispersal compensates for the increased chance of landing in an occupied patch (Tilman et al. 1997), leaving a small effect of local vs. global dispersal.

#### *Model utility, limitation, and future directions*

The SPOM employed here was introduced to provide a population dynamic model that is capable of capturing species–environment correlations, which have traditionally been addressed using statistical approaches such as gradient analysis. We show how both environment-specific and demographic parameters can be estimated from survey data by coupling the SPOM to a Bayesian analysis (Tables 2 and 3). The end result is a fairly simple, yet semi-mechanistic parameterized population dynamic model that is capable of reproducing many observed features of the distributions of three co-occurring oak species (Figs. 2–4). The excellent fit to the survey data could not have been achieved by separate applications of gradient analyses (lack spatially dependent processes; see Fig. A1 in Appendix A), traditional SPOMs (lack environmental forcing), or more complex individual-based or ecosystem models (not readily parameterized from data). Importantly, the SPOM presented provides an initial framework for estimating the importance of different ecological processes in maintaining the current state of the vegetation (Figs. 2, 5). It also yields predictions of potential responses to perturbations in the biotic and abiotic environment (Figs. 6–7), highlighting directions for future modeling and field studies. Species responses to the environment result from complex interactions between a number of abiotic and biotic processes (Crawley 1997), but as we illustrate here, it is possible to develop and apply a simple modeling framework that captures part of this complexity.

We deliberately simplified many aspects of the model for a few key reasons. First, one of our goals was to illustrate the modeling approach by beginning with a fairly simple model that can be described by a few key processes. Second, although the survey plots provide an enormous amount of data, as do many inventory datasets, the data were limited in temporal scope (i.e., a snapshot) and the types of measurements represented (only presences–absence information). Third, we chose to maximize computational efficiency, and developed a simple model with tractable behavior that could be parameterized from survey data in a transparent manner. These simplifications call for caution in interpreting results, because alternative formulations may have given a similar fit to data (model uncertainty; Higgins et al. 2004). However, more biological detail can

be easily incorporated into the SPOM, providing a potentially more realistic modeling framework.

There are several ways that the SPOM could be modified to make it more realistic. First, and most important, we were forced to assume that current distributions are near equilibrium (this assumption is shared by most studies of species–environment relationships, including gradient analysis). But this may not be the case, because land use in this region has undergone significant changes over recent decades (Bramsnaes 1992, Blondel and Aronson 1995, Gomez-Limon and Fernandez 1999, Grove and Rackham 2001). However, if time-series data are available then this assumption can be relaxed. For example, Wu et al. (2002) introduced a method to parameterize a SPOM using information on historical changes in habitat cover, without assuming equilibrium. Although there is an increasing availability of data both on the historical distributions of plant species (e.g., Brewer et al. [2002] for European *Quercus*) and on historical land-use change (e.g., Ramankutty and Foley 1999, Hurtt et al. 2002, 2006), such data are not currently available for the woodlands in this study. Combining survey and historical data to estimate parameters for landscape models would greatly improve our ability to predict vegetation responses to future climate change.

Second, our SPOM does not include any interspecific interactions, which have the potential to modify species distributions (Pacala and Hurtt 1993, Leathwick and Austin 2001; see Zavala and Zea 2004 for Mediterranean pine–oak woodlands). These effects were not included here because the survey data were only available for a subset of the woody species in the landscape, so only a small subset of competitive interactions could have been studied. In addition, observations suggest that facilitation, as well as competition, is likely to be important in this region (Gomez-Aparicio et al. 2004), which leaves great uncertainty in how interspecific interactions should be incorporated. (For example, the realized niches of these species could be *larger* than their fundamental niches, due to the effects of facilitation.)

Competitive and facilitative interactions can be incorporated into the SPOM framework readily, for example, by making establishment and/or disturbance depend on the species that currently occupy a site (e.g., see Prakash and de Roos 2004). However, the result is a metacommunity model, the theoretical foundations and applicability of which have only just begun to be explored (see Leibold et al. 2004). In the oak woodland case, such a metacommunity model would most likely be more flexible than the SPOM we used, but it could not be properly constrained by the limited (snapshot, presence–absence, limited number of species) data. In situations where more complete data are available, parameterized metacommunity models could offer a powerful approach to understanding the interactions

between environmental forcing, dispersal processes, and interspecific interactions in complex landscapes.

Third, the approximation to the likelihood (Eqs. 9–13), which is also contingent upon the equilibrium assumption, is another potential source of error affecting the model parameterization and predictions. We could have treated  $S_{j,i}$  as a latent variable, thereby bypassing the need for the approximation in Eq. 13. The Bayesian framework is particularly well-suited for dealing with latent variables (Cappé and Robert 2000, Clark 2005), but in this case the approach would have added over 30 000 unknown quantities (i.e., an  $S_{j,i}$  for each species and site). The modeling approach we applied it is already computationally demanding, hence we choose to forego the latent variable option because it would have greatly increased the number of SPOM and M-H iterations required for obtaining posterior samples. In addition, the fit to observations presented here suggests that the approximation we used was capable of estimating appropriate parameter values (e.g., Fig. 4). However, implementation of a full Bayesian analysis that includes a hierarchical representation of data, latent variables, and parameters (e.g., Clark 2005) is clearly a method that deserves further attention for fitting SPOM models to field data.

Finally, it is important to bear in mind the key differences between the SPOM approach, gradient analysis, and individual-based modeling. Unlike gradient modeling the SPOM offers a model grounded in population dynamic processes, including dispersal, with an explicit timescale. It can therefore address biological questions such as the effects of demographic rates (local extinction, colonization, dispersal [e.g., Fig. 2]) and applied questions such as changes in habitat cover, disturbance, and climatic conditions (e.g., Fig. 7). But in the form presented here, the SPOM approach shares some important limitations with gradient analysis, including the equilibrium assumption discussed above. Unlike individual-based models, SPOMs can be implemented efficiently at large spatial scales, and can be parameterized readily from survey data; but individual-based models offer a level of biological realism that cannot be approached by models that reduce the state of a plant community to the presence or absence of the component species. Because of these different benefits and limitations, we anticipate that a variety of modeling approaches will continue to be important in studying regional vegetation dynamics in the future. However, unlike the other approaches, the SPOM approach had not been applied to studying large-scale vegetation dynamics previously. The novel modeling approach presented has the potential to improve our heuristic and predictive understanding of vegetation dynamics.

#### ACKNOWLEDGMENTS

We thank J. A. Villanueva and DGCONA (Ministerio de Medio Ambiente) for access to the forest inventory databases; Jose Manuel Moreira and Javier Quijada for allowing us access to data on fire frequency from Red de Información Ambiental

de Andalucía (REDIAM); and Itziar R. Urbieto for assistance with forest inventory data processing. We thank Jonathan Dushoff, Pedro Jordano, Jose M. Gómez, and Christian Wirth for useful discussions. Rampal Etienne and two anonymous reviewers provided extremely helpful comments on a previous version of this manuscript. This work was supported by the Andrew Mellon Foundation (D. W. Purves) and is a product of the Carbon Mitigation Initiative (CMI) of the Princeton Environmental Institute at Princeton University. CMI is sponsored by BP and Ford. The work was also supported by projects REN2000-745 and REN2002-04041-C02-02/GLO (Spanish Ministry of Science and Technology). K. Ogle acknowledges support by the National Science Foundation under a grant awarded in 2003. We acknowledge the networks REDBOME (Junta de Andalucía) and GLOBIMED (Spanish Ministry of Science and Technology) for hosting meetings and fostering scientific interactions.

## LITERATURE CITED

- Archibold, O. W. 1995. Ecology of world vegetation. Chapman and Hall, London, UK.
- Arianoutsou, M., and V. Papanastasis. 2004. Ecology, conservation and sustainable management of mediterranean type ecosystems of the world. Millpress, Rotterdam, The Netherlands.
- Austin, M. P., A. O. Nicholls, and C. R. Margules. 1990. Measurement of the realized qualitative niche: environmental niches of five *Eucalyptus* species. *Ecological Monographs* 60: 161–177.
- Blondel, J., and J. Aronson. 1995. Biodiversity and ecosystem function in the Mediterranean basin: human and non-human determinants. Pages 43–119 in G. W. Davis and D. M. Richardson, editors. *Mediterranean-type ecosystems: the function of biodiversity*. Springer-Verlag, Berlin, Germany.
- Bolker, B. 2003. Combining endogenous and exogenous spatial variability in analytical population models. *Theoretical Population Biology* 64:255–270.
- Bolker, B., S. W. Pacala, and S. A. Levin. 2000. Moment methods for ecological processes in continuous space. Pages 388–410 in U. Dieckmann, R. Law, and J. A. J. Metz, editors. *The geometry of ecological interactions: simplifying spatial complexity*. Cambridge University Press, New York, New York, USA.
- Bossema, I. 1979. Jays and oaks: an eco-ethological study of a symbiosis. *Behavior* 70:1–117.
- Bramsnaes, A. 1992. The landscape of agriculture: planning and regulation of agricultural landscapes in European Community. *Landscape and Urban Planning* 22:17–30.
- Brewer, S., R. Cheddadi, J. L. de Beaulieu, and M. Reille. 2002. The spread of deciduous *Quercus* throughout Europe since the last glacial period. *Forest Ecology and Management* 156: 27–48.
- Cappé, O., and C. P. Robert. 2000. Markov chain Monte Carlo: 10 years and still running! *Journal of the American Statistical Association* 95:1282–1286.
- Chib, S., and E. Greenberg. 1995. Understanding the Metropolis-Hasting algorithm. *American Statistician* 49:327–335.
- Clark, J. S. 2005. Why environmental scientists are becoming Bayesians. *Ecology Letters* 8:2–14.
- Condit, R., et al. 2000. Spatial patterns in the distribution of tropical tree species. *Science* 288:1414–1418.
- Corcuera, L., J. J. Camarero, and E. Gil-Pelegrin. 2002. Functional groups in *Quercus* species derived from the analysis of pressure-volume curves. *Trees: structure and function* 16:465–472.
- Costa, M., C. Morla, and H. Sainz. 1998. Los Bosques de la Península Ibérica: una interpretación geobotánica. Geoplaneta, Barcelona, Spain.
- Cowles, H. C. 1899. The ecological relations of the vegetation on the sand dunes of Lake Michigan, part I, geographical relations of the dune floras. *Botanical Gazette* 27:95–117.
- Crawley, M. J. 1997. Plant–herbivore dynamics. Pages 401–474 in M. J. Crawley, editor. *Plant ecology*. Blackwell Science, Oxford, UK.
- Davies, K. F., C. Gascon, and C. Margules. 2001. Habitat fragmentation: consequences, management, and future research priorities. Pages 81–98 in M. E. Soule and G. H. Orians, editors. *Conservation biology: research priorities for the next decade*. Island Press, New York, New York, USA.
- Davis, A. J., L. S. Jenkinson, J. L. Lawton, B. Shorrocks, and S. Wood. 1998a. Making mistakes when predicting shifts in species range in response to global warming. *Nature* 391:783–786.
- Davis, A. J., J. L. Lawton, B. Shorrocks, and L. S. Jenkinson. 1998b. Individualistic species responses invalidate simple physiological models of community dynamics under global environmental change. *Journal of Animal Ecology* 67:600–612.
- de Miguel, J. M. 1999. The nature of the agricultural-forestry-pasture landscape in the conservation of biological diversity in Spain. *Revista Chilena de Historia Natural* 72:547–557.
- Desanker, P. V. 1996. Development of a MIOMBO woodland dynamics model in Zambesian Africa using Malawi as a case study. *Climatic Change* 34:279–288.
- Etienne, R. S. 2000. Local populations of different sizes, mechanistic rescue effect and patch preference in the Levins metapopulation model. *Bulletin of Mathematical Biology* 62: 943–958.
- Etienne, R. S., C. J. F. ter Braak, and C. C. Vos. 2004. Application of stochastic patch occupancy model to real metapopulations. Pages 105–132 in I. Hanski and O. E. Gaggiotti, editors. *Ecology, genetics, and evolution of metapopulations*. Elsevier, Burlington, Massachusetts, USA.
- Fotelli, M. N., K. M. Radoglou, and H. I. A. Constantinidou. 2000. Water stress responses of seedlings of four Mediterranean oak species. *Tree Physiology* 20:1065–1075.
- Franklin, J. 1998. Predicting the distribution of shrub species in southern California from climate and terrain-derived variables. *Journal of Vegetation Science* 9:733–748.
- Freckleton, R. P., and A. R. Watkinson. 2002. Large-scale spatial dynamics of plants: metapopulations, regional ensembles and patchy populations. *Journal of Ecology* 90:419–434.
- Gavilán, R., and F. Fernández-González. 1997. Climatic discrimination of Mediterranean broad-leaved sclerophyllous and deciduous forests in central Spain. *Journal of Vegetation Science* 8:377–386.
- Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 2004. *Bayesian data analysis*. Chapman and Hall/CRC, Boca Raton, Florida, USA.
- Gomez, J. M. 2003. Spatial patterns in long-distance dispersal of *Quercus* acorns by jays in a heterogeneous landscape. *Ecography* 26:573–584.
- Gomez-Aparicio, L., R. Zamora, J. M. Gomez, J. A. Hodar, J. Castro, and E. Baraza. 2004. Applying plant facilitation to forest restoration: a meta-analysis of the use of shrubs as nurse plants. *Ecological Applications* 14:1128–1138.
- Gomez-Limón, J., and J. V. D. Fernandez. 1999. Changes in use and landscape preferences on the agricultural-livestock landscapes of the central Iberian Peninsula (Madrid, Spain). *Landscape and Urban Planning* 44:165–175.
- Grove, A. T., and O. Rackham. 2001. *The nature of mediterranean Europe: an ecological history*. Yale University Press, New Haven, Connecticut, USA.
- Hanski, I. A. 1994. A practical model of metapopulation dynamics. *Journal of Animal Ecology* 63:151–162.
- Hanski, I. A. 1997. Metapopulation dynamics: from concepts and observations to predictive models. Pages 69–122 in I. A. Hanski and M. E. Gilpin, editors. *Metapopulation biology*. Academic Press, San Diego, California, USA.
- Hanski, I. A., and O. E. Gaggiotti, editors. 2004. *Ecology, genetics, and evolution of metapopulations*. Elsevier, Burlington, Massachusetts, USA.

- Hanski, I. A., and M. E. Gilpin. 1997. *Metapopulation biology*. Academic Press, San Diego, California, USA.
- Hanski, I. A., and D. Simberloff. 1997. The metapopulation approach, its history, conceptual domain, and application to conservation. Pages 5–26 in I. A. Hanski and M. E. Gilpin, editors. *Metapopulation biology*. Academic Press, San Diego, California, USA.
- Heegaard, E. 2002. A model of alpine species distribution in relation to snowmelt time and altitude. *Journal of Vegetation Science* 13:493–504.
- Hiebeler, D. 2000. Populations on fragmented landscapes with spatially structured heterogeneities: landscape generation and local dispersal. *Ecology* 81:1629–1641.
- Higgins, S. I., J. S. Clark, R. Nathan, T. Hovestadt, F. Schurr, J. M. V. Fragoso, M. R. Aguiar, E. Ribbens, and S. Lavorel. 2004. Forecasting plant migration rates: managing uncertainty for risk assessment. *Journal of Ecology* 91:341–347.
- Hilborn, R., and M. Mangel. 1997. *The ecological detective: confronting models with data*. Princeton University Press, Princeton, New Jersey, USA.
- Hurtt, G. C., S. Frolking, M. G. Fearon, B. Moore III, E. Shevliakova, S. Malyshev, S. W. Pacala, and R. A. Houghton. 2006. The underpinnings of land-use history: three centuries of global gridded land-use transitions, wood harvest activity, and resulting secondary lands. *Global Change Biology* 12:1208–1229.
- Hurtt, G. C., S. W. Pacala, P. R. Moorcroft, J. Caspersen, E. Shevliakova, R. A. Houghton, and B. Moore. 2002. Projecting the future of the US carbon sink. *Proceedings of the National Academy of Sciences (USA)* 99:1389–1394.
- Intergovernmental Panel on Climate Change. 2001. *Climate change 2001: impacts, adaptation and vulnerability*. Cambridge University Press, Cambridge, UK.
- Inventario Forestal Nacional. 1995. Segundo inventario forestal nacional. Ministerio de Agricultura, Pesca y Alimentación, Madrid, Spain.
- Iverson, L. R., and A. M. Prasad. 1998. Predicting abundance of 80 tree species following climate change in the eastern United States. *Ecological Monographs* 68:465–485.
- Iwasa, Y. 2000. Lattice models and pair approximation in ecology. Pages 227–251 in U. Dieckmann, R. Law, and J. A. J. Metz, editors. *The geometry of ecological interactions: simplifying spatial complexity*. Cambridge University Press, New York, New York, USA.
- Johnson, W. C., C. S. Adkisson, T. R. Crow, and M. D. Dixon. 1997. Nut caching by Blue Jays (*Cyanocitta cristata* L.): implications for tree demography. *American Midland Naturalist* 138:357–370.
- King, A. W., and K. A. With. 2002. Dispersal success on spatially structured landscapes: when do spatial pattern and dispersal behavior really matter? *Ecological Modelling* 147: 23–39.
- Lanner, R. M. 1996. *Made for each other: a symbiosis of birds and pines*. Oxford University Press, Oxford, UK.
- Law, R., and U. Dieckmann. 2000. Moment approximations of individual-based models. Pages 252–270 in U. Dieckmann, R. Law, and J. A. J. Metz, editors. *The geometry of ecological interactions: simplifying spatial complexity*. Cambridge University Press, New York, New York, USA.
- Law, R., D. J. Murrell, and U. Dieckmann. 2003. On population growth in space and time: spatial logistic equations. *Ecology* 84:252–262.
- Lawton, J. L. 2000. Concluding remarks: a review of some open questions. Pages 401–424 in M. J. Hutchings, E. John, and A. J. A. Stewart, editors. *Ecological consequences of environmental heterogeneity*. Cambridge University Press, Cambridge, UK.
- Leathwick, J. R., and M. P. Austin. 2001. Competitive interactions between tree species in New Zealand's indigenous tree species. *Functional Ecology* 15:233–242.
- Leathwick, J. R., and D. Whitehead. 2001. Soil and atmospheric water deficits and the distribution of New Zealand's indigenous tree species. *Functional Ecology* 15:233–242.
- Leibold, M. A., M. Holyoak, N. Mouquet, and others. 2004. The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters* 7:601–613.
- Levine, J. M., and D. J. Murrell. 2003. The community-level consequences of seed dispersal patterns. *Annual Reviews of Ecology and Systematics* 34:549–574.
- Levins, R. 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological Society of America* 12:237–240.
- Lopez, J. E., and C. A. Pfister. 2001. Local population dynamics in metapopulation models: Implications for conservation. *Conservation Biology* 15:1700–1709.
- López-Soria, L., and C. Castell. 1992. Comparative genet survival after fire in woody Mediterranean species. *Oecologia* 53:493–499.
- Marzluff, J. M., and R. P. Balda. 1992. *The Pinyon Jay*. T. and A. D. Poyser, London, UK.
- Moorcroft, P. R., G. C. Hurtt, and S. W. Pacala. 2001. A method for scaling vegetation dynamics: the ecosystem demography model (ED). *Ecological Monographs* 71:557–585.
- Naveh, Z. 1990. Fire in the Mediterranean: a landscape ecological perspective. Pages 1–20 in J. G. Goldammer and M. J. Jenkins, editors. *Fire ecosystem dynamics*. SPB Academic Publishing, The Hague, Germany.
- O'Hara, R. B., E. Arjas, H. Toivonen, and I. Hanski. 2002. Bayesian analysis of metapopulation data. *Ecology* 83:2408–2415.
- Pacala, S. W., C. D. Canham, J. Saponara, J. A. Silander, R. K. Kobe, and E. Ribbens. 1996. Forest models defined by field measurements: estimation, error analysis and dynamics. *Ecological Monographs* 66:1–43.
- Pacala, S. W., and G. C. Hurtt. 1993. Terrestrial vegetation and climate change: integrating models and experiments. Pages 57–74 in P. Kareiva and J. Kingsolver, editors. *Biotic interactions and global change*. Sinauer Associates, Sunderland, Massachusetts, USA.
- Paine, R. T. 1966. Food web complexity and species diversity. *American Naturalist* 100:65–75.
- Pearson, R. G., and T. P. Dawson. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography* 12:361–371.
- Powell, J. A., and N. E. Zimmermann. 2004. Multiscale analysis of active seed dispersal contributes to resolving Reid's paradox. *Ecology* 85:490–506.
- Prakash, S., and A. M. de Roos. 2004. Habitat destruction in mutualistic metacommunities. *Theoretical Population Biology* 65:153–163.
- Primack, R. B. 1998. *Essentials of conservation biology*. Sinauer Associates, Sunderland, Massachusetts, USA.
- Pulido, F. J., and M. Diaz. 2005. Regeneration of a Mediterranean oak: a whole-cycle approach. *EcoScience* 12: 92–102.
- Purves, D. W., and J. Dushoff. 2005. Directed seed dispersal and metapopulation response to habitat loss and disturbance: application to *Eichhornia paniculata*. *Journal of Ecology* 93: 658–669.
- Purves, D. W., and R. Law. 2002. Fine-scale spatial structure in a grassland community: quantifying the plant's-eye view. *Journal of Ecology* 90:121–129.
- Ramankutty, N., and J. A. Foley. 1999. Estimating historical changes in land cover: North American croplands from 1850 to 1992. *Global Ecology and Biogeography* 8:381–396.
- Retana, J., J. M. Espelta, M. Gracia, and M. Riba. 1999. Seedling recruitment. Pages 89–103 in F. Rodà, J. Retana, C. Gracia, and J. Bellot, editors. *Ecology of Mediterranean*

- evergreen oak forests. Springer-Verlag, New York, New York, USA.
- Rey, J. M. 1999. Modelling potential evapotranspiration of potential vegetation. *Ecological Modelling* 123:141–159.
- Rey Benayas, J. M. 1998. Growth and survival in *Quercus ilex* L. seedlings after irrigation and artificial shading on Mediterranean set-aside agricultural lands. *Annals of Forest Science* 55:801–807.
- Rey Benayas, J. M., J. Navarro, T. Espigares, J. Nicolau, and M. A. Zavala. 2005. Limiting effects of high radiation and weed competition on reforestation of Mediterranean abandoned cropland with contrasting *Quercus* species. *Forest Ecology and Management* 212:302–314.
- Ripley, B. D. 1981. *Spatial statistics*. John Wiley and Sons, New York, New York, USA.
- Robert, C. P., and G. Casella. 1999. *Monte Carlo statistical methods*. Springer-Verlag, New York, New York, USA.
- Snyder, R. E., and P. Chesson. 2003. Local dispersal can facilitate coexistence in the presence of permanent spatial heterogeneity. *Ecology Letters* 6:301–309.
- Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. van der Linde. 2002. Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society Series B* 64:583–640.
- Svenning, J. C. 2001. Environmental heterogeneity, recruitment limitation and the mesoscale distribution of palms in a tropical montane rain forest (Maquipucuna, Ecuador). *Journal of Tropical Ecology* 17:97–113.
- Svenning, J. C., and F. Skov. 2002. Mesoscale distribution of understorey plants in temperate forest (Kalo, Denmark): the importance of environment and dispersal. *Plant Ecology* 160: 169–185.
- ter Braak, C. J. F., and R. S. Etienne. 2003. Improved Bayesian analysis of metapopulation data with an application to a tree frog metapopulation. *Ecology* 84:231–241.
- Terradas, J., and R. Savé. 1992. The influence of summer and winter stress and water relationships on the distribution of *Quercus ilex* L. *Vegetatio* 99–100:137–145.
- Tilman, D., C. L. Lehman, and P. Kareiva. 1997. Population dynamics in spatial habitats. Pages 3–20 in D. Tilman and P. Kareiva, editors. *Spatial ecology: the role of space in population dynamics and interspecific interactions*. Princeton University Press, Princeton, New Jersey, USA.
- Valladares, L., E. Balaguer, E. Martínez-Ferri, E. Pérez-Corona, and E. Manrique. 2002. Plasticity, instability and canalization: is the phenotypic variation in seedlings of sclerophyll oaks consistent with the environmental unpredictability of Mediterranean ecosystems? *New Phytologist* 156:457–467.
- Vander Wall, S. B., and R. P. Balda. 1977. Co-adaptations of the Clark's Nutcracker and the Piñon pine for efficient seed harvest and dispersal. *Ecological Monographs* 47:89–111.
- Vázquez, A., B. Pérez, F. Fernández-González, and J. M. Moreno. 2002. Recent fire regime characteristics and potential natural vegetation relationships in Spain. *Journal of Vegetation Science* 13:663–676.
- Villanueva, J. A. 1993. El segundo inventario forestal nacional de España. *Revista Montes* 32:52–62.
- Villar-Salvador, P., P. Castro-Diez, C. Perez-Rontome, and G. Montserrat-Martí. 1997. Stem xylem features in three *Quercus* (*Fagaceae*) species along a climatic gradient in NE Spain. *Trees: Structure and Function* 12:90–96.
- Vos, C. C., C. J. F. ter Braak, and W. Nieuwenhuizen. 2000. Incidence function modelling and conservation of the tree frog *Hyla aborea* in the Netherlands. *Ecological Bulletin* 48: 165–180.
- Walter, H. 1973. *Vegetation of the earth in relation to climate and eco-physiological conditions*. [J. Wieser, translator.] Springer-Verlag, New York, New York, USA.
- Walter, H. 2002. *Walter's vegetation of the earth: the ecological systems of the geo-biosphere*. [G. Lawlor and D. Lawlor, translators.] Springer, Berlin, Germany.
- Whittaker, R. H. 1956. Vegetation of the great smoky mountains. *Ecological Monographs* 26:1–80.
- Whittaker, R. H. 1960. Vegetation of Siskiyou Mountains, Oregon and Washington. *Ecological Monographs* 30:279–338.
- Wolfenden, G. E., and J. W. Fitzpatrick. 1984. *The Florida Scrub-Jay: demography of a cooperative-breeding bird*. Princeton University Press, Princeton, New Jersey, USA.
- Wu, W. D., R. Heikkilä, and I. Hanski. 2002. Estimating the consequences of habitat fragmentation on extinction risk in dynamic landscapes. *Landscape Ecology* 17:699–710.
- Zavala, M. A., and G. E. Zea. 2004. Mechanisms maintaining biodiversity in Mediterranean pine-oak forests: insights from a spatial simulation model. *Plant Ecology* 171:197–207.

#### APPENDIX A

Summary of gradient analysis for three *Quercus* species in central Spain (*Ecological Archives* M077-004-A1).

#### APPENDIX B

Perturbation experiments (*Ecological Archives* M077-004-A2).

#### APPENDIX C

Statistical model of fire frequency in Mediterranean Spain (*Ecological Archives* M077-004-A3).