Forest recovery through applied nucleation: effects of tree islet size and disperser mobility on tree recruitment in a temperate landscape

Forest Ecology and Management 550 (2023) 121508

https://doi.org/10.1016/j.foreco.2023.121508

Electronic Supplementary Material

Teresa Morán-López [1,2*], Javier Rodríguez-Pérez [3], Isabel Donoso [4], Daniel Martínez[5], Juan Manuel Morales[2,6] and Daniel García[1]

1. Depto. de Biología de Organismos y Sistemas (Universidad de Oviedo, Oviedo), and Instituto Mixto de Investigación en Biodiversidad (Universidad de Oviedo-CSIC-Principado de Asturias, Mieres), Asturias, Spain.
2. Grupo de Ecología Cuantitativa, INIBIOMA-CONICET, Univ. Nacional del Comahue. Bariloche, Río Negro, Argentina.
3. Depto. Ciencias del Medio Natural, Centro Jerónimo de Ayanz, Universidad Pública de Navarra (UPNA), Pamplona, Navarra, Spain.
4. Instituto Mediterráneo de Estudios Avanzados (IMEDA, CSIC, UIB), Grupo de Oceanografía y Cambio Global, Esporles, Islas Baleares, Spain.
5. Servicios de Gestión Medioambiental Sigma, Oviedo, Asturias, Spain.
6. School of Biodiversity, One Health & Veterinary Medicine, University of Glasgow, Glasgow, Scotland, UK.

* Corresponding author (moranteresa@uniovi.es)

Appendix A: Simulation of landscapes, seed dispersal and seedling recruitment

A.1. Landscape properties: tree cover, composition of fruiting species and fruit production

In our simulations, matrix cells transformed into forest (restored cells) mimicked the composition of the secondary forest in the study area. Thus, we needed information about (i) tree cover in forest cells (proportion) (ii) total cover of fleshy-fruited species (m^2) and (iii) species-specific composition (m^2). To attain this information, we first classified landscape cells into *forest* and *matrix* habitat according to their tree cover. To establish a threshold value of such classification, we quantified Intraclass Correlation Coefficient (ICC) of groups (i.e., forest *vs* matrix habitats) according to increasing values of tree cover (from 0 to 0.5, proportion). ICC measures how data-structured groups resemble each other (Koch et al. 1982) and values higher than 0.9 indicate a high consistency within groups. We classified cells into forest habitats when tree cover was equal or higher than 0.2 (when ICC reaches values of 0.9, Fig. A1). Such classification visually resembled the forest patches found in the study site (Fig. A2).

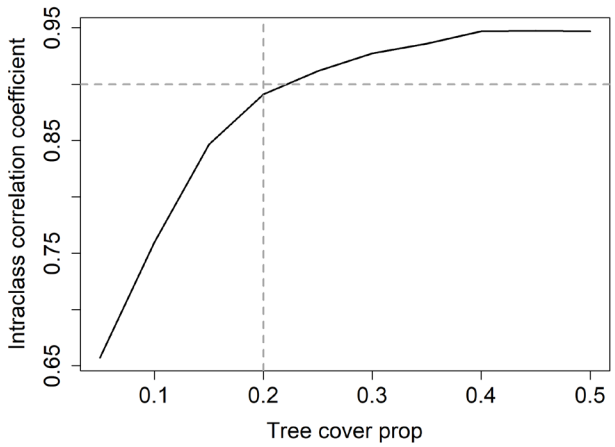


Fig. A1. Intraclass correlation coefficient (ICC) of cells into categories of habitat (forest *vs* matrix) at increasing levels of tree cover as a threshold in the classification. Grey dotted lines depict the threshold value used in our classification and its ICC estimation.

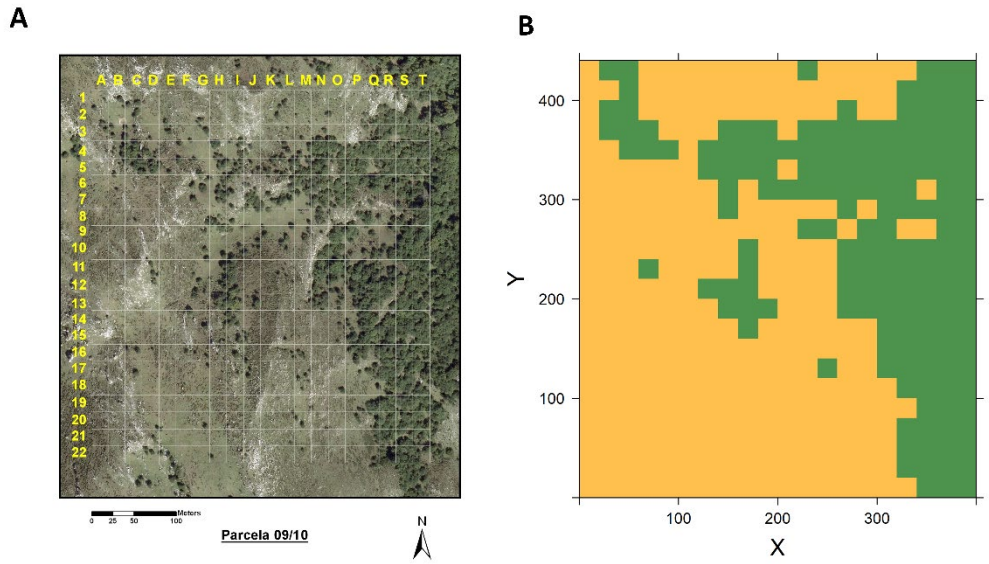


Fig. A2. (A) Aerial photograph of the study plot and (B) map of cell classification with a threshold value of 0.2 of cover for forest cells (green cells classified as forest, orange cells classified as matrix).

In our simulations, forest restored cells were characterized by a certain value of tree cover (m^2) and a proportion of such cover that corresponded to fleshy-fruited tree species. These values were sampled from density distributions fitted to field data (Fig. A3).

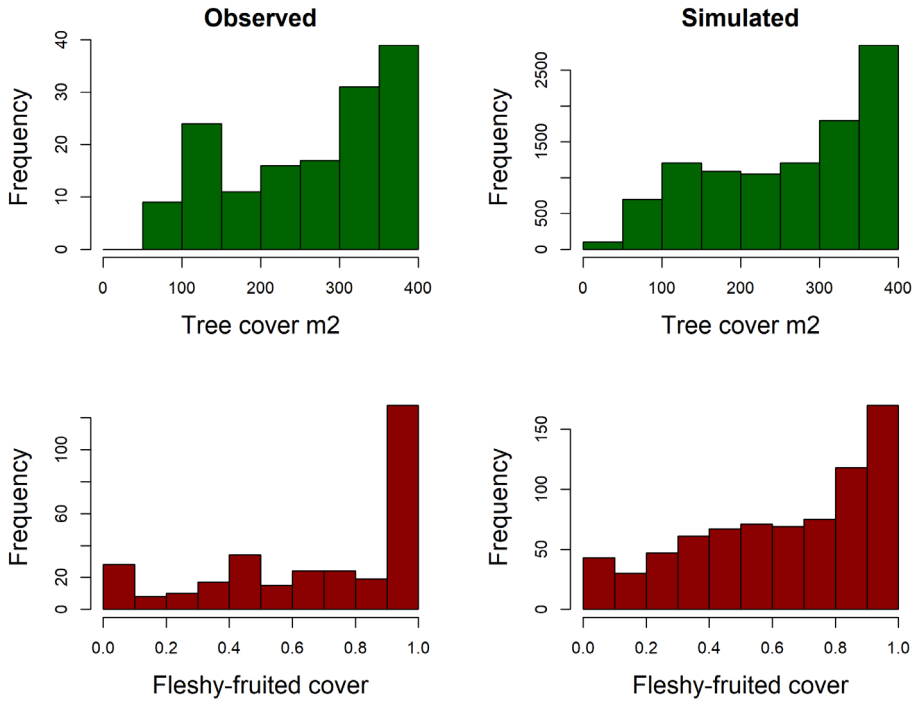
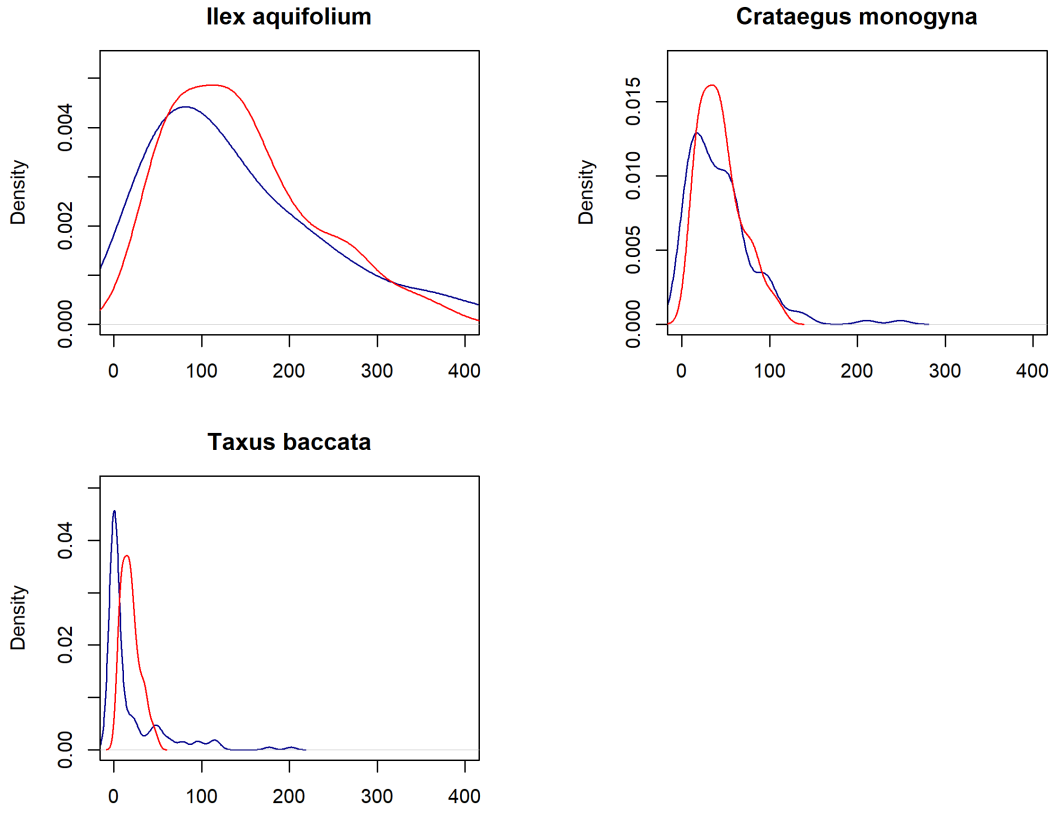


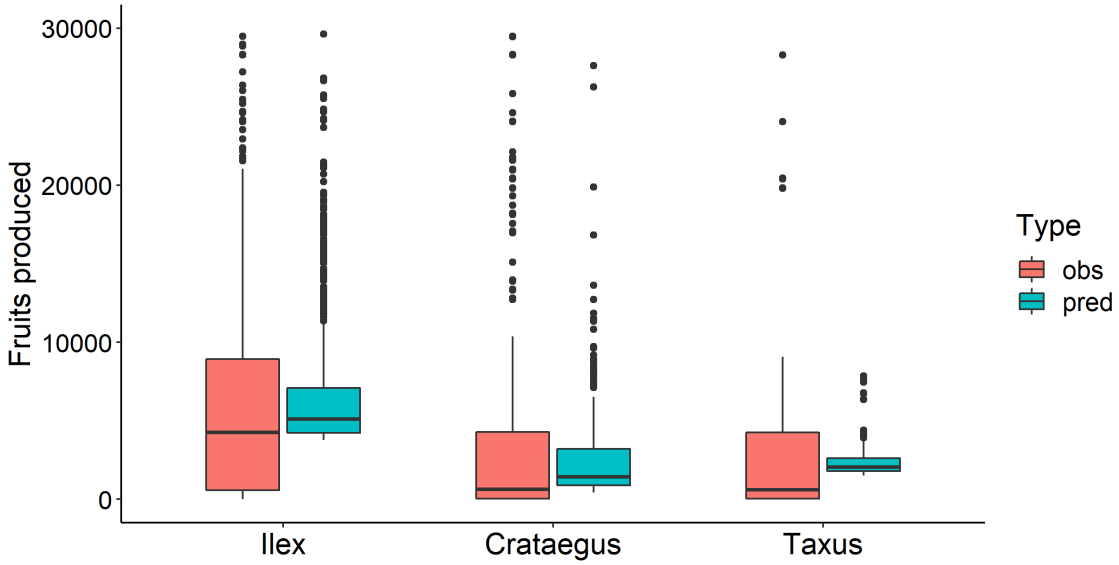
Fig. A3. Histograms of values of tree cover (m^2) in forest cells (upper panels) and the proportion of cover corresponding to fleshy-fruited species (lower panels). Left panels correspond to observed values and right panels correspond to simulated values sampled from density distributions fitted to field data.

To assign a composition of fleshy-fruited species to restored cells we fitted a multinomial regression to field data, which assigned a number of m^2 of cover of fleshy-fruited species to each species (*Ilex aquifolium, Crataegus monogyna* and *Taxus baccata*). We used a Bayesian approach with the Stan program (Stan development team). After checking for model convergence (Rhat < 1.1) and effective sample sizes (>1000), we performed a posterior predictive check to evaluate model fit (Fig. A4). Estimated proportion of cover of fruiting species per cell were 0.69 (credibility interval= [0.68, 0.70]), 0.21 (CI= [0.2, 0.22]) and 0.09 (CI= [0.09, 0.1]) for *Ilex aquifolium, Crataegus monogyna* and *Taxus baccata,* respectively. We multiplied these values by cell size (400 m^2) to attain species-specific covers (m^2).



Figs. A4. Posterior predictive check of multinomial model (response variable m^2 of cover per cell). Blue lines correspond to observed values of fleshy fruited species cover in forest cells, red lines simulated values by sampling the posterior distributions of species-specific probabilities of the multinomial model.

To simulate fruit production across the landscape (i.e., in forest, matrix and restored cells), for each fruiting species we fitted a zero-inflated regression model using data of fruit counts per cell of the study area for 5 years (2007-2011). In all species, the probability of fruit production was a parameter to estimate and represented the probability that tree cover corresponded to male (e.g., *Taxus baccata* and *Ilex aquifolium*) and not fruiting individuals (e.g., *Crataegus monogyna*). In all cases, fruit production of cells depended on the tree cover of each fruiting species. We included year as a random term in the intercept to account for interannual variability in fruit production. Models were fitted with weakly informative priors: the probability of production was sampled from a uniform distribution ranging from 0 to 1; and the intercept and tree cover effects from a normal distribution with mean 0 and standard deviation of 1. We evaluated chain convergence (Rhat<1.1) and number of effective size (Neff>1000). Subsequently, we performed posterior predictive check to assess model fit. Our model adjusted well fruit production for *Crataegus monogyna* and *Ilex aquifolium* but overestimated fruit production of *Taxus baccata* (Fig. A5). Nonetheless, *Taxus baccata* represented a small fraction of overall fruit production (6.4% on average across years). Thus, such overrepresentation would not affect movement or foraging choices of simulated birds which depended on overall and relative fruit abundances (see below).



Figs. A5. Posterior predictive check of crop sizes (fruit production > 0) observed in data and predicted by our zero-inflated regressions.

A.2. Seed dispersal as a consequence of bird movement and foraging choices

In our simulations birds move and forage following the same behavioral rules to those of a previous model fitted and validated with real data in the study site (Morales et al., 2013). Simulated birds move throughout the landscape balancing the costs and benefits of reaching a new destination cell, as following:

$$logit(v_{ik}) = a_{0k} + b_{0k}B_i \qquad \text{eq.A1}$$

$$D_{ik} = 1 - tanh\left(\left(d_{ij}/a_{dk}\right)^{b_{dk}}\right) \qquad \text{eq.A2}$$

$$C_{ik} = tanh((c_i/a_{ck})^{b_{ck}}) \qquad \text{eq.A3}$$

$$F_{ik} = tanh\left(\left(f_i/a_{fk}\right)^{b_{fk}}\right) \qquad \text{eq.A4}$$

$$\beta_{ik} = D_{ik} * C_{ik} * F_{ik} \Big/ \sum_{n=1}^{N} D_{ik}C_{ik}F_{ik} \qquad \text{eq.A5}$$

Before each movement, birds decide whether to leave the study area or not depending on the distance from the current location to the edge of the plot (B_i) and species-specific parameters that modulate such probabilities (a_{0k} and b_{0k}) (Fig. A6a). If birds stay within the plot, they decide where to move next based on eq. A2-5. D_{ik} represents the attractiveness of the i-th cell for a bird of the k-th bird species based on its distance from its current location and the species-specific scale of movement (based on a_{dk} and b_{dk}) (Fig. A6b). C_{ik} and F_{ik} depict the attractiveness of the i-th cell to a bird of the k-th species based on its tree cover (c_i) and fruit availability (f_i) as well as the preference of the k-th species towards covered and fruit-rich areas (based on a_{ck} and b_{ck}; a_{fk} and b_{fk}, respectively; Fig. A6c-d). Finally, a probability distribution is built by multiplying all biases of movement (eq. A2-4) and dividing by the sum of all cells so that β_{ik} sums 1 across all options (i.e., cells within the landscape). In Morales et al. (2013) species-specific parameters of movement were fitted by means of maximum likelihood estimation using observed trajectories of each thrush species in the study plot for 2007, 2008 and 2009 (238 accumulated hours of observation). Models with different combinations of factors (distance, cover and fruit availability) were fitted and the one with the lowest AIC value retained. In the case of *Turdus iliacus, T. merula, T. viscivorus* and *T. philomelos* all factors modulated their movement. In contrast, fruit availability was unimportant for *T. pilaris* and tree cover within cells for *T. torquatus*.

Once a bird reaches the target cell, the probability of landing on open areas (l_{ik}) depends on the amount of tree cover of the destination cell (C_i) and species-specific preferences towards landing on tree cover (eq. A6. $logit(l_{ik}) = a_{lk} + b_{lk}C_i$; Fig. A6f). For each bird species, we fitted a Bernoulli regression model in which the probability of landing on open microhabitats depended on the amount of tree cover available in the target cell. To this end, we used a total of 1645 observations (from 2008 to 2011) of the stratum of landing (tree canopy *vs* open areas) of thrushes. We decided to include this behavior because we were interested in modeling post-dispersal stages, where the microhabitat of deposition can be a strong modulator of seed predation and seedling survival (Donoso et al., 2016). We modeled Bernoulli regressions using a Bayesian approach with JAGS program (Plummer, 2003). After checking for chain convergence (Rhat<1.1) and sufficient effective sample size of model parameters (>1000) and performed posterior predictive checks (Fig.A7).

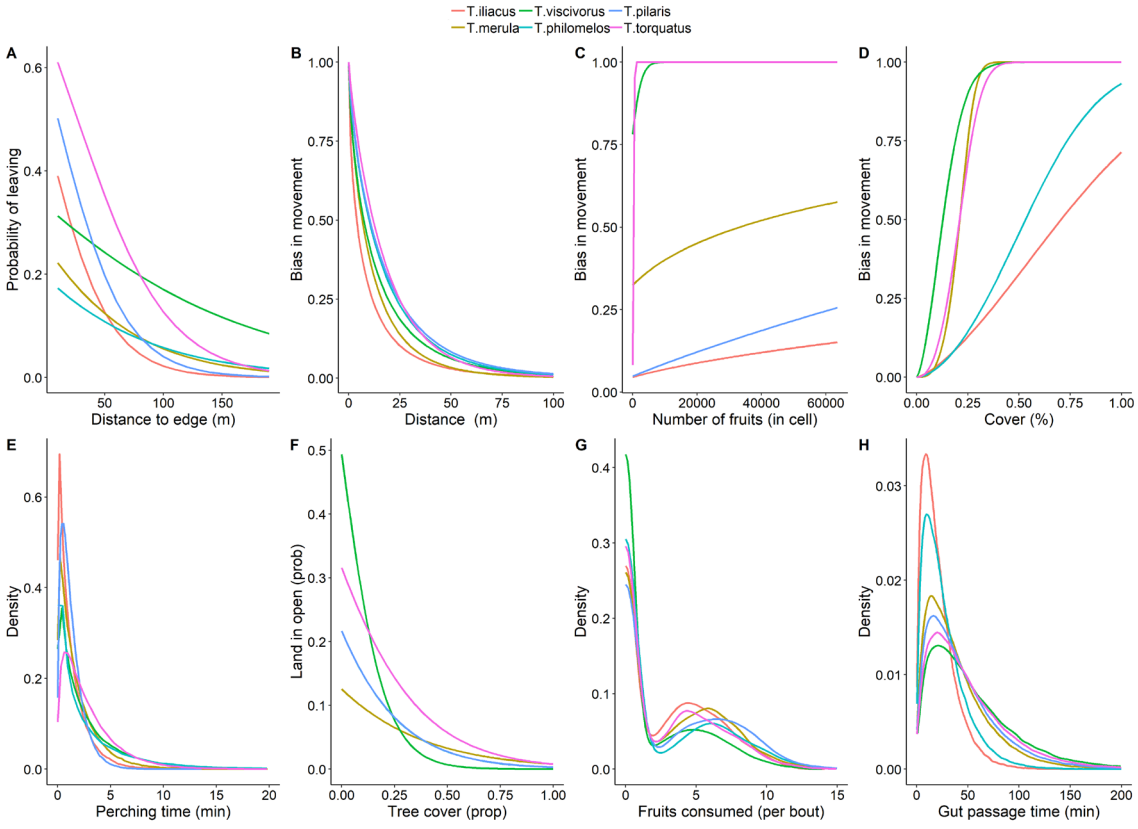


Fig. A6. Bird movement and foraging rules. (a) Probability of leaving the study plot as a function of the distance to the edge; bias in movement decisions related to (b) distance, (c) number of fruits, (d) tree cover. (e) Density distribution of perching times, (f) probability of landing in open areas as a function of tree cover. Density distribution of (g) fruit consumption and (h) gut passage times. Each line corresponds to a thrush species.

While visiting cells, if there are fruits available, the simulated bird decides how many fruits to consume based on a species-specific zero-inflated Poisson distribution, with a probability of fruit consumption (1-ϕ_k) and an expected number of fruits consumed (λ_k) (Fig. A6g). During a foraging bout, the model assigns a plant species identity to fruits consumed, based on the relative abundance of fruits of each species within each cell (i.e., there is no particular fruiting preference; García et al., 2013). After consumption, the model transforms the number of fruits into seeds consumed (based on estimates of the number of seeds per fruit for the different species, from a sample of 20–30 fruits from each of 10–15 plants per species collected in 2001 and 2002 in the study area (García et al., 2005). The time birds expend in cells is drawn from a gamma distribution with species-specific shape and rate parameters (Fig. A6e). Parameters to model fruit consumption and perching times were fitted using observations of thrush behavior in the study area from 2008 to 2011 (N=705).

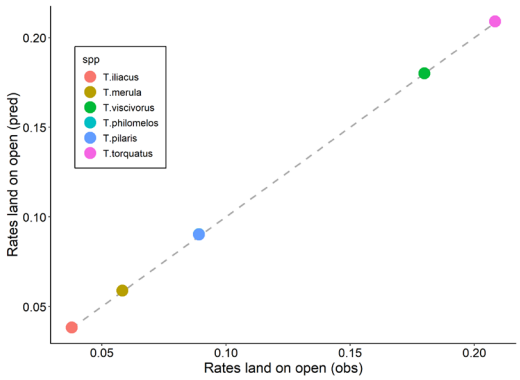


Fig. A7. Posterior predictive check of rates of landing on open microhabitats after arriving to a cell. Each point represents a species (observed and mean predicted values), grey dashed line 1:1 relationship (perfect match).

After perching time expires, the bird decides where to go next and, if it is to a different cell, it moves at a speed of 6 meters per second (Morales et al., 2013). While moving (or perching) birds defecate seeds of consumed fruits according to the time lapse since they were ingested and species-specific gut passage times. Following Morales *et al.* (2013) gut passage times (GPT) are drawn from a gamma distribution with a common shape parameter for all thrush species, but with a species-specific rate parameter (related to mean GPT values; Fig. A6g). After defecation, the model tracks the cell and the microhabitat where the seed is deposited. If the seed is defecated while the bird is perching on a tree, the microhabitat of deposition will be under tree cover. In contrast, if

the seed is defecated during bird displacement, the microhabitat of deposition is randomly sampled with a probability equal to its proportion of tree cover on the cell where the seed located.

In each model run we simulated 5,000 sequences in which simulated birds moved, foraged, and defecated. The identity of simulated birds was sampled from their mean relative abundance observed in the study area between 2008 and 2011 (*Turdus iliacus, T. merula, T. viscivorus, T. philomelos, T. pilaris* and *T. torquatus* with probabilities of 0.34, 0.28, 0.18, 0.17, 0.02, 0.01, respectively). In our landscape, forest areas act as sources of birds, hence, the initial location of birds was sampled randomly from forest cells. After each simulation, the model recorded the location of defecated seeds and the microhabitat of deposition.

A.3. Post-dispersal predation, germination and one-year survival

After seed deposition the model simulated species-specific post-dispersal predation, seed germination and one-year survival of seedlings accounting for the microhabitat of deposition. The parametrization of these demographic rates for the three target plant species (*I. aquifolium*, *C. monogyna* and *T. baccata*) was based on field studies developed in the study plot during 2009-2011 and, mostly, in areas of similar habitat physiognomy and composition nearby the study plot, in the Sierra de Peña Mayor (43º 179N, 5º 309W, 900 m a.s.l., Asturias Province), during 2002-2004 (for a comprehensive location and field methodology description see, for example, García et al. 2005).

Post-dispersal seed predation was parameterized using data from a field survey (conducted in winter 2002 and 2003) that monitored survival of seeds of the target plant species in 150 sampling stations distributed across different microhabitats (115 stations under the cover of the target plant species and 35 stations open), in areas close to the study plot. We placed 8 seed depots composed of plastic mesh (1.5-mm pore) triangles (6-cm side), in each sampling station, by nailing them to the ground at a distance of 50 cm from each other. One seed of each of the target plant species was glued to a different vertex of the triangle, using a low odor, rainproof thermoplastic glue. Seeds were glued firmly to the plastic triangles to prevent removal by wind and rain. We did not find trampling or the digging-up of triangles by animals during experiments. Thus, we considered that a seed had been preyed if: (1) it was missing from the plastic mesh; or

(2) it was still on the mesh but was gnawed and empty. We considered that seed predation was almost exclusively attributable to forest mice (*Apodemus* spp.). Depots were monitored after 2 and 4 weeks.

The germination of seeds was monitored in 2003-2004 in areas close to the study plot by means of 100 sampling stations (75 stations under the cover of the target of plant species and 25 in open microhabitats) separated at least 5 m one from each other. In each sampling station, we placed in February 2003 a set of 10 seeds dispersed by birds in the previous fruiting season (i.e., autumn 2002 and early winter 2003) and *a priori* viable (according to the buoyancy test) inside 5 cm x 5 cm glass-fiber bags (1 mm pore diameter) and buried them in the topsoil surface layer (at a depth of 3 cm). To quantify seed germination, bags were monitored in November 2004.

Seedling emergence and survival was monitored in two study periods: 2002-2003 and 2009-2011. In 2002-2003, we established permanently labeled quadrats (50 x 50 cm) in the 100 sampling stations described for germination survey. Quadrats were periodically revisited during spring and summer 2002 and spring 2003, searching, in each survey, for all emerged seedlings of the target plant species. Seedlings were individually identified (distinguished on the basis of the presence of cotyledons and stem color, see (Peterken & Lloyd, 1967; Thomas & Polwart, 2003), and their survival monitored monthly during spring and summer and after one-year. In 2009-2011, we selected 220 cells of the study plot following a chessboard pattern (for a comprehensive description of the sampling design, see Martínez & García (2017). Five sampling stations, separated from each other by 4 meters, were placed along the central north-south axis of each cell. In each station, we set a permanently labeled 50 × 50 cm quadrat on the ground and noted the type of microhabitat (under tree cover or open) where the quadrat was located. During spring-summer 2011 and 2012, we marked and counted the emerged seedlings of the target plant species following the same procedure described above for estimating survival during spring and summer and after one-year. The sample size of each species depended on the number of emerged seedlings, making it challenging to adjust data for rare species (i.e., *T. baccata*). Thus, to quantify the probability of one-year survival we pooled data from both sampling periods.

Rates of survival of dispersed seeds to post-dispersal predation, germination and survival of seedlings after one year were modeled as a binomial regression that

depended on the microhabitat of deposition (*open*, tree uncovered ground; or *cover*, under tree cover). In our regressions we set open microhabitat in the intercept. In some cases (i.e., predation rates of *Ilex aquifolium* and *Taxus baccata*) we fitted a beta-binomial regression to account for overdispersion in data. We fitted models with a Bayesian approach with Jags software (Plummer, 2003) using weakly informative priors. Intercept and microhabitat effects were sampled from a normal distribution with mean 0 and standard deviation of 10. In the case of beta-binomial regression we included an overdispersion parameter (phi) sampled from a normal distribution with mean 0 and standard deviation of 1 and exponentiated to ensure positive values. In all cases, after checking for chain convergence (Rhat<1.1) and suitable effective sample sizes (>1000), we performed a posterior predictive check to evaluate model fit (Fig. A8).

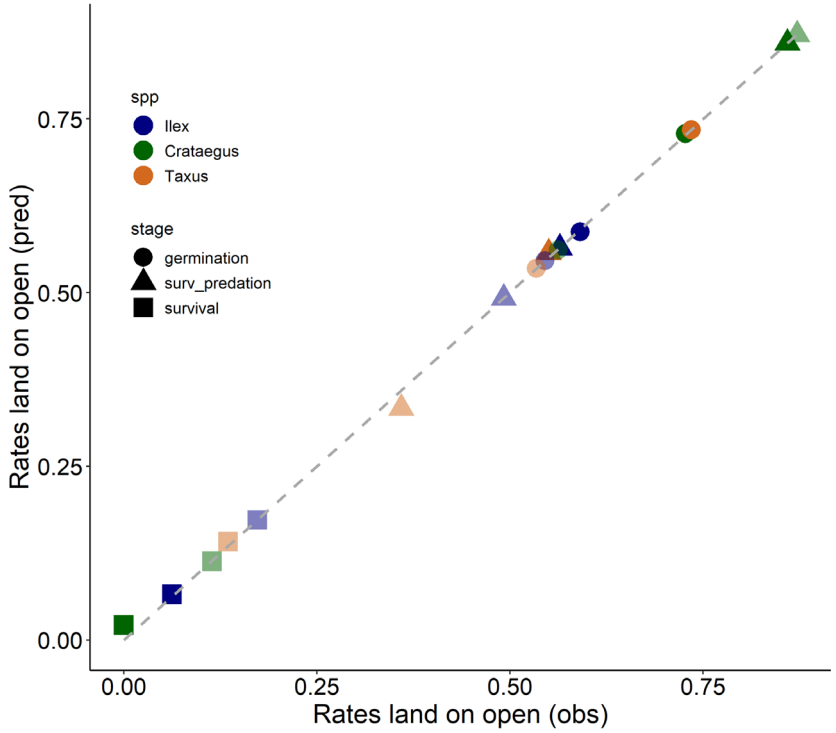


Fig. A8. Summary of posterior predictive checks of probabilities of survival to predation, germination rates and probability of one-year survival of seedlings. Transparency of points depicts the microhabitat of deposition (being transparent dots cover microhabitat). Dashed grey line represents the 1:1 relationship (perfect match).

To simulate post-dispersal demographic stages, we used a probability transition model. The probability of survival to post-dispersal seed predation was modeled as following:

$$y_{imp} \sim Binom(p_{pm}, T_{ipm}) \begin{cases} logit(p_{pm}) = \beta_{0p} + \beta_{1p} C_m & if\ p = 1\ (Ilex\ aquifolium) \quad \text{eq. A6} \\ & else \\ logit(mu_{pm}) = \beta_{0p} + \beta_{1p} C_m & \text{eq. A7} \\ alpha_{pm} = mu_{pm} \Phi_p & \text{eq. A8} \\ beta_{pm} = (1 - mu_{pm}) \Phi_p & \text{eq. A9} \\ p_{pm} \sim Beta(alpha_{pm}, beta_{pm}) & \text{eq. A10} \end{cases}$$

In each i-th cell, the number of seeds of each p-th plant species surviving per m-th microhabitat was sampled from a binomial distribution with probability of survival (p_{pm}) and a number of trials equal to the number of seeds present (T_{ipm}). The probability of survival of dispersed seeds depended on species-specific probabilities of survival and on the microhabitat where it was located ($\beta_{0p} + \beta_{1p} C_m$, being $C_m = 1$ if the microhabitat is cover). In the case of *Ilex aquifolium,* the probability of survival was deterministic (eq. A6), whereas in the rest of species it was sampled from a beta distribution with a mean value (mu_{pm}) that depended on the microhabitat of deposition (eq. A7) and a species-specific dispersion parameter (Φ_p).

Subsequently, the model simulated seed germination. In each cell, the number of seeds germinating was sampled from a binomial distribution ($G_{imp} \sim Binom(\psi_{pm}, \hat{y}_{ipm})$, eq. A12) with a number of trials equal to the number of seeds surviving post-dispersal predation ($\hat{y}_{ipm}$) and a probability of germination (ψ_{pm}) that depended on the plant species and the microhabitat ($logit(\psi_{pm}) = \alpha_{0p} + \alpha_{1p} C_m$, eq. A13). Similarly, the number of emerged seedlings that survived one year was drawn from a binomial distribution with a probability of one-year survival ($S_{imp} \sim Binom(\eta_{pm}, \hat{G}_{ipm})$ eq. A14) that depended on the microhabitat of deposition and the species of seedlings ($\eta_{pm} = \gamma_{0p} + \gamma_{1p} C_m$, eq. A15). During each run the model tracked in each landscape cell the number of seeds surviving predation, seedlings emerging and surviving after one year and the microhabitat where they were located.

References

Donoso, I., García, D., Rodríguez-Pérez, J., & Martínez, D. (2016). Incorporating seed fate into plant–frugivore networks increases interaction diversity across plant regeneration stages. *Oikos*, *125*(12), 1762–1771. https://doi.org/10.1111/OIK.02509

García, D., Martínez, D., Herrera, J. M., & Morales, J. M. (2013). Functional heterogeneity in a plant–frugivore assemblage enhances seed dispersal resilience to habitat loss. *Ecography*, *36*(2), 197–208. https://doi.org/10.1111/J.1600-0587.2012.07519.X

García, D., Obeso, J. R., & Martínez, I. (2005). Rodent seed predation promotes differential recruitment among bird-dispersed trees in temperate secondary forests. *Oecologia*, *144*(3), 435–446.

Martínez, D., & García, D. (2017). Role of avian seed dispersers in tree recruitment in woodland pastures. *Ecosystems*, *20*(3), 616–629.

Morales, J. M., García, D., Martínez, D., Rodriguez-Pérez, J., & Herrera, J. M. (2013). Frugivore Behavioural Details Matter for Seed Dispersal: a Multi-Species Model for Cantabrian Thrushes and Trees. *Plos One*, *8*, e65216. https://doi.org/10.1371/journal.pone.0065216

Peterken, G. F., & Lloyd, P. S. (1967). Ilex aquifolium L. *Journal of Ecology*, *55*(3), 841–858.

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, *124*(125.10), 1–10.

Thomas, P. A., & Polwart, A. (2003). Taxus baccata L. *Journal of Ecology*, *91*(3), 489–524.

A.4. Summary tables of model parameters

Table. A1. Summary of model parameters used to model the composition of matrix cells turning into forest areas. In both cases, we used a density distribution fitted to observed data to simulate new values. Tree cover corresponds only to cells classified as forests (proportion of cover > 0.2) and proportion of fleshy-fruited species to landscape cells with some level of tree cover.

Parameter	Interpretation	Mean	sd	Median	L ($q_{0.05}$)	U($q_{0.95}$)
Tree cover (m^2)	Number of m^2 of cover of the cell that correspond to tree cover	265.15	103.39	289.65	95.71	400.00
Prop fleshy-fruited (proportion)	Proportion of tree cover within cells that belong to fleshy-fruited species	0.73	0.25	0.73	0.27	1.00

Table A2. Summary of model parameters involved in the composition of fleshy-fruited species in forest recovered cells, and in fruit production in all cells within the landscape.

Subprocess	Parameter	Output value	*Ilex aquifolium*	*Crataegus monogyna*	*Taxus baccata*
Composition	Φ_p	Probability that a m^2 of cover of fleshy-fruited species belongs that the p-th species	0.70	0.21	0.09
Fruit production	P_p	Probability that a cell with cover of the p-th species produces fruits	0.66	0.54	0.36
	f_{0p}	Mean fruit production of the p-th species (log-transformed)	9.09	8.02	8.73
	f_{1p}	Effects of cover of the p-th species on its fruit production	0.43	0.37	0.33
	Mean	Mean values of cover (m^2) in landscape cells where the p-th plant species is present*	99.01	35.38	41.52
	sd	Standard deviation (m^2) of cover in landscape cells with cover where the p-th plant species is present*	105.83	34.87	41.75

*Values that are necessary in simulations because regressions were fitted to scaled values of cover.

Table A3. Summary of model parameters related to frugivory and seed dispersal sub-processes.

Behavior/Param	Output	Parameter	*T. iliacus*	*T. merula*	*T.philomelos*	*T. pilaris*	*T.torquatus*	*T.viscivorus*	Eq.	Obs
Relative abundance	Species of thrush	δ_k	0.34	0.28	0.17	0.02	0.01	0.18		
Leave the plot	Probability	a_{0k}	-0.07	-1.08	-1.43	0.36	0.71	-0.70	A1.	Input (units): $m/10^2$
		b_{0k}	-3.72	-1.75	-1.36	-3.53	-2.64	-0.89		
Bias distance	0-1 bias	a_{dk}	0.01	0.02	0.02	0.03	0.03	0.02	A2.	Input (units): $m/10^3$
		b_{dk}	0.52	0.65	0.67	0.67	0.79	0.60		
Bias cover	0-1 bias	a_{ck}	1.08	0.25	0.75	0.27	--	0.19	A3.	
		b_{ck}	1.40	3.28	1.76	2.38	--	1.48		
Bias fruit production	0-1 bias	a_{fk}	16.83	2.58	3.58	--	0.11	0.10	A4.	Input (units): Fruits/10^5
		b_{fk}	0.60	0.34	0.85	--	50.20	2.20		
Landing on open areas	Probability	a_{lk}	-1.94	-1.94	-1.28	-1.28	-0.77	-0.02	A6.	
		a_{lk}	-2.89	-2.89	-4.57	-4.56	-4.06	-9.83		
Fruit consumption	Probability	(1-ϕ_k)	0.47	0.48	0.63	0.50	0.55	0.67		
	Number of fruits	λ_k	5.17	5.92	6.57	6.56	5.56	5.15		
Perching time	Minutes	$shape_k$	0.85	0.93	0.57	1.47	1.26	0.75		
		$rate_k$	0.64	0.49	0.20	1.11	0.44	0.27		
Gut passage time	Minutes	$shape'$	1.59	1.59	1.59	1.59	1.59	1.59		
		$rate'_k$	0.07	0.04	0.06	0.04	0.03	0.03		

Table A4. Summary of model parameters involved in post-dispersal demographic processes.

Subprocess	**Microhabitat**	***Crataegus monogyna***	***Ilex aquifolium***	***Taxus baccata***
Post-dispersal survival	Open	0.86	0.54	0.56
	Cover	0.87	0.49	0.33
Probability of germination	Open	0.73	0.59	0.74
	Cover	0.56	0.55	0.53
One-year seedling survival	Open	0.01	0.06	0.03
	Cover	0.11	0.17	0.005

Appendix B: Simulation of restoration scenarios

B.1. Applied nucleation vs natural regeneration

To compare the effectiveness of applied nucleation with respect to natural regeneration on attracting frugivorous birds and promoting tree recruitment into the matrix, we simulated landscapes with increasing levels of matrix restored (from 10 to 50% of matrix transformed into forest), but with contrasting spatial patterns (Fig. B1).

We simulated natural regeneration following a growth-from-edge pattern (i.e., a diffusion of new forest cover from the border of extant forest patches to the matrix), which is frequently found in spontaneous forest recovery (Mouillot et al., 2005). To this end, we randomly and iteratively sampled a matrix cell adjacent to the forest and transformed it into forest habitat. In each step, habitat classification of cells was updated so that new edges candidate to grow emerged. In the case of applied nucleation, we created tree islets of 40 x 40 m so that the width of tree islets matched the scale of movement of *Turdus* (according to species-specific distance-bias functions, Fig. A6B). To this end, we first sampled a number of cells (initial cells, hereafter) across the landscape. The number of initial cells was calculated as a function of the total area of matrix transformed into forest (m^2) and the size of tree islets (0.16 ha). Initial cells were sampled randomly from the matrix and expanded orthogonally to form a square shape. We established a buffer of 1 cell width around the tree islet to avoid coalescence among them. In this way we could control for islet size in scenarios in which high values of matrix transformed into forest cells. Otherwise, we could have created too large islets due to the coalescence of several of them. In both restoration scenarios, the creation of new forest cells stopped when we reached the target values of matrix transformed into forest.

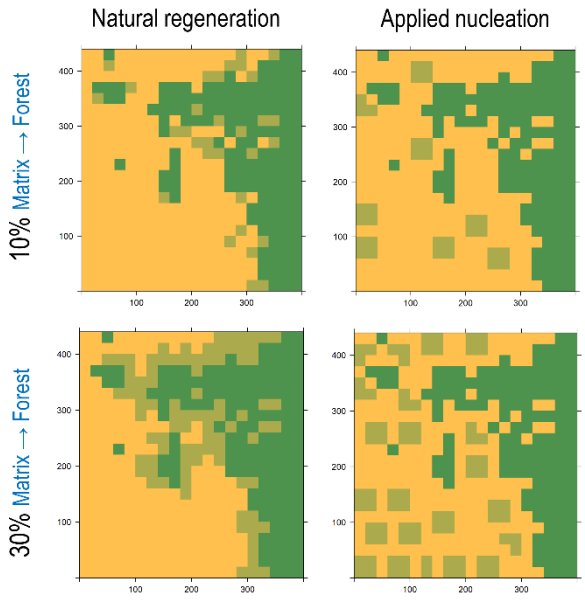


Fig. B1. Example of simulated landscapes in which matrix was transformed into forest under scenarios of natural regeneration (with a growth-from edge pattern, left) and applied nucleation (right). Upper and lower panels depict different levels of matrix turning into forest (10 and 30%, respectively). Dark green and orange colors represent cells of forest and matrix habitat in our study plot (real landscape), light green depicts forest recovered (i.e., matrix cells during simulation set-up transformed into forest).

B.2. Simulating tree islets of contrasting size

To evaluate the effect of islet size and area recovered on tree recruitment into the matrix we created patches of new forest with a square shape and a width of 20, 40, 60 and 80 m. The algorithm for generating new patches was the same as in the section B.1. According to the area of matrix to be transformed and tree islet size we randomly located a number of initial cells across the matrix. These initial cells expanded orthogonally to create a square patch of new forest. Then, we created a buffer of 1 cell-width around tree islets in which no new forest cell could be established. Finally, the creation of new tree islets stopped once we reached the target values of forest recovery were attained.

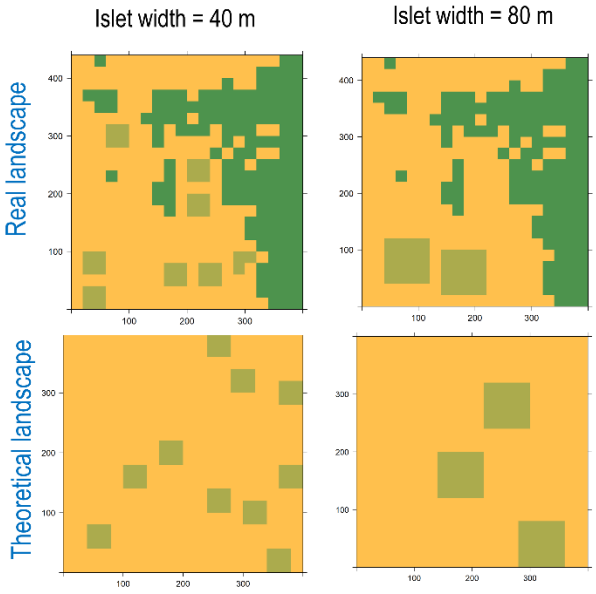


Fig. B2. Example of simulated applied nucleation with different tree islet sizes in real (upper) and theoretical (lower) landscapes. In both cases 20% of the matrix was turned into forest, but the *a priori* size of tree islets differed. Dark green and orange colors represent cells of forest and matrix habitat in our study plot (real landscape), light green depicts forest recovered (i.e., matrix cells during simulation set-up transformed into forest).

References

Mouillot, F., Ratte, J. P., Joffre, R., Mouillot, D., & Rambal, S. (2005). Long-term forest dynamic after land abandonment in a fire prone Mediterranean landscape (central Corsica, France). Landscape Ecology 2005 20:1, 20(1), 101–112. https://doi.org/10.1007/S10980-004-1297-5

Appendix C: Supplementary results

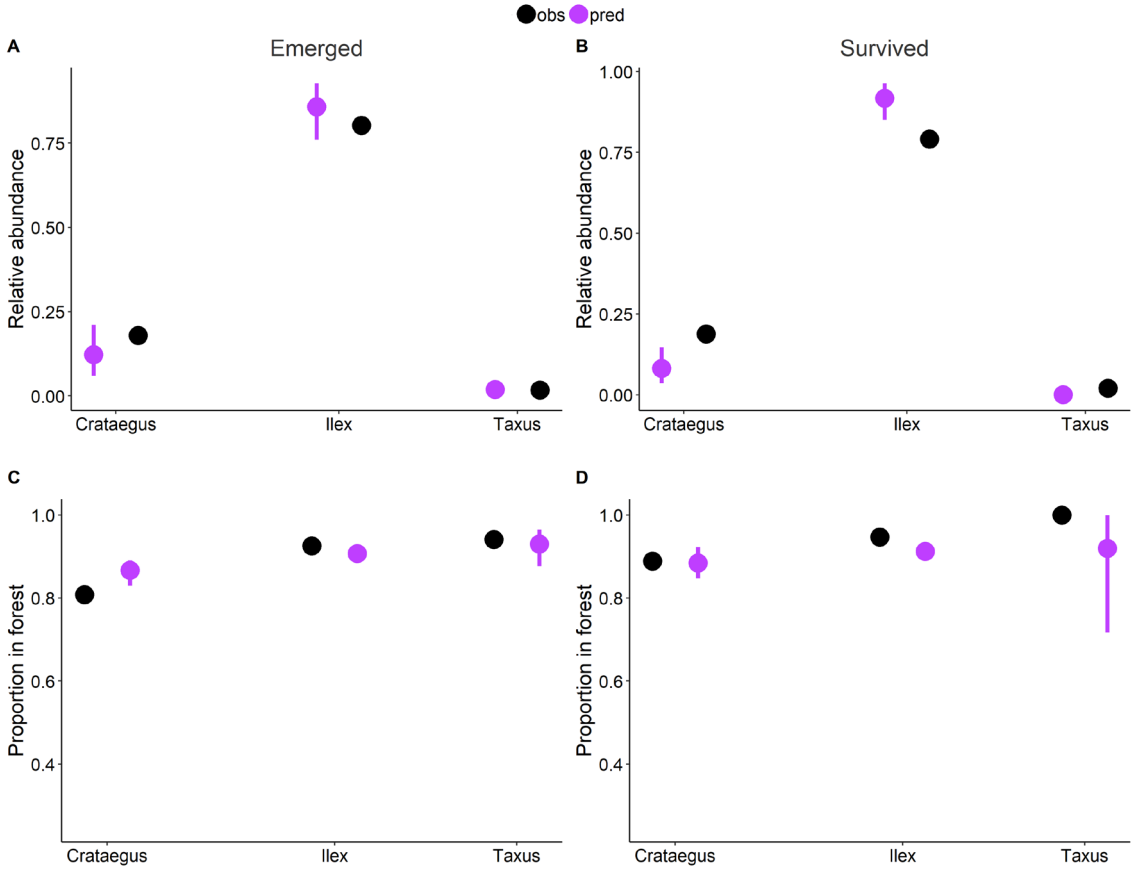


Fig. C1. Comparison of model outputs performed in the real landscape and patterns of recruitment observed in 2011 in the study plot. Panels A-B relative abundance of seedlings of each species and C-D proportion of seedlings observed in forest habitat (*vs* matrix). Magenta dots represent mean values across 50 repetitions and lines q005 and q095 quantiles. Black dots depict observed values.

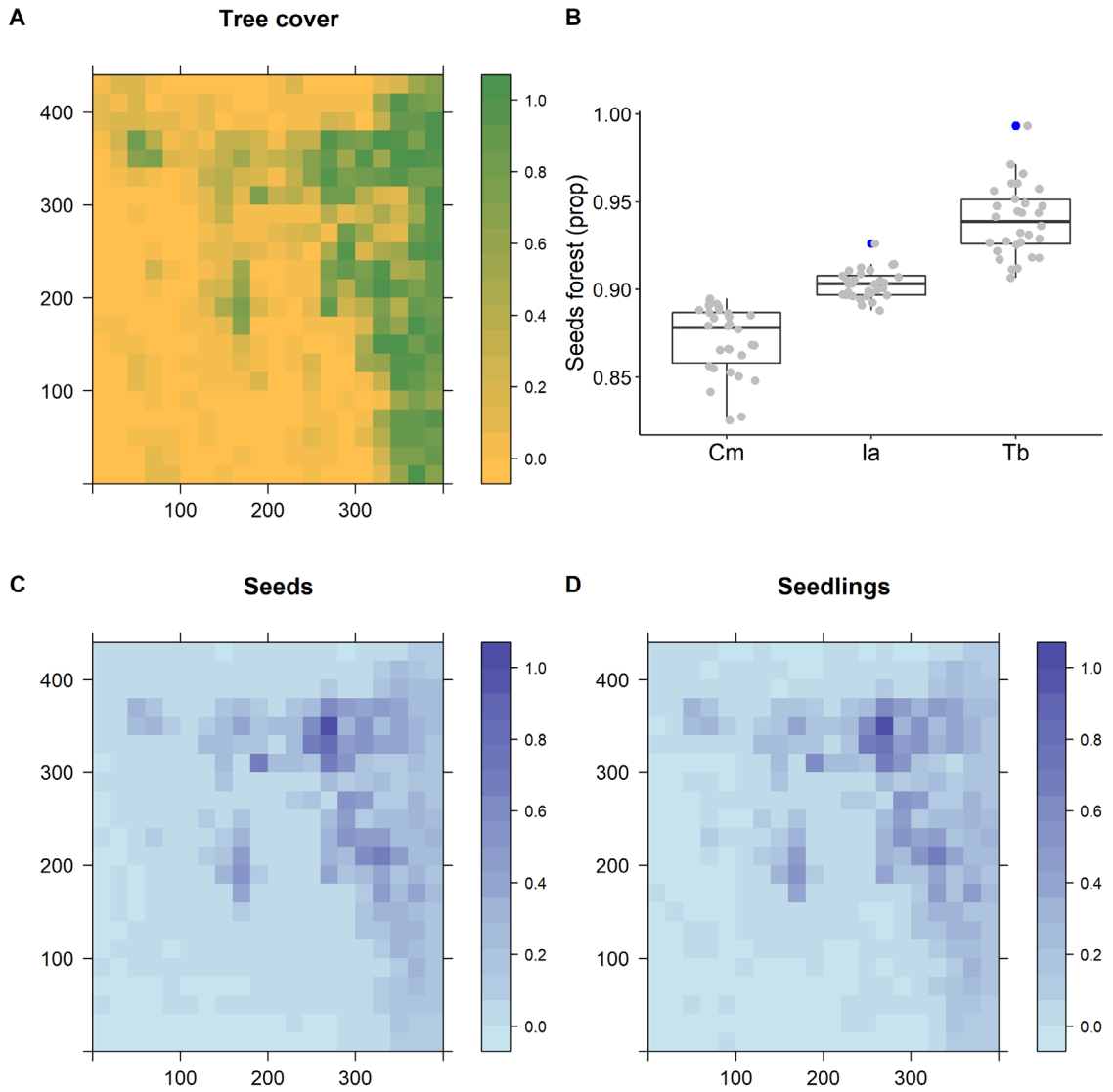


Fig. C2. (A) Map of tree cover in real landscape, (B) proportion of seeds arriving to forest habitats of the three fleshy-fruited species present in the area. (C) Seeds arriving and (D) seedlings recruiting (min-max normalized) in each landscape cell.

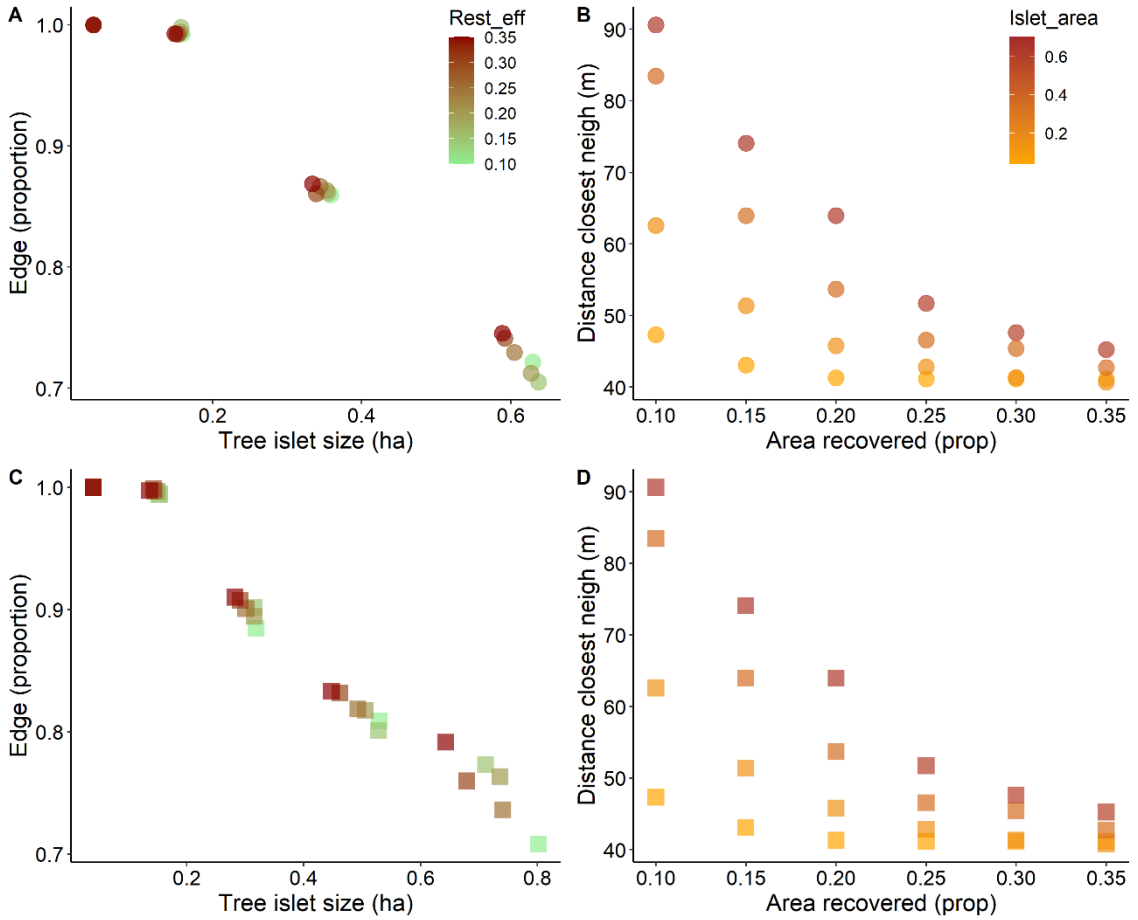


Fig. C3. (Left pannels) Effects of size (ha) on the proportion of edge of tree islets in the theoretical (A) and (C) real landscape. Green to brown colors depict the proportion of matrix transformed into forest (area recovered). (Right pannels) Effect of the proportion of matrix area recovered on the minimum distance between islets in the (B) theoretical and (D) real landscape. Red to yellow colors depict islets of different sizes. In all pannels points correspond to mean values across 30 repetitions.

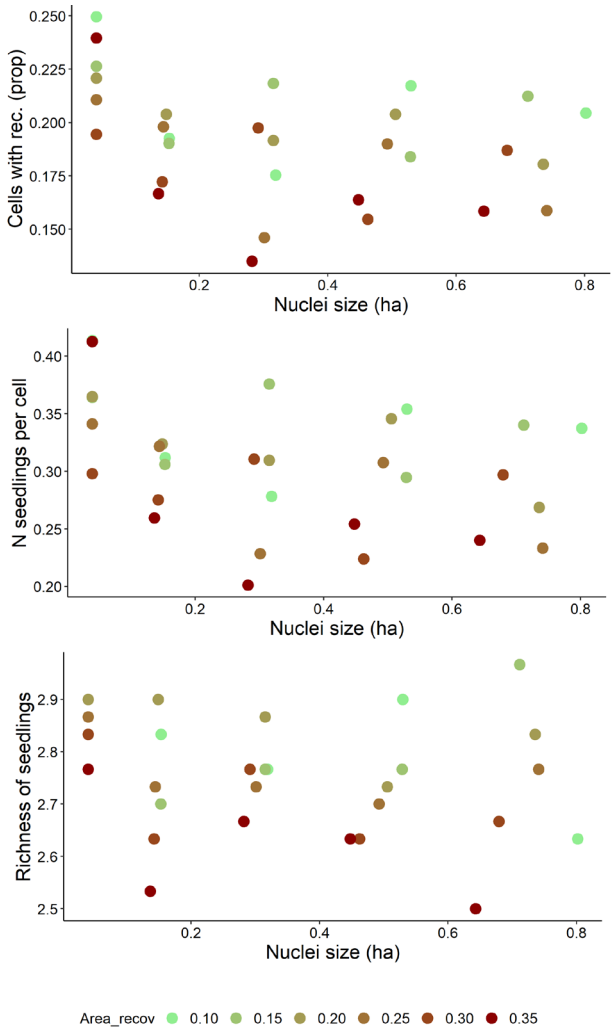


Fig. C4. Effects of size of tree islets (ha) and area recovered (proportion of matrix tranformed into forest) on the (A) proportion of matrix cells recruiting seedlings, (B) number of seedlings per matrix cell and (C) their richness. Dots depict mean values of response variables and size of tree islets (ha) across 30 repetitions of varying islet width (from 20 to 80 m) and the proportion of matrix transformed into forest (restoration effort, color code). In all cases tree islets were generated in our real landscape, which is characterized by a large forest patch and small patches scattered throughout the matrix (Fig. A2).

Table C1. Effects of tree islets size and area recovered (as the proportion of matrix area transformed into forest cells) in the (i) proportion of matrix cells recruiting one-year old seedlings, (ii) number of seedlings per matrix cell and (iii) their richness. Mean values of posterior distributions, credibility intervals ($q_{0.025}$, $q_{0.975}$), Rhat, number of effective sample size (Neff) and proportion of the posterior with the same sign as the mean (f). Effects with an f value > 0.95 are in bold. Tree islets were generated in our real landscape, which is characterized by a large forest patch and small patches scattered throughout the matrix (Fig. B4).

Response	Covariate	Mean	CI	Rhat	Neff	f
Matrix cells with seedling recruitment	Intercept	-1.44	[-1.51, -1.36]	1	67500	1
	Islet size	**-0.07**	**[-0.14, 0]**	**1**	**67500**	**0.97**
	Area recovered	**-0.09**	**[-0.17, -0.02]**	**1**	**67500**	**0.99**
	Size*Area	-0.03	[-0.11, 0.05]	1	60139	0.79
Seedlings per cell	Intercept	3.79	[3.19, 4.3]	1	1917	1
	Islet size	**-0.07**	**[-0.13, -0.01]**	**1**	**31092**	**0.99**
	Area recovered	**-0.09**	**[-0.15, -0.04]**	**1**	**74250**	**1**
	Size*Area	-0.04	[-0.1, 0.02]	1	74250	0.88
Richness	Intercept	2.76	[2.72, 2.79]	1	67500	1
	Islet size	**-0.03**	**[-0.07, 0]**	**1**	**67500**	**0.97**
	Area recovered	**-0.07**	**[-0.1, -0.03]**	**1**	**67500**	**1**
	Size*Area	-0.01	[-0.05, 0.03]	1	63230	0.68

Table C2. Summary of regressions between matrix colonization metrics and islet size in different scenarios of bird mobility (m). For the proportion of matrix cells with recruitment and number of seedlings per cell, we fitted a negative exponential function. In the case of richness, we fitted a linear regression. Scatterplots with mean fitted responses can be found in Fig. 3 in the main text. CI depicts confidence interval of estimated parameters.

Response	**Scale of mov.**	**Intercept**	**$IC_{intercept}$**	**Slope**	**CI_{slope}**
Matrix cells recruiting seedlings	20	0.14	[0.12, 0.16]	5.61	[4.26, 7.42]
	40	0.22	[0.21, 0.24]	2.2	[1.89, 2.54]
	80	0.3	[0.28, 0.31]	1.12	[0.94, 1.31]
Number of seedlings per cell	20	0.31	[0.27, 0.36]	4.5	[6.23, 10.41]
	40	0.39	[0.36, 0.42]	5.5	[2.06, 2.90]
	80	0.53	[0.49, 0.57]	6.5	[0.97, 1.56]
Seedling richness	20	2.49	[2.35, 0.64]	-1.58	[-1.99, -1.18]
	40	2.91	[2.85, 2.97]	-0.95	[-1.13, -0.78]
	80	2.88	[2.84, 2.94]	-0.2	[-0.31,-0.08]