



# Calibrating an individual-based movement model to predict functional connectivity for little owls

SEVERIN HAUENSTEIN <sup>1,8</sup> JULIEN FATTEBERT <sup>2,3</sup> MARTIN U. GRÜEBLER <sup>2</sup> BEAT NAEF-DAENZER <sup>2</sup>  
GUY PE'ER <sup>4,5,6</sup> AND FLORIAN HARTIG <sup>1,7</sup>

<sup>1</sup>Department of Biometry and Environmental System Analysis, University of Freiburg, 79106 Freiburg, Germany

<sup>2</sup>Swiss Ornithological Institute, CH-6204 Sempach, Switzerland

<sup>3</sup>School of Life Sciences, University of KwaZulu-Natal, 4000 Durban, South Africa

<sup>4</sup>German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, 04103 Leipzig, Germany

<sup>5</sup>Department of Conservation Biology, UFZ–Helmholtz Centre for Environmental Research, Department of Economics and Department Ecosystem Services, 04318 Leipzig, Germany

<sup>6</sup>University of Leipzig, 04109 Leipzig, Germany

<sup>7</sup>Theoretical Ecology, University of Regensburg, 93053 Regensburg, Germany

*Citation:* Hauenstein, S., J. Fattebert, M. U. Gruebler, B. Naef-Daenzer, G. Pe'er, and F. Hartig. 2019. Calibrating an individual-based movement model to predict functional connectivity for little owls. *Ecological Applications* 00(00):e01873. 10.1002/eap.1873

**Abstract.** Dispersal is crucial for population viability and thus a popular target for conservation measures. However, the ability of individuals to move between habitat patches is notoriously difficult to estimate. One solution is to quantify functional connectivity via realistic individual-based movement models. Such simulation models, however, are difficult to build and even more difficult to parameterize. Here, we use the example of natal little owl (*Athene noctua*) dispersal to develop a new analysis chain for the calibration of individual-based dispersal models using a hybrid of statistical parameter estimation and Approximate Bayesian Computation (ABC). Specifically, we use locations of 126 radio-tracked juveniles to first estimate habitat utilization by generalized additive models (GAMs) and the biased random bridges (BRB) method. We then include the estimated parameters in a spatially explicit individual-based model (IBM) of little owl dispersal and calibrate further movement parameters using ABC. To derive efficient summary statistics, we use a new dimension reduction method based on random forest (RF) regression. Finally, we use the calibrated IBM to predict the dispersal potential of little owls from local populations in southwestern Germany to suitable habitat patches in northern Switzerland. We show that pre-calibrating habitat preference parameters while inferring movement behavioral parameters via ABC is a computationally efficient solution to obtain a plausible IBM parameterization. We also find that dimension reduction via RF regression outperforms the widely used least squares regression, which we applied as a benchmark approach. Estimated movement parameters for the individuals reveal plausible inter-individual and inter-sexual differences in movement behavior during natal dispersal. In agreement with a sex-biased dispersal distance in little owls, females show longer individual flights and higher directional persistence. Simulations from the fitted model indicate that a (re)colonization of northern Switzerland is generally possible, albeit restricted. We conclude that the presented analysis chain is a sensible work-flow to assess dispersal connectivity across species and ecosystems. It embraces species- and individual-specific behavioral responses to the landscape and allows likelihood-based calibration, despite an irregular sampling design. Our study highlights existing, yet narrow dispersal corridors, which may require enhancements to facilitate a recolonization of little owl habitat patches in northern Switzerland.

**Key words:** *Approximate Bayesian Computation; Athene noctua; choosing optimal summary statistics; IBM parameterization; landscape connectivity; movement ecology.*

## INTRODUCTION

Understanding and predicting animal dispersal is central for ecology and conservation. One important aspect of dispersal is connectivity: the flow of

individuals or genes within or among populations (e.g., Jangjoo et al. 2016), which is crucial for their long-term persistence (Hanski 1998). Estimating the degree of connectedness for particular species, functional groups, or individuals has led to the concept of functional connectivity, which recognizes that effective ecological connectivity emerges from the interplay of landscape features with species traits (Taylor et al. 1993, Baguette and Van Dyck 2007).

Manuscript received 31 August 2018; revised 19 December 2018; accepted 1 February 2019. Corresponding Editor: Marissa L. Baskett.

<sup>8</sup> E-mail: severin.hauenstein@biom.uni-freiburg.de

While the concept of functional connectivity is intuitively attractive, its application is often complicated (Pe'er et al. 2011). The most direct way to quantify functional connectivity is to directly measure habitat use and movements of dispersing individuals using data-driven models (Killeen et al. 2014, Fattbert et al. 2015), but tracking a sufficient number of dispersing individuals across all possible landscape configurations is virtually impossible. Widely used are simple semi-mechanistic approaches, such as least-cost path analysis (e.g., Stevens et al. 2006, Rabinowitz and Zeller 2010) and circuit theory (e.g., McRae et al. 2008, Pelletier et al. 2014) and, less frequently, neutral percolation models (see With 1997). While useful for approximately estimating gene flow and identifying dispersal corridors, these approaches usually fail to consider the often available detailed ecological knowledge about the target species and make the unwarranted assumption of perfect knowledge of the landscape by dispersing individuals.

For more detailed simulations of functional connectivity, we would ideally build individual-based models (IBMs), which simulate dispersal trajectories while accounting for both the landscape structure and an individual's movement processes explicitly (e.g., Allen et al. 2016). In principle, IBMs can consider any ecological process. The challenge, however, is that these processes must be calibrated to data, for example empirical observations of a species' habitat use and movement during dispersal (Nafus et al. 2017, Signer and Ovaskainen 2017, Buderman et al. 2018). The calibration of such stochastic simulation models is not trivial, because the hierarchy of the stochastic processes makes their likelihood usually intractable and prevents the possibility of using established statistical parameter estimation methods (Hartig et al. 2011). Irregular sampling designs, which are common for (radio)telemetry data, exacerbate this problem.

A common approach for calibrating complex IBMs with intractable likelihoods is Approximate Bayesian Computation (ABC; Beaumont et al. 2002, Hartig et al. 2011). ABC uses a simulation-based approximation of the model likelihood, based on summary statistics that measure the distance between observed and simulated data. As a positive side effect, problems of standard statistical methods that frequently occur with movement data, such as missing or irregular observations, can be seamlessly resolved by this approach. However, ABC can also be computationally costly, and these costs increase substantially with the number of parameters to estimate. A second problem of ABC is that efficient parameter inference hinges on the definition of appropriate summary statistics that reduce the dimensionality of the simulated and observed data, while retaining the important information for the inference (sufficiency). As a consequence, few studies have used ABC to parameterize movement models so far (but see, e.g., Marchand et al. 2017, Zhang et al. 2017), and there is very little guidance about the best methods to do so.

Here, we consider these questions in the light of a typical conservation problem, the dispersal potential of the little owl (*Athene noctua*) populations around the border triangle of Germany, Switzerland, and France. The little owl has been affected by land use changes in central European farmland areas over the last 60 yr and has suffered extreme declines in many European countries (e.g., Van Nieuwenhuyse et al. 2008). Recently, three local populations in southwestern Germany have started to recover after implementing nest box programs, producing dispersers beyond local populations (Fiedler 2013). However, it remains poorly known whether the dispersing individuals from these populations are able to reach suitable habitat patches and potentially (re)colonize new or improved former habitats in northern Switzerland.

To assess whether such recolonization is possible, we built a spatially explicit IBM to simulate movement of dispersing juvenile little owls during their transfer phases. The model evaluates relocation decisions based on habitat preferences and movement behavior. For model calibration, we use radio telemetry data of 126 radio-tracked juveniles from the largest little owl population in southwestern Germany. We propose two novel technical modifications to remedy the outlined problems of current ABC approaches. Firstly, we reduce the number of parameters to estimate via ABC using strong priors of habitat preference parameters from an independent likelihood-based parameter estimation. Specifically, we estimate habitat suitability with a Generalized Additive Model (GAM; Hastie and Tibshirani 1990), modeling the utilization distributions of juvenile little owls, derived from the biased random bridges kernel method (BRB; Benhamou 2011), as a function of several environmental covariates. These estimates are incorporated in the IBM as a priori fixed values and linked to the movement process through an individual-specific importance weight. We then use ABC rejection sampling to estimate four movement behavioral parameters, including a habitat suitability importance parameter. As our second technical modification, we use random forest (RF) regression for the dimension reduction of the summary statistics. A dimension reduction is a standard step in ABC approaches to improve the efficiency of the estimation, but usually simpler (linear) approaches are used for this task.

We assess the plausibility of the estimated parameter values through visual exploration of the little owl movement behavior of little owls, with a focus on inter-individual and inter-sexual differences. We also test the ability of RF regression as dimension reduction approach in ABC rejection sampling and compare it to the widely used least squares regression. Finally, we use the calibrated IBM to predict the functional connectivity between the expanding southwestern German little owl populations and the unoccupied habitat patches in northern Switzerland with the objective to identify potential dispersal corridors and barriers.

## METHODS

*Study area and species*

Our primary study area, where movement data were collected, is located in the southwestern federal German state of Baden-Württemberg (BW) in the district of Ludwigsburg. The area covers 687 km<sup>2</sup>, and includes a large part of the Neckarbecken, a large valley in central and northern Baden-Württemberg (Fig. 1b). Altitudes within the Neckarbecken range between 120 and 250 m above sea level (asl), the maximum altitude for the recorded data is 540 m asl. The landscape mosaic within the area is composed of farmland (58%; predominately intensively used agricultural fields, as well as meadows, orchards, and vineyards), forests (25%), and urban settlements (17%). With 850 residents/km<sup>2</sup>, the area is relatively densely populated.

For predictions, we were interested in an extended area, comprising the entire state of Baden-Württemberg, plus 80-km extensions southwards into northern Switzerland (CH) and westwards into northeastern France (F). These areas cover important targets for little owl recolonization into Switzerland with potential dispersal corridors. Important landscape features are the Rhinriver, Lake Constance and the elevations of the Black Forest and the Swabian Alb in BW, the foothills

of the Swiss Jura mountains and the Alps in northern Switzerland, and the Vosges mountains in northeastern France (see Fig. 1a).

The little owl (*A. noctua*) is a primarily nocturnal generalist mesopredator, inhabiting various types of open, agricultural landscapes. Little owls form monogamous territorial pairs, show biparental care, and fledglings of both sex are obligate dispersers (Van Nieuwenhuysse et al. 2008). As priority species in Switzerland and the subject of strict protection in Germany, the little owl has recently become a target of extensive single-species conservation management (Keller et al. 2010). In the past 50 yr and over much of Europe, it suffered vast population declines (Van Nieuwenhuysse et al. 2008). As a hemerophilic species with a preference for open agricultural landscapes, the little owl has been particularly threatened by the consequences of agricultural intensification and landscape fragmentation (Devictor et al. 2008). The three little owl populations in BW, however, are currently expanding, producing dispersers beyond the range of the local populations (Fiedler 2013). The largest population is spread across the Neckarbecken and the other two along the Upper Rhine valley around the Kaiserstuhl and near Lörrach (as indicated by the ringing records between 2011 and 2015 in Fig. 1a; Vogelwarte Radolfzell 2016). Also, there is at least one little owl population west of the river Rhine in northeastern

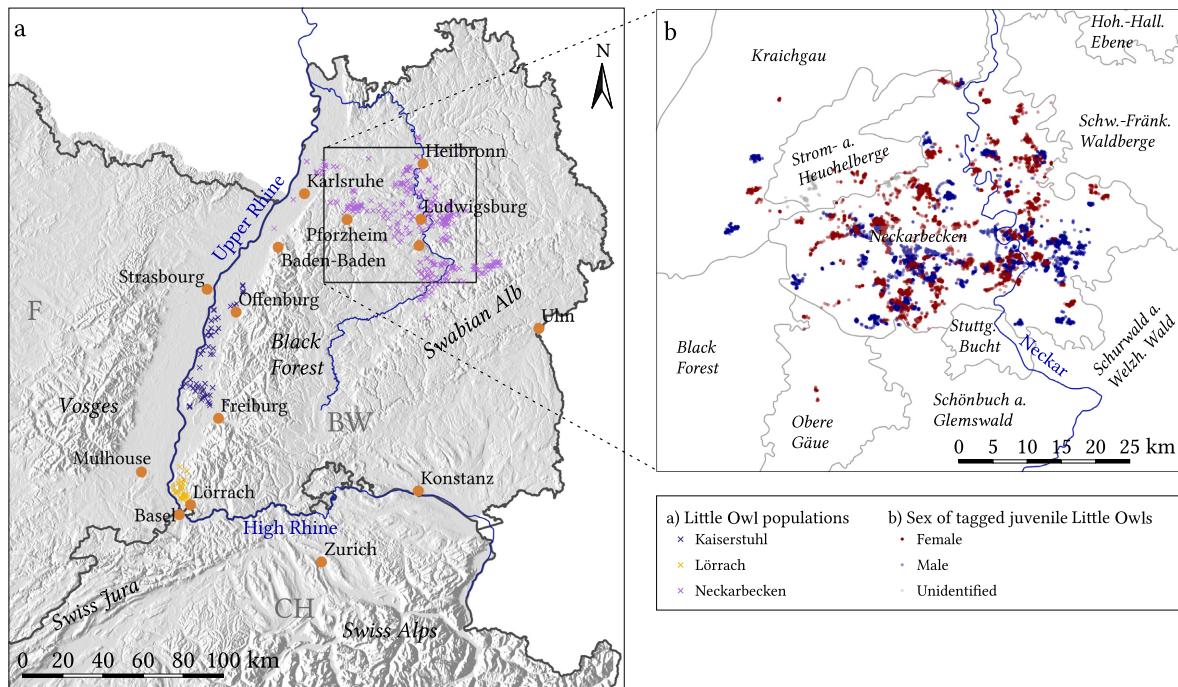


FIG. 1. (a) Extended study area (terrain map) illustrating landscape features including natural areas, larger cities (orange), and major rivers (blue). Ringing records of juvenile little owls between 2011 and 2015 (Vogelwarte Radolfzell 2016) in Baden-Württemberg (BW) are displayed as yellow (population at Lörrach), blue (Kaiserstuhl), and purple (Neckarbecken) crosses. (b) Primary study area with radio telemetry locations of juvenile little owls tagged in the Neckarbecken between 2009 and 2011 (females in red, males in blue, unidentified sex in gray), river Neckar (blue line) and natural regions.

France, for which we could not obtain any data. In Switzerland there are only sporadic little owl sightings and no recent breeding records (Knaus et al. 2018).

*Bird tracking data*

The little owl population within the primary study area (Neckarbecken) consists of approximately 220 breeding pairs (H. Keil, unpublished data). Between 2009 and 2011, 127 juveniles from this population were equipped with miniature, VHF radio transmitters (Naef-Daenzer et al. 2005) at their nesting sites before they fledged. Of these, 14 individuals (3 males, 5 females, 6 unidentified) were tagged in 2009, 48 individuals (27 males, 20 females, 1 unidentified) in 2010, and 65 individuals (25 males, 40 females) in 2011 (Bock et al. 2013). All of them were tagged in their nest boxes at around day 30 after hatching. With 6.9–7.2 g (slight mass differences due to manual production), the VHF transmitters weighed approximately 4–5% of the little owls’ body mass. The transmitters had an expected lifespan of 380 d and were fit to the bird with a standard figure-8 harness (Kenward 2000). The estimated survival probabilities of the tagged owls were not different compared to estimates based on ring recoveries in other little owl studies (Naef-Daenzer et al. 2017).

During their first year, the tagged owls were located by multiple observers at least once per week, mostly at night. Upon encounter, the homing-in method was used for locating the birds (Kenward 2000). Since the objective was to record dispersal movements, the fixes were taken in 5-min intervals, but only if the individual was actively moving during the observation session. In that case, they were followed until the signal was lost or the individual stayed in a location for more than 15 min. The data were collected between June 2009 and February 2013 with an average recording period of 300 d, a maximum of 597 d and a minimum of 72 d per individual. For 48 individuals, deaths were recorded within the tracking period. We removed the data of one individual that was hit by a train, nursed for a while and then released again.

52 of all tagged individuals (28 females, 24 males) were also subject to an experimental feeding study, for which details can be found in Michel et al. (2017). We used the information to test whether experimental feeding is confounded with inter-sexual differences in movement behavior, but did not consider it further.

We used the data of all 126 individuals for the statistical habitat suitability model, but removed individuals with fewer than 50 recorded locations from the ABC calibration. The latter resulted in 96 individuals, of which 47 were females, 43 males, and 6 of unidentified sex.

*Individual-based little owl dispersal model*

To describe the IBM, we broadly follow the updated Overview, Design concepts, Detail (ODD) protocol

(Grimm et al. 2010), a standardized documentation format for agent-based models (ABMs). Fig. 2 presents a schematic workflow and visualizes the links between input variables, IBM and predictions. The model was implemented in C++ and interfaced to R (R Core Team 2018) in the LOSim package (see *Data Availability*).

*Purpose.*—This spatially explicit individual-based model (IBM) is designed to simulate the natal dispersal of little owls with the objective to predict the functional connectivity between expanding little owl populations and unoccupied habitat patches.

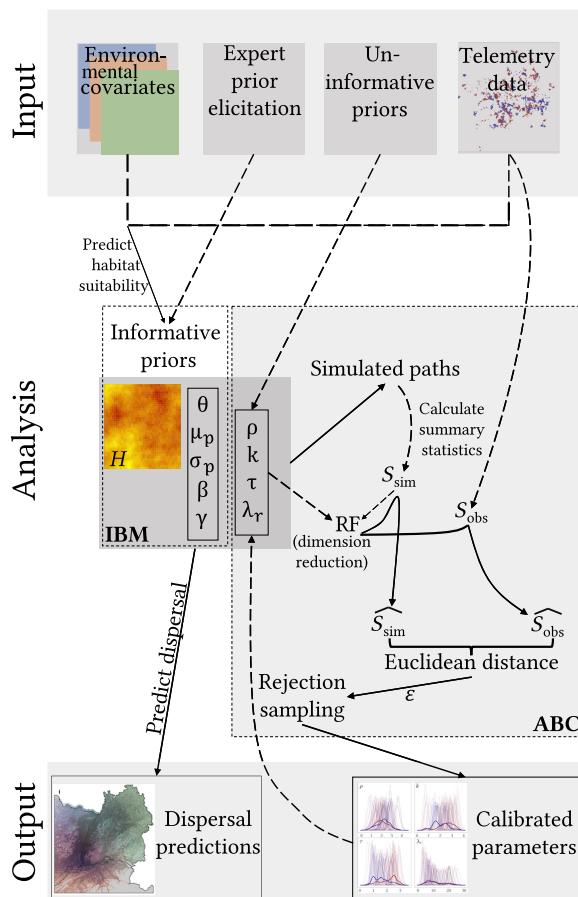


FIG. 2. Schematic workflow subdivided into input, analysis, and output. The analysis comprises the habitat suitability ( $H$ ) estimation as base layer for the individual-based model (IBM), the IBM, and its calibration through Approximate Bayesian Computation (ABC). The main parts of the latter are the simulation of little owl dispersal, dimension reduction by choosing appropriate summary statistics via random forest (RF) regression, and rejection sampling. Dashed arrows represent the inflow of variables (e.g., environmental covariates) or other measures (e.g., summary statistics) into any part of the analysis. Solid arrows display the respective output or prediction. Parameters are defined in Table 1.

*Scales, state variables, parameters.*—The model simulates a juvenile little owl’s transfer dispersal movements on a  $20 \times 20$  m habitat suitability ( $H$ ) grid (see *Input data* and Fig. 3a). Simulated time progresses on a continuous scale. The transfer phase (i.e., the simulated period) is assumed to last 60 d, based on population-aggregated visual inspection of the empirical data. The duration of this transfer phase likely varies between individuals, yet based on the scarce and irregular telemetry data, an individual-specific estimation was not possible. We believe, however, that this source of inter-individual variability is negligible in a population-based assessment of functional-connectivity. The state of each individual is characterized by its current location, time, daytime, and its current energy and activity status. Individuals differ in their start location and model parameters. Parameters influencing the simulated dispersal trajectory are the habitat-specialization level (importance weight of  $H$ ), preference of flight distances, directional persistence, duration of roosting, duration of perching (for the difference of roosting, day rest, and perching see *Process overview and scheduling*), and energy efficiency (see Table 1).

*Process overview and scheduling.*—In the IBM, we differentiate three modes of movement: active dispersal, day rest, and roosting. Active dispersal only happens between sunset and sunrise. During the day an individual always rests. When energy is low, it might roost for several days (including nights) and only starts dispersing again after sunset some days later. During active

dispersal, relocation decisions are modeled as a function of  $H$ , flight distance and turning angle (see Fig. 3b). First movements after a day-rest or roosting period are based on  $H$  and flight distance, but not turning angle.

Before any simulated movement, the model evaluates the individual’s activity status: active during the night, day rest during the day. In the latter case daytime skips forward to sunset and the owl becomes active, the elapsed time progresses accordingly. During active dispersal, an individual takes subsequent movement decisions. The time between these relocations (i.e., perching plus flight time) is determined by a log-normal distribution. After each relocation, the model evaluates the individual’s position and energy status, as well as elapsed time and daytime. If the latter passes sunrise the owl’s activity status switches to day rest, or if the energy budget is low to roosting, during which the energy budget rises again. These steps are repeated until the elapsed time reaches the 60 d limit or a predefined maximum number of relocations (set to 40,000, which was never reached in our simulations).

### Design concepts

*Basic principles.*—The model describes active little owl dispersal as a sequence of many short travel distances, interspersed by primarily short rest on perches. Fitting with the biology of the species, a relatively small wing area and high wing loading, the little owl relocates in relatively short flights, often from perch to perch. Hunting flights are typically of even shorter distance (Glutz von

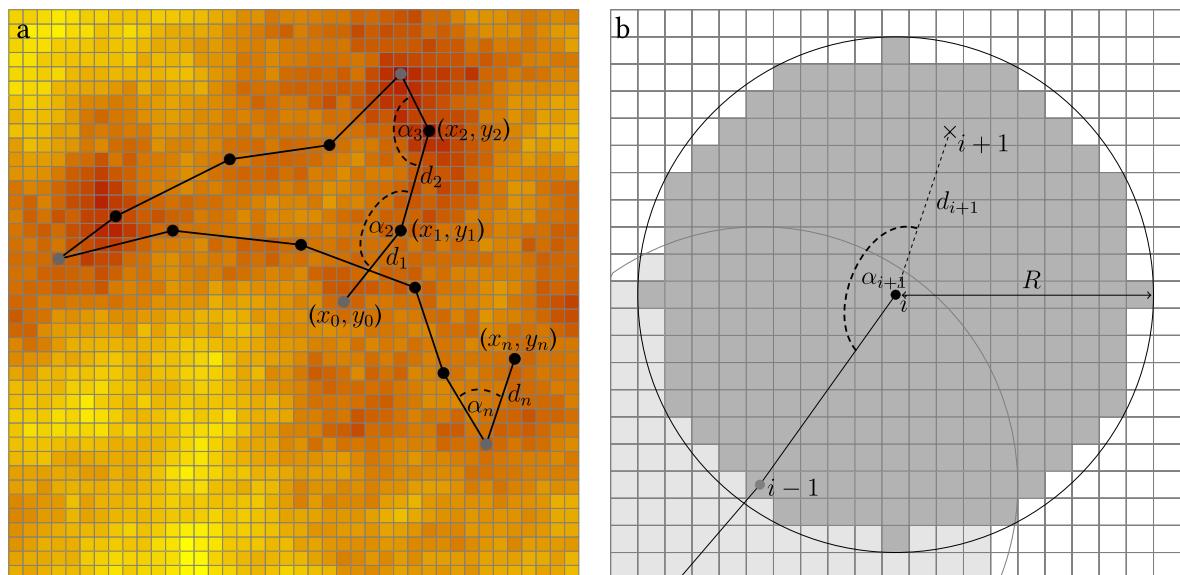


FIG. 3. (a) Overview of modeled movement on a simulated habitat suitability grid (higher suitability in red/dark colors): black dots represent simulated locations in dispersal mode, gray dots locations of roosting.  $\alpha$  is the turning angle corresponding to the previous step and  $(x,y)$  the location coordinates. (b) Concept figure illustrating relocation decisions: the radius of the circle(s) represents the perceptual range  $R$ . For all cells of the underlying habitat suitability grid with their centroid located within  $R$  (darker gray for current location  $(x,y)$ ), lighter gray for the previous location  $(x,y)_{i-1}$ ), weights are calculated based on weighted habitat suitability, distance  $d_{i+1}$ , and (if in dispersal mode) turning angle  $\alpha_{i+1}$ . The next location  $(x,y)_{i+1}$  is evaluated from a multinomial model based on these weights.

TABLE 1. Mathematical symbols, description and initial values of IBM parameters.

Parameter	Name and description	Initial value
$\rho$	Habitat-specialization level: weight defining the importance of habitat suitability in the selection of a new location	$\mathcal{U}(0.1, 5)$
$k$	Flight-distance shape: shape parameter in a gamma distribution describing the flight-distance preference	$\mathcal{U}(0.1, 4)$
$\theta$	Flight-distance scale: scale parameter in a gamma distribution describing the flight-distance preference	2
$\tau$	Directional persistence: inverse standard deviation in a Gaussian distribution (with $\mu = 0$ ) of the arc-cosine of turning angle $\alpha$	$\mathcal{U}(0.1, 3.5)$
$\mu_p$	Log-scale mean in a log-normal distribution describing perching duration in minutes	1.5
$\sigma_p$	Log-scale standard deviation in a log-normal distribution describing perching duration in minutes	0.7
$\beta$	Energy uptake efficiency: number of minutes to refill the energy budget completely, i.e., increase by 1	4,320
$\gamma$	Maximum effort: distance in meters to exhaust energy budget completely, i.e., reduce by 1	6,000
$\lambda_r$	Average roosting duration in days	$\mathcal{U}(0.01, 30)$

*Note:* Initial values are either given as values taken constant throughout all simulations informed by little owl experts or as a uniform prior distribution with relatively exhaustive ranges.

Blotzheim and Bauer 1980). To our knowledge, non-stop long-distance flights have not been documented for this species.

We assume in the model that an individual takes the next best location within a perceptual range. Generally, juvenile little owls disperse without preference for any specific cardinal direction (Van Nieuwenhuysen et al. 2008). Our field observations (B. Naef-Daenzer and M. U. Gruebler) suggest that an individual's dispersal path is segmented in various rather directional excursions (of multiple short flights), yet it seems that the direction between such directional excursions changes rather randomly.

The model implements several species-typical behavioral states, including resting and roosting, accounting for the species' circadian activity pattern. For the primarily crepuscular and nocturnal bird species, the activity phase (i.e., actual movement, including foraging and dispersal) typically begins 5 min before to 40 min after sunset and ends approximately 55 min before to 5 min after sunrise. During the day, little owls usually rest (Glutz von Blotzheim and Bauer 1980, Van Nieuwenhuysen et al. 2008).

*Emergence.*—Due to the randomness in movement decisions, the sole dispersal path of a simulated individual

shows low interpretability concerning the functional connectivity of the landscape. By generating such trajectories for a large number individuals, we project dispersal movement on the landscape and assess the functional connectivity as an emergent property of this model.

*Stochasticity.*—A simulated individual's movement decisions are based on weighted habitat suitability ( $H$ ), flight distance and directional persistence. The resulting preference values for the surrounding cells enter a multinomial distribution, which results in a stochastic decision about the next movement step. The degree of randomness in the simulated movement paths can be controlled by altering the model parameters. For example, a higher-weighted preference for  $H$  (i.e., higher habitat specialization) and flight distance, as well as an increased directional persistence decrease the degree of randomness.

Other sources of stochasticity are the duration of relocations and their corresponding stopovers at perches, and the number roosting days. The former are drawn from a log-normal distribution, the latter from a Poisson distribution.

*Observation.*—The recorded model outputs per individual include the coordinates and elapsed time for all simulated locations. For the calibration, we used the location and elapsed time only for the simulated steps whose elapsed times correspond to the elapsed times of the real observations. This procedure highlights the importance of accounting for the circadian rhythm of the modeled species in the IBM. Synchronizing both daytime and elapsed time improves the matching of movement modes in real and simulated observations. As the radio tracking via triangulation leads to some imprecision in the recorded locations, we added a Gaussian observation model on top of the model outputs, essentially adding additional stochasticity to the data that entered the ABC procedure. From the model output, we calculated a number of summary statistics for the ABC procedure, including several metrics derived from the obtained flight distances, turning angles, habitat suitability, as well as a resource selection coefficient (RSC):

$$\text{RSC} = \frac{N_{H_j > H_{\text{observed}}}}{N_{H_j < H_{\text{observed}}}} \quad (1)$$

where  $H_j$  is the habitat suitability of the cells  $j$  within a 200 m radius around the observed cell,  $H_{\text{observed}}$  the habitat suitability of the currently occupied cell, and  $N$  the number of cells where  $H_j$  is larger (numerator) or smaller (denominator) than  $H_{\text{observed}}$ . For a list of all calculated summary statistics, see Appendix S3.

*Initialization.*—For the model calibration, starting conditions (location, daytime, day of year) were determined for each individual by the time and location from which it began its first excursion further than 400 m away from

its fledging site (i.e., initiation of dispersal-like movement). For the predictions from the calibrated model, we re-sampled the start locations from ringing records of juvenile little owls between 2011 and 2015 (Vogelwarte Radolfzell 2016). Initial daytimes and day of the year were randomly sampled from between 00:00 and 24:00 and 205 to 352, respectively.

The model parameters habitat-specialization level ( $\rho$ ), flight-distance shape ( $k$ ), directional persistence ( $\tau$ ), and average roosting duration ( $\lambda_r$ ) were calibrated on the observed radio telemetry data (see *IBM calibration*). For predictions, their values were drawn randomly from the multivariate ABC posterior distribution. The parameters flight-distance scale ( $\theta$ ), log-scale mean ( $\mu_p$ ) and standard deviation ( $\sigma_p$ ) of perching duration, and the energy efficiency parameters  $\beta$  and  $\delta$  were kept constant for all simulations. Their values were derived from three independent little owl field experts, who were first familiarized with the behavioral processes and the effect of the respective parameter of interest and then asked for their best guess for the respective parameter value. We used the mean per parameter, which are presented in Table 1. The meaning and implication of all parameters is explained in *Submodels*. The mean and standard deviation of the Gaussian observation models used the recorded accuracy estimates of the radio telemetry data.

*Input data.*—The only external input of the model is the habitat suitability base grid, which was estimated from all recorded radio telemetry locations (i.e., not restricted to the transfer phase), as a function of a number of a priori selected environmental predictors (see Fig. 2). The relationship was modeled using a generalized additive model (GAM; Hastie and Tibshirani 1990).

Because of the highly irregular sampling interval in the radio telemetry data, which prevents us from using use/availability regressions, such as resource and step selection functions, we computed the utilization distribution (UD; Van Winkle 1975) for each owl individually within the extent of the recorded locations to use it as response variable. We estimated the UD using the biased random bridges kernel method (BRB; Benhamou and Corn elis 2010, Benhamou 2011), which incorporates local movement information “provided by serial correla-

individual and by UD. Note that this sample is based on 17,801 recorded telemetry locations.

As predictors of habitat suitability we considered the availability of breeding cavities and high-quality food resources as well as the presence of predators and human-built structures, which are factors known to affect the occurrence of little owls (e.g., Van Nieuwenhuysse et al. 2008, Gottschalk et al. 2011, Michel et al. 2016). We thus included land cover as a factor, while elevation and the distances to the nearest forest patches, nesting boxes, urban settlements, and major roads were fitted as thin plate regression splines with an additional shrinkage penalty to reduce over-fitting and increase predictive power (Wood 2017). To account for inter-individual variability, we fitted a random intercept for the individual owls. For GIS processing and references for the environmental data, see Appendix S1.

### Submodels

*Activity status evaluation.*—At each iteration  $i$ , the model evaluates the current activity status,  $A_i$ , of an individual. Because this decision is based on daytime, the model constantly keeps track of the time of day (tod) and the location-specific and day-of-year-specific times of sunrise and sunset. The latter are corrected for the slight shift in the little owl’s response to sunrise and sunset, such that  $s_0 = \text{sunrise} - 25 \text{ min}$  and  $s_1 = \text{sunset} + 25 \text{ min}$ .  $A$  is classified in active and inactive, and evaluated so that

$$A_i = \begin{cases} \text{active,} & \text{if } s_{0,i} \leq \text{tod}_i \leq s_{1,i} \\ \text{inactive,} & \text{otherwise.} \end{cases} \quad (2)$$

In inactive movement mode, we further consider the individual’s current energy budget,  $E_i$ , and the summed up habitat suitability,  $\sum H$ , within its perceptual range in a Bernoulli model to determine whether the owl only rests for the one day (i.e., higher energy/lower habitat suitability; resulting in  $t_r = 1$ ), or whether it roosts (i.e., lower energy/higher habitat suitability), and if so, we draw the number of roosting days,  $t_r$  from a Poisson distribution:

$$\text{if } A_i = \text{inactive; } t_r = \begin{cases} 2 + x, & \text{if Bern}(p = \sum H + 1 - E_i); \text{ where } x \sim \text{Pois}(\lambda_r) \\ 1, & \text{otherwise,} \end{cases} \quad (3)$$

tion between successive relocations” (Benhamou 2011:1), and thus prevents overestimation of generally avoided regions. Thereafter we standardized the estimated UD subsequently, so that its sum was 1 over the extent of the study area. From the standardized UD grid, we sampled 50,657 locations stratified by

where  $\lambda_r$  represents the average roosting duration. During the time of roosting the individual stays at the current location.

*Movement decisions.*—In active movement mode, the individual relocates within a perceptual range,  $R$ , which

is defined as the rounded up 95% quantile of a gamma cumulative distribution function:

$$R = \left\lceil \inf \left\{ x \in (0, \infty) : 0.95 \leq \frac{1}{\Gamma(k)} \gamma \left( k, \frac{x}{\theta} \right) \right\} \right\rceil, \quad (4)$$

where  $k$  and  $\theta$  represent the shape and scale parameter of the gamma distribution. We chose this distribution as it provides sufficient flexibility to accommodate substantially different distance preferences among individual little owls.

For each cell within a circle around the current location  $(x_i, y_i)$  of radius  $R$ , a selection weight,  $\omega_j$ , is calculated based on the cell specific habitat suitability,  $H(x_j, y_j)$ , distance  $d_j$  to  $(x_i, y_i)$ , and, if the previous step was also marked as active dispersal movement (i.e.,  $A_{i-1} = \text{active}$ ), turning angle  $\alpha_j$ . Thus

$$\omega_j = \begin{cases} H(x_j, y_j)^{\rho-1} \frac{1}{\Gamma(k)\theta^k} d_j^{k-1} e^{-\frac{d_j}{\theta}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\arccos(\alpha_j)^2}{2\sigma^2}} & \text{if } A_{i-1} = \text{active} \\ H(x_j, y_j)^{\rho-1} \frac{1}{\Gamma(k)\theta^k} d_j^{k-1} e^{-\frac{d_j}{\theta}} & \text{otherwise,} \end{cases} \quad (5)$$

where  $j$  is the index of the respective cell within perceptual range,  $R$ . In other words, selection weight,  $\omega_j$ , is the product of (1) habitat suitability weighed by habitat-specialization level  $\rho$ , (2) the probability density of a gamma distribution (see Eq. 4) for the flight distance,  $d_j$ , between  $x_j, y_j$  and  $x_i, y_i$ , and, if  $A_{i-1} = \text{active}$ , (3) the probability density of a Gaussian distribution (with  $\mu = 0$  and  $\sigma = \frac{1}{\tau}$ , where  $\tau$  represents directional persistence) for the arc-cosine of turning angle,  $\alpha_j$ . The latter is defined as the angle between potential location  $x_j, y_j$ , current location  $x_i, y_i$  and previous location  $x_{i-1}, y_{i-1}$ .

Through model parameter  $\rho$  we introduce an individual-specific weight of the importance of habitat suitability, such that each simulated individual can display different habitat-specialization levels. Individuals can also differ in their preference for flight distance, which is controlled by  $k$  and  $\theta$ , and develop a certain degree of directional persistence through  $\tau$ .

During active dispersal, time progresses by random draws from a log-normal distribution, such that

$$t_{i+1} = t_i + x; \quad \text{where } x \sim \log\mathcal{N}(\mu_p, \sigma_p). \quad (6)$$

$\mu_p$  and  $\sigma_p$  represent the log-scale mean and standard deviation, respectively, of perching time. Daytime and, if necessary, the day of the year are altered accordingly.

**Energy budget.**—An individuals' energy budget,  $E_i$ , is updated based on a simple principle: during active dispersal the energy reduces linearly with traveled distance. The amount lost per unit distance depends on the efficiency parameter,  $\delta$ .  $E_i$  increases only in inactive

movement modes (i.e., day rest or roosting). The amount of uptake depends on the habitat suitability within perceptual range,  $R$ , such that

$$E_{i+1} = \begin{cases} E_i - \frac{d_{i+1}}{\delta}, & \text{if } A_i = \text{active} \\ E_i + \sum H \frac{(t_{i+1}-t_i)}{\beta}, & \text{otherwise,} \end{cases} \quad (7)$$

where  $\delta$  represents the distance after which the energy budget,  $E_i$ , is emptied completely and  $\beta$  is the number of minutes it takes to increase  $E_i$  by 1.  $\sum H$  represents the summed up habitat suitability within  $R$ . We define  $E \in (0, 1)$ . Should these limits be exceeded,  $E_i$  is set to 0 or 1, respectively.

### IBM calibration

We calibrated the model parameters, habitat-specialization level ( $\rho$ ), preference for flight distance

( $k$ ), directional persistence ( $\tau$ ) and average roosting duration ( $\lambda_r$ ), using Approximate Bayesian Computation (ABC; Beaumont et al. 2002, Beaumont 2010, Csilléry et al. 2010). The general idea of ABC is to approximate the likelihood of simulation models by noting that the probability that simulated data from the model  $D_{\text{sim}}$ , are close to the observed data,  $D_{\text{obs}}$ , is approximately equal to  $p(D_{\text{obs}}|\phi)$ , the likelihood of obtaining  $D_{\text{obs}}$  from the model given the parameters  $\phi$ . This insight can then be used in other Bayesian sampling algorithms, such as rejection sampling, Markov chain Monte Carlo (MCMC), or Sequential Monte Carlo (SMC) to approximate the likelihood or posterior distribution for a given model. The computational efficiency of the approach can be improved by comparing the simulations and observed data based on summary statistics. An appropriate choice of summary statistics is key for the efficiency of the method, and much recent research has been devoted to this topic (Blum et al. 2013, Prangle 2018). Here, we aggregated the information in the calculated (raw) summary statistics using an intermediate random forest (RF) model. RF is a flexible machine learning algorithm, which is particularly useful to predict highly non-linear, complex relationships and thus ideal for dimension reduction of multidimensional summary statistics. On the basis of the estimated summary summary statistics we employ ABC rejection sampling (Tavaré et al. 1997, see also pseudocode in Hartig et al. 2011). To test the performance of RF dimension reduction, we also use least squares regression as a methodological

benchmark to estimate summary statistics (see Fearnhead and Prangle 2012).

For the ABC rejection procedure (see Fig. 2), we first sampled  $n_p = 500,000$  parameter values from wide uniform prior distributions, such that  $\rho \sim \mathcal{U}(0.1, 5)$ ,  $k \sim \mathcal{U}(0.1, 4)$ ,  $\tau \sim \mathcal{U}(0.1, 3.5)$  and  $\lambda_r \sim \mathcal{U}(0.01, 30)$ . For each little owl individual, we then simulated dispersal paths for all  $n_p$  parameter combinations, with the remaining model parameters kept constant (see Table 1). For both simulations and observations we calculated 49 summary statistics ( $S$ ). The calculated metrics comprise, for example, the mean and standard deviation of all habitat suitability values at (simulated) observed locations, the distance between start and end location, the convex hull area of the full dispersal path, and the 10%, 20%, 80% and 90% quantile of the recorded turning angles. For a list of all calculated summary statistics, see Appendix S3.

To aggregate the information in the summary statistics, we performed dimension reduction using individual-specific intermediate RF models. The idea of this approach is similar to the partial least squares regression method by Wegmann et al. (2009): we used RF to predict  $\rho$  (habitat-specialization level),  $k$  (preference for flight distance),  $\tau$  (directional persistence), and  $\lambda_r$  (average roosting duration) as a function of the 49 summary statistics calculated from the simulated dispersal paths ( $S_{\text{sim}}$ ). To avoid correlations among the estimated summary statistics, we successively included the predicted values as additional predictors, such that

$$\begin{aligned} \rho &\sim S_{\text{sim},1} + S_{\text{sim},2} + \dots + S_{\text{sim},49} \\ k &\sim S_{\text{sim},1} + S_{\text{sim},2} + \dots + S_{\text{sim},49} + \hat{\rho} \\ \tau &\sim S_{\text{sim},1} + S_{\text{sim},2} + \dots + S_{\text{sim},49} + \hat{\rho} + \hat{k} \\ \lambda_r &\sim S_{\text{sim},1} + S_{\text{sim},2} + \dots + S_{\text{sim},49} + \hat{\rho} + \hat{k} + \hat{\tau}. \end{aligned} \quad (8)$$

We fitted models for each individual separately, using the ranger R package (Wright and Ziegler 2017), and thereafter used the fitted RF models to calculate four reduced summary statistics from the previously 49 summary statistics for simulated ( $S_{\text{sim}}$ ) and observed data ( $S_{\text{obs}}$ ).

For the (reduced) four-dimensional space, we compared the estimated summary statistics of the simulations,  $\widehat{S}_{\text{sim}}$ , and those of the observations,  $\widehat{S}_{\text{obs}}$  via calculating the Euclidean distance, standardized to the range of the respective summary statistics in prior space. To approximate the IBM posterior distribution, we chose the cut-off value,  $\epsilon$ , so that 1,000 of the 500,000 parameter combinations were accepted. We reduced the problem of an ABC posterior biased towards the prior using post-sampling regression adjustment (Blum 2010).

*Validation of the ABC approach.*—The quality of the ABC approximation hinges on the sufficiency of the

used summary statistics. Because sufficiency can usually not be analytically predicted, it is common practice to test the approach on simulated data before using it on real data. For this, we randomly sampled 1,000 from the 500,000 simulations and applied ABC rejection sampling and post-sampling regression adjustment on them, treating them as observations. We repeated this procedure with a second dimension-reduction technique for the aggregation of information in the calculated summary statistics: least squares regression as part of the semi-automatic ABC approach (Fearnhead and Prangle 2012), implemented in the abctools R package's saABC function (Nunes and Prangle 2015). The estimates obtained from this approach were not considered further in the little owl study; they merely served as a reference to the aggregated summary statistics derived from RF regression.

### *Predicting functional connectivity*

From the calibrated IBM, we generated natal dispersal movements of little owls from the three potential source populations in southwestern Germany at Neckarbecken, Kaiserstuhl, and Lörrach (from north to south). As start locations for the first simulated dispersing generation, we resampled the coordinates of juvenile little owl ringing records from the years 2011 to 2015, which were assigned to the respective source population (see color scheme in Fig. 1a; Vogelwarte Radolfzell 2016). To achieve a stable connectivity map, we drew 5,000 locations from each of the three populations. For these 15,000 individuals, we randomly sampled parameter values from the ABC posterior distribution to simulate 60-d dispersal paths from the IBM. The end locations of all individuals served as start locations for the next generation. The parameter values were sampled new for each simulated generation. This resulted in dispersal trajectories of 55 simulated generations of 5,000 simulated individuals per potential source population. These were converted into densities counting the number of locations within  $500 \times 500$  m cells.

Note: the number of simulated juveniles (5,000) per generation presumably exceeds the number of fledged offspring substantially. We simulated dispersal for such a large number to obtain a diffusion map for the analysis of functional connectivity. Also note: we only simulated movement during the transfer phase of dispersal while neglecting other demographic processes entirely. Therefore the time of these simulated generations should only be interpreted as a relative measure.

## RESULTS

### *Validation of the calibration method*

To validate the applied calibration method, we first considered the inferential capacity of the summary statistics obtained by random forest (RF) regression, that is, how well they can describe the respective

parameter of interest. We found that the estimated RF summary statistics ( $\widehat{S}_1 - \widehat{S}_4$ ) were highly correlated with the respective parameter values (Pearson's  $r^2 > 0.95$  for all observed little owl individuals), and virtually uncorrelated among themselves (see Appendix S4: Fig. S1a). Both these results indicate that the estimated statistics describe the simulated movement trajectories sufficiently well to discriminate the individual model parameters.

For methodological comparison, we applied semi-automatic ABC (Fearnhead and Prangle 2012), which uses least squares regression to aggregate the calculated summary statistics. We found that parameter samples and respective estimated summary statistics were substantially less correlated. Furthermore,  $\widehat{S}_1$ ,  $\widehat{S}_2$ ,  $\widehat{S}_3$  and  $\widehat{S}_4$  partially correlated with each other (see Appendix S4: Fig. S1b), which reduces the ability to discriminate between the individual model parameters.

We proceeded to test the ability of the full ABC method (using ABC rejection sampling) to retrieve parameter estimates from data created by the model with known parameter combinations. We found that ABC using RF to estimate summary statistics produced posterior estimates close to the respective true parameter values (as indicated by the red identity line in Fig. 4a). In particular, the average roosting duration ( $\lambda_r$ ) could be fitted with high precision. The posterior distributions of habitat-specialization level ( $\rho$ ), preference for flight distance ( $k$ ), and directional persistence ( $\tau$ ) were slightly shifted toward lower and upper values. The obtained

80% ABC posterior range was rather narrow for this aggregation method. While the medians of the posterior distributions derived from ABC rejection sampling based on summary statistics aggregated by least squares regression were not dramatically worse, the 80% ABC posterior credible intervals were substantially wider (see Fig. 4b).

#### Estimated movement parameters of little owl dispersal

The marginal posterior estimates for all four calibrated IBM parameters, habitat-specialization level ( $\rho$ ), preference for flight distance ( $k$ ), directional persistence ( $\tau$ ), and average roosting duration ( $\lambda_r$ ), showed substantial inter-individual variability (see Fig. 5). For flight distance and directional persistence, distinct sex differences were recognizable, with the posteriors of females and males scattering around sex-specific mean values. With a posterior median of  $\widetilde{k}_f = 2.53$  (80% CI: 1.26–3.28, which translates to a maximum probability for a flight distance of 61 m; 80% CI: 10–91 m) females showed preferences for substantially greater flight distances than males with  $\widetilde{k}_m = 1.93$  (80% CI: 1.29–2.85, i.e., a maximum probability for a flight distance of 37 m; 80% CI: 11–74 m). Female little owls also displayed higher directional persistence between single flights ( $\widetilde{\tau}_f = 2.26$ ; 80% CI: 1.12–1.82 and  $\widetilde{\tau}_m = 1.55$ ; 80% CI: 0.93–2.45). These values mean that at the median posterior value, females consider in 80% of the cases locations within  $33^\circ$  around the direction of the previous flight (80% CI:  $65\text{--}26^\circ$ ). Males are less directionally persistent

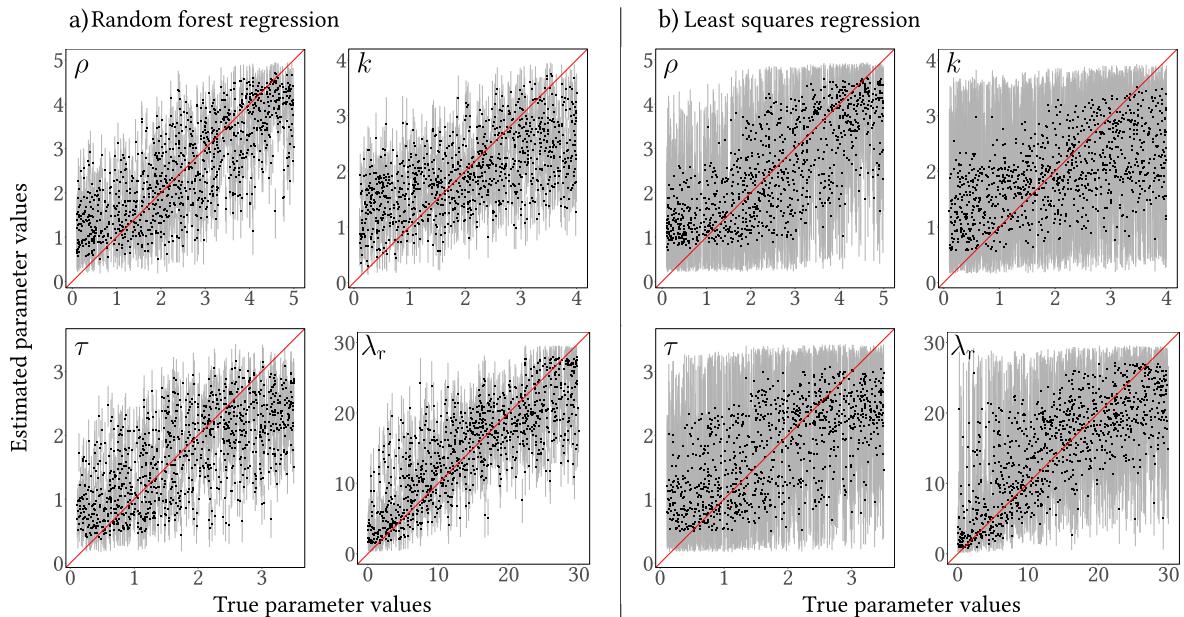


FIG. 4. Estimated parameter values against true parameter values for the two summary statistics dimension-reduction approaches: (a) random forest and (b) least squares regression, and for the four ABC calibrated model parameters: habitat-specialization level ( $\rho$ ; top left), preference for flight distance ( $k$ ; top right), directional persistence ( $\tau$ ; bottom left), and average roosting duration ( $\lambda_r$ ; bottom right), respectively. Black dots represent the median of the respective ABC posterior, gray bars 80% CIs and red diagonals represent identity of the estimated and true parameter values.

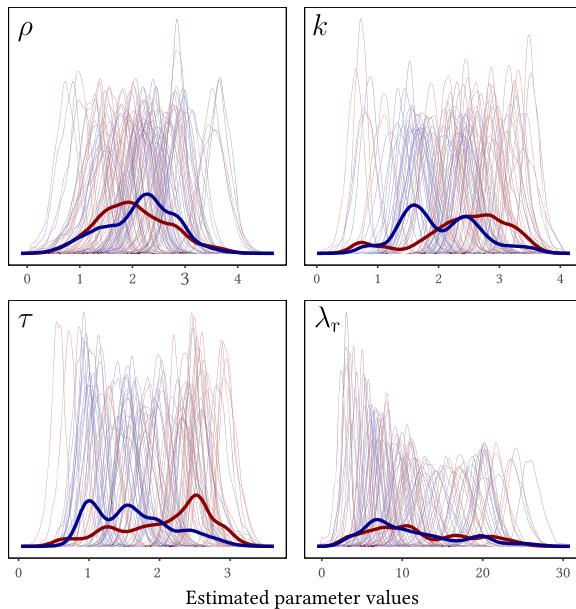


FIG. 5. Estimated posterior densities for the ABC-calibrated IBM parameters: habitat-specialization level ( $\rho$ ; top left), preference for flight distance ( $k$ ; top right), directional persistence ( $\tau$ ; bottom left), and average roosting duration ( $\lambda_r$ ; bottom right). Thin lines represent individual ABC posterior distributions, thick lines represent ABC posterior distributions pooled by sex (red for females, blue for males), illustrating the sex differences in movement behavior.

and choose in 80% of the cases from locations within  $47^\circ$  around the direction of the previous flight at their posterior median (80% CI:  $79\text{--}30^\circ$ ). The estimated average roosting durations were only slightly higher for female juvenile little owls ( $\lambda_{r,f} = 10.75$  days and  $\lambda_{r,m} = 9.70$  days). In contrast, males displayed moderately higher habitat-specialization levels ( $\tilde{\rho}_m = 2.22$ ; 80% CI: 1.19–2.96 and  $\tilde{\rho}_f = 1.97$ ; 80% CI: 1.16–2.94). That is to say, the importance of habitat suitability in the process of selecting a new location is raised to the power of 8.19 (80% CI: 2.29–18.34) for male and 6.20 (80% CI: 2.18–17.98) for female little owls.

#### *Predicted functional connectivity*

The predicted functional connectivity after the longest simulated time (55 simulated generations, 5,000 individuals, three source populations) was non-zero in most parts of the underlying landscape, except for the alpine land in the very south of the map (Fig. 6, right panel), but also showed substantial spatial heterogeneity. As expected, the highest densities were found in and close to the area of the source populations, but also some of the valleys into northern Switzerland showed moderately high densities. Good predicted connectivity was also found in the Neckarbecken and along the northern part of the Swabian Alb (compare Fig. 6f, g). Not surprisingly, areas of low estimated habitat suitability also

showed low predicted connectivity (see Appendix S1); these are primarily higher altitudes including the Black Forest in southwestern Germany, the foothills of the Jura in northwestern Switzerland, and the Vosges Mountains in northeastern France.

We also found that simulations from Lörrach (red) showed an increased north-south dispersion with considerably more individuals dispersing northward, away from Switzerland. Highest densities were found on the eastern side of the Rhine. Simulated dispersers from the Kaiserstuhl population (blue) showed the same north-south movements within the Rhine valley, also with higher densities on the eastern side of the Rhine valley. Simulated dispersal movements of individuals from the Neckarbecken population (green) were highest in the Neckarbecken before the movements extended east towards the limit of the extended study area. Westward dispersal movement seemed to be limited. The overlap of simulated dispersal movement was most prominent between the populations at Lörrach and Kaiserstuhl, while the connectivity between the two population in the Rhine valley and the one at the Neckarbecken seemed limited. Compared to the movement density within the Rhine valley, dispersion into northern Switzerland is low. This was particularly true for the first seven simulated generations, where only few individuals started crossing the Northern hills of the Jura Mountains. Only from generation 20 on simulated movements into northern Switzerland increased.

#### DISCUSSION

In this study, we tested a new approach for calibrating an individual-based model of natal dispersal in little owls, comprising a combination of likelihood-based and approximate model parametrization and using random forest (RF) regression for more efficient ABC rejection sampling. We validated the proposed RF-enhanced ABC calibration internally and compared its inferential capacity to estimates derived from summary statistics chosen by least squares regression. Finally, we used the fitted IBM as an inferential tool, to compare movement differences between males and females, and to project dispersal movement into southwestern Germany, northern Switzerland, and northeastern France for an analysis of functional connectivity.

From a methodological perspective, our results indicate that the hybrid of statistical methods and ABC for IBM parameterization, combined with an RF-based selection of the summary statistics in ABC worked well for calibrating a relatively complex movement model to highly irregular radio telemetry data. Using a statistical model for estimation of habitat preference parameters reduced the number of parameters to calibrate via ABC substantially, and thus facilitated the basis for realistic dispersal simulations. Furthermore, RF clearly outperformed alternative approaches for choosing ABC

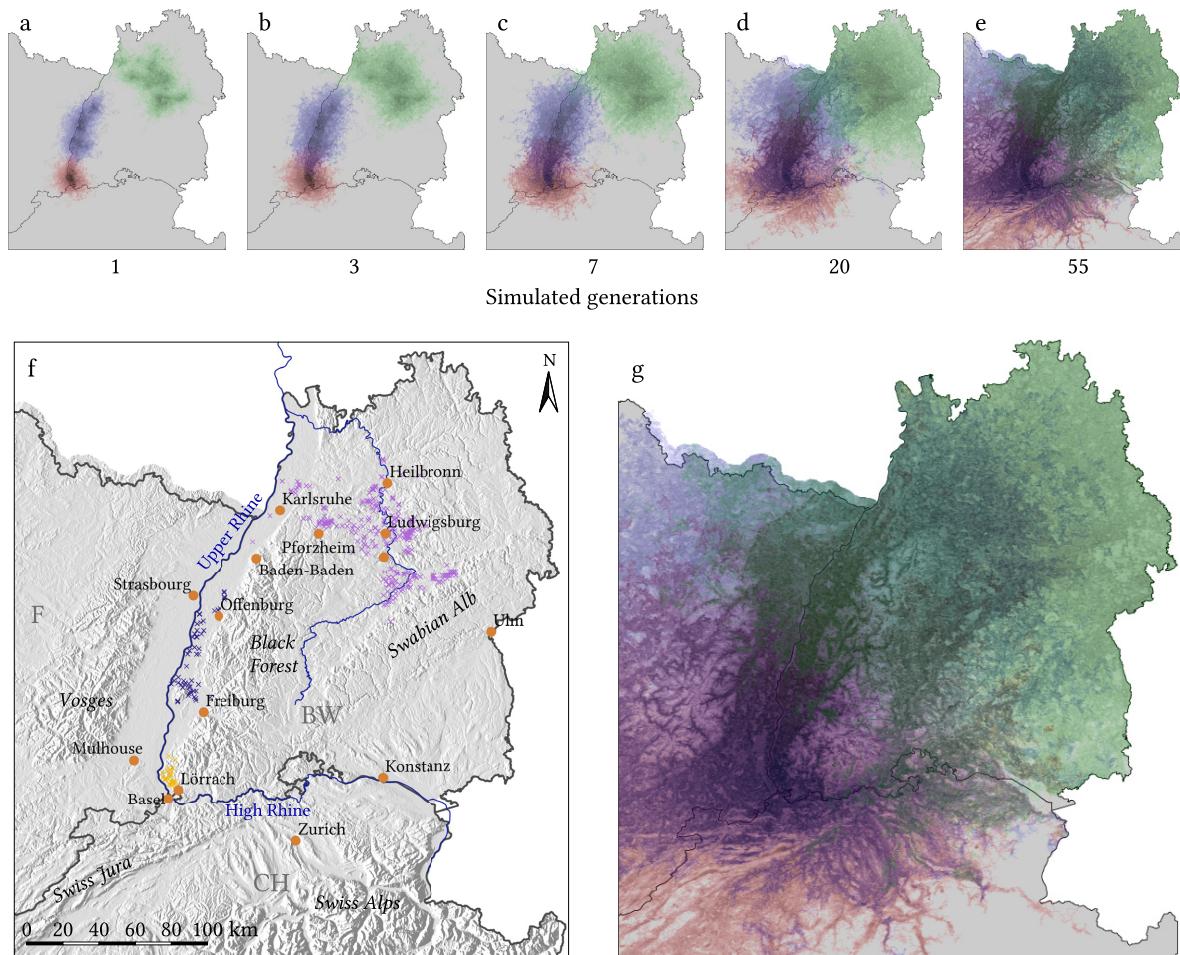


FIG. 6. (a–e) Dispersal density estimated from a total of 15,000 individuals after 1, 3, 7, 20, and 55 simulated dispersal periods. Purple, blue, and yellow each display the density of 5,000 simulated individuals from potential source populations at the Neckar-becken, Kaiserstuhl, and Lörrach, respectively. (f) Terrain map illustrating landscape features for comparison with predicted functional connectivity and initial start locations represented by colored crosses. (g) Same as panel e; predicted functional connectivity after 55 simulated generations.

summary statistics leading to overall better posterior estimates.

From an ecological perspective, the fitted model revealed clear differences between the dispersal movement behavior of male and female little owls, with females showing longer and more directed dispersal, during which they exhibited lower habitat specialization than males. Projections of the model to assess functional connectivity demonstrated that a (re-)colonization of northern Switzerland seems generally possible. Yet, the simulations also indicated a rather slow population expansion, limiting the ability to re-colonize in the near future.

#### *Hybrid calibration: combining statistical inference with ABC*

Modeling the movement of animals is challenging as it requires an adequate representation of the often

complex stochastic movement process to account for all factors influencing relocation decisions (Nathan et al. 2008). Both the complexity required in a movement model and the stochasticity in the simulations render the estimation of model parameters a difficult undertaking.

For the problem of calibrating ecological models whose likelihood function is analytically intractable, Approximate Bayesian Computation (ABC) has become a widely applied solution (e.g., Kosmala et al. 2016, Marchand et al. 2017). Yet, for complex models, the calibration of all parameters via ABC can become impractical as the number of required simulations often exceeds the costs that are computationally feasible (e.g., Boyd et al. 2018). Many studies avoid these issues through prior elicitation or by deriving estimates from experts or literature instead of using data-driven approaches to inform some or even all of the model parameters (e.g., Arrignon et al. 2007, Graf et al. 2007, Pe'er and Kramer-Schadt 2008). Our study presents a hybrid

calibration approach as an improved means to disentangle parameterization of environmental processes (i.e., habitat suitability) and behavioral processes, which, provided that the environmental influence on a species' or an individual's movement is multi-factorial, greatly reduces the remaining number of parameters to calibrate via ABC. Specifically, we use statistical, likelihood-based methods to estimate the model parameters describing the influence of several environmental factors on movement, and subsequently ABC to estimate some of the parameters describing movement behavior. For only a few parameters we used prior elicitation. From a Bayesian's perspective, this approach might be interpreted as using the estimated habitat suitability as an informative prior for the environmental influence on the dispersal movement of the little owl. Certainly, a prerequisite for this approach to work is the possibility of using statistical estimation techniques for a subset of the IBM parameters (e.g., parameters describing habitat preferences).

For species that exhibit substantial inter-individual variability in the way they respond to their environments, it is important to allow for such variability in the IBM, because the estimated habitat suitability (i.e., estimated means in form of a map) cannot convey this information in a computationally feasible way. Our approach conserves this variability by implementing the habitat-specialization-level parameter: a weight defining the importance of habitat suitability in the selection of a new location, which can be different for each observed individual. It is worth noting that parameter values chosen by prior elicitation cannot conserve such inter-individual variability and, because they are guesses based on field observations, introduce some uncertainty not accounted for in the model calibration.

#### *Using random forest for choosing summary statistics in ABC*

To approximate a posterior distribution via ABC, minimal sufficiency of the summary statistics is essential, which means, on the one hand, that the calculated summary statistics must “produce an aggregation of the [simulated] data that contains the same information as the original data [...]” for the purpose of parameter inference (Hartig et al. 2011:4), and on the other hand that their dimensionality should be as low as possible.

The problem of choosing the optimal summary statistics has become a new field of research. Wegmann et al. (2009), for example, suggested employing a partial least squares regression for choosing dimensionally reduced summary statistics. A readily available implementation of a similar approach can be found in the semi-automatic ABC, which also uses an additional estimation stage of the optimal summary statistics (least squares regression). This method was proposed by Fearnhead and Prangle (2012), who describe it as robust way to

choose the summary statistics. Both approaches, however, assume linear relationships between summary statistics and the parameter sample. Instead of aggregating the information of many “simple” summary statistics, Zhang et al. (2017) used statistical inferential movement models (IMMs) to produce fewer, but more informative summary statistics in the first place. Since the predictive capacity of the methods employed to reduce the dimensionality of calculated summary statistics is the sole purpose, algorithmic models, such as decision trees or neural nets are likely to be advantageous over statistical models (or “data models” as referred to in Breiman 2001b).

In this study, we have shown that random forest (RF) regression (Breiman 2001a) can reliably detect the often non-linear relationships between the summary statistics and the parameter sample. In a cross-validation using simulated data, we have shown that RF produces posterior estimates close to the true parameter values with high precision (i.e., narrow CIs). Least squares regression showed dramatically wider CIs. In such an analysis, this observation clearly signals a higher quality of the summary statistics derived from RF regression. As already noted by Saulnier et al. (2017), dimension reduction via RF is also robust against correlated raw summary statistics, which makes their selection a less elaborate task. We believe that we are the first to employ this technique for dimension reduction in ABC rejection sampling, although Raynal et al. (2017) used it to directly estimate model parameter values from the parameter sample. Whilst useful, their approach does not approximate the parameter posterior distribution. This becomes disadvantageous, if model parameters are correlated or do not exhibit a unimodal posterior distribution.

#### *Sex differences in little owl dispersal movement*

Female juvenile little owls are known to disperse farther from their parental territory than male conspecifics (see Glutz von Blotzheim and Bauer 1980, Van Nieuwenhuyse et al. 2008). In a study of 28 ringed male juveniles and 30 ringed female juveniles, Kämpfer and Lederer (1988) showed that the mean dispersal distance is 2.3 km for males and 6.3 km for females. The maximum dispersal distance of females (up to 100 km) exceeds that of males (~ 20 km) by far. Yet, these maxima represent rather isolated incidences. The individual specific movement characteristics analyzed in our study comprise individual-specific habitat-specialization levels, flight-distance preferences, directional persistence, and roosting duration, all of which vary prominently among the observed individuals. Nonetheless, female juvenile little owls tend to have lower habitat specialization, are highly persistent in their flight direction and exhibit substantially longer individual flights. This behavioral pattern of females coincides with their tendency to disperse farther away from the parental territory.

One might interpret these inter-sexual differences in the obtained parameter estimates as a mere consequence of the spatial arrangement of the telemetry data (i.e., males ending up close to the parental territory, females dispersing substantially further). Yet, if this were the case, we would expect a substantially shorter mean roosting duration of female juveniles, as this would have been a parsimonious way to achieving longer overall dispersal distances in females. We also tested the correlation of body mass (which was measured while tagging the individuals) and whether individuals were artificially fed or not between the 14th and the 50th day after hatching with the respective posterior medians to rule out potential confounding variables (for the latter, see Appendix S5: Fig. S1). It seems that artificially fed individuals showed no different movement behavior, except for shorter roosting duration. As we did not find any correlation for any of the parameters with the juveniles' body mass, we consider it likely that the observed patterns in the posterior distributions reflect differences in behavioral traits between males and females.

#### *Little owl dispersal connectivity*

Our analysis of functional connectivity suggests that the habitat patches in northern Switzerland, which are currently unoccupied (Knaus et al. 2018), are theoretically connected at least to the two populations at Kaiserstuhl and Lörrach. At the same time, the surrounding topographic conditions constrain the movement of juvenile little owls. Mountainous areas, such as the Black Forest or the foothills of the Jura mountains, exhibit a clear barrier effect, resulting in decreased functional connectivity in northern Switzerland compared to areas of similar distance within the Rhine valley. Furthermore, the diffusion map (Fig. 6) shows that dispersal movement into Switzerland occurs in rather constrained paths, which identify potential dispersal corridors.

We also find the three little owl populations in southwestern Germany to be likely functionally connected. A regular exchange between these populations has been reported by Fiedler (2013). The same study also showed that there is more exchange between the two populations located along the Upper Rhine valley. This seems consequential not only because of the distance between the two populations in the Rhine valley, but also because their direct paths to the population in the Neckarbecken are constrained by the mountainous Black Forest and Swabian Alb. Potential linkage between the populations in the Neckarbecken and at the Kaiserstuhl seems most probable along the valleys between Baden-Baden and Rastatt in the Rhine valley and the city of Pforzheim through the towns Gaggenau, Gernsbach and Bad Herrenalb.

The functional connectivity with the Swiss Plateau seems to be limited, although a few ring recoveries from little owls of one of the three German populations have been recorded there in the last 20 yr. Dispersing

individuals might reach the areas south of Basel from the German upper Rhine Valley through the edges of the city of Basel. However, the potential corridors from the German Rhine Valley to the central Swiss Plateau might be redirected eastwards along the Rhine River towards Lake Constance. This redirection happens because of the hilly area of the Swiss Jura Mountains, consisting of a mosaic of forests and agricultural fields, acting as a barrier toward the south. The first potential connections to the south occur either around Brugg through the valley of Frick or through the lower Aare Valley. Within the Swiss Plateau, the potential dispersal corridors mainly follow the valleys of the large rivers of Aare, Reuss, Rhine and Thur.

Beyond a simple habitat suitability model (e.g., Fatterbert et al. 2018) that can serve as a basis to identify locations for habitat restoration (e.g., restoring orchards) or nest box placement to improve micro-habitat quality and support breeding, the presented IBM can help optimize interventions to promote dispersal and connectivity either by developing strategies for targeted restoration in areas of connectivity-potential or clear gaps; and these could be tested for their potential by model-experimentation. Because of the diffusing nature of little owl dispersal, nesting-placement strategies can be informed by identifying locations or patches, within distance from extant populations, which dispersing individuals may reach by using existing corridors.

While our IBM is useful for simulating dispersal movements of individuals during their transfer phases, it does not allow estimation of the time until a likely recolonization of any given location. Based on the mapped diffusion along the simulated generations we can only speculate that a recolonization of the northern Swiss habitats from the three potential source population in southwestern Germany might take a long time. An exact quantification would require accounting for demographic processes and intra-specific interactions, such as density dependent dispersal and Allee effects.

#### CONCLUSIONS

Being able to better estimate functional connectivity among sub-populations is critically important for conservation science and management. We believe that the analysis we present here provides an interesting example, and a few smaller innovations, of applying a hybrid ABC-statistical approach for calibrating individual-based movement models that can predict functional connectivity. Our study also demonstrates that such an approach can be simultaneously used for inference (e.g., male/female differences) and prediction. Regarding the latter, our results suggest that the population connectivity status of the little owl might not be as hopeless as sometimes suggested: a (re)colonization of northern Switzerland seems possible. However, simulation results suggest that populations expand slowly. Without further support, such as habitat improvement measures to

enhance potential dispersal corridors, recolonization of many suitable but uninhabited areas may not be likely in the near future.

#### ACKNOWLEDGMENTS

All authors conceived the project. B. Naef-Daenzer and M. U. Grüebler designed and executed the telemetry data collection. S. Hauenstein built the IBM, undertook the analysis, and wrote the manuscript. All authors contributed critically to the drafting and revision of this manuscript. Little owls were radio-tagged under the permission of the regional council of Baden-Württemberg, Germany (permission No. 35-9185.81/0288) and tracked under license of Deutsche Bundesnetzagentur (license No. 37 55 5413). This work was funded by the Swiss National Science Foundation (Grant 3100A 132951/1 to B. Naef-Daenzer and M. U. Grüebler), the Hirschmann foundation, and the Karl Mayer foundation. We acknowledge support by the state of Baden-Württemberg through bwHPC and the German Research Foundation (DFG) through grant No. INST 39/963-1 FUGG. G. Pe'er is funded by the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig. We also thank H. Keil for his support during the telemetry data collection, W. Fiedler, F. Kloibhofer, K. Lechner, and P. Mende for sharing their expert knowledge on little owl ecology and movement behavior and A. Schmitz for preparatory work.

#### LITERATURE CITED

- Allen, C. H., L. Parrott, and C. Kyle. 2016. An individual-based modelling approach to estimate landscape connectivity for bighorn sheep (*Ovis canadensis*). *PeerJ* 4:e2001.
- Arrignon, F., M. Deconchat, J.-P. Sarthou, G. Balent, and C. Monteil. 2007. Modelling the overwintering strategy of a beneficial insect in a heterogeneous landscape using a multiagent system. *Ecological Modelling* 205:423–436.
- Baguette, M., and H. Van Dyck. 2007. Landscape connectivity and animal behavior: functional grain as a key determinant for dispersal. *Landscape Ecology* 22:1117–1129.
- Beaumont, M. A. 2010. Approximate Bayesian computation in evolution and ecology. *Annual Review of Ecology, Evolution and Systematics* 41:379–406.
- Beaumont, M. A., W. Zhang, and D. J. Balding. 2002. Approximate Bayesian computation in population genetics. *Genetics* 162:2025–2035.
- Benhamou, S. 2011. Dynamic approach to space and habitat use based on biased random bridges. *PLoS ONE* 6:e14592.
- Benhamou, S., and D. Cornélis. 2010. Incorporating movement behavior and barriers to improve kernel home range space use estimates. *Journal of Wildlife Management* 74:1353–1360.
- Blum, M. G. 2010. Choosing the summary statistics and the acceptance rate in approximate Bayesian computation. Pages 47–56 in Y. Lechevallier and G. Saporta, editors. *Proceedings of COMPSTAT'2010: 19th International Conference on Computational Statistics*, Paris, France. Physica-Verlag, Heidelberg, Germany.
- Blum, M. G. B., M. A. Nunes, D. Prangle, and S. A. Sisson. 2013. A comparative review of dimension reduction methods in approximate Bayesian computation. *Statistical Science* 28:189–208.
- Bock, A., B. Naef-Daenzer, H. Keil, F. Korner-Nievergelt, M. Perrig, and M. U. Grüebler. 2013. Roost site selection by little owls (*Athene noctua*) in relation to environmental conditions and life-history stages. *Ibis* 155:847–856.
- Boyd, R., S. Roy, R. Sibly, R. Thorpe, and K. Hyder. 2018. A general approach to incorporating spatial and temporal variation in individual-based models of fish populations with application to Atlantic mackerel. *Ecological Modelling* 382:9–17.
- Breiman, L. 2001a. Random forests. *Machine Learning* 45:5–32.
- Breiman, L. 2001b. Statistical modeling: The two cultures. *Statistical Science* 16:199–231.
- Buderman, F. E., M. B. Hooten, J. S. Ivan, and T. M. Shenk. 2018. Large-scale movement behavior in a reintroduced predator population. *Ecography* 41:126–139.
- Csilléry, K., M. G. Blum, O. E. Gaggiotti, and O. François. 2010. Approximate Bayesian computation (ABC) in practice. *Trends in Ecology & Evolution* 25:410–418.
- Devictor, V., R. Julliard, and F. Jiguet. 2008. Distribution of specialist and generalist species along spatial gradients of habitat disturbance and fragmentation. *Oikos* 117:507–514.
- Fattebert, J., H. S. Robinson, G. Balme, R. Slotow, and L. Hunter. 2015. Structural habitat predicts functional dispersal habitat of a large carnivore: How leopards change spots. *Ecological Applications* 25:1911–1921.
- Fattebert, J., V. Michel, P. Scherler, B. Naef-Daenzer, P. Milanesi, and M. U. Grüebler. 2018. Little owls in big landscapes: Informing conservation using multi-level resource selection functions. *Biological Conservation* 228:1–9.
- Fearnhead, P., and D. Prangle. 2012. Constructing summary statistics for approximate Bayesian computation: semi-automatic approximate Bayesian computation. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 74:419–474.
- Fiedler, W. 2013. Positive Bestandsentwicklung durch Nisthilfen: Steinkauz-Streuoibstkauz. *Der Falke: Journal für Vogelbeobachter* 60:462–463.
- Glutz von Blotzheim, U., and K. Bauer. 1980. *Handbuch der Vögel Mitteleuropas*. Volume 9. Second edition. Akademische Verlagsgesellschaft, Wiesbaden, Germany.
- Gottschalk, T. K., K. Ekschmitt, and V. Wolters. 2011. Efficient placement of nest boxes for the little owl (*Athene noctua*). *Journal of Raptor Research* 45:1–14.
- Graf, R. F., S. Kramer-Schadt, N. Fernández, and V. Grimm. 2007. What you see is where you go? Modeling dispersal in mountainous landscapes. *Landscape Ecology* 22:853–866.
- Grimm, V., U. Berger, D. L. DeAngelis, J. G. Polhill, J. Giske, and S. F. Railsback. 2010. The ODD protocol: a review and first update. *Ecological Modelling* 221:2760–2768.
- Hanski, I. 1998. Metapopulation dynamics. *Nature* 396:41–49.
- Hartig, F., J. M. Calabrese, B. Reineking, T. Wiegand, and A. Huth. 2011. Statistical inference for stochastic simulation models—theory and application. *Ecology Letters* 14:816–827.
- Hastie, T. J., and R. J. Tibshirani. 1990. *Generalized additive models*. Volume 43. Chapman and Hall/CRC, London, UK.
- Jangjoo, M., S. F. Matter, J. Roland, and N. Keyghobadi. 2016. Connectivity rescues genetic diversity after a demographic bottleneck in a butterfly population network. *Proceedings of the National Academy of Science USA* 113:10914–10919.
- Kämpfer, A., and W. Lederer. 1988. Dismigration des Steinkauzes (*Athene noctua*) in Mittelwestfalen. *Vogelwelt* 109:155–164.
- Keller, V., R. Ayé, W. Müller, R. Spaar, and N. Zbinden. 2010. Die prioritären Vogelarten der Schweiz: Dokumentation zur Revision 2010. Schweizerische Vogelwarte, Sempach, und Schweizer Vogelschutz SVS/BirdLife Schweiz, Zürich, Switzerland.
- Kenward, R. E. 2000. *A manual for wildlife radio tagging*. Academic Press, London, UK.

- Killeen, J., H. Thurfjell, S. Ciuti, D. Paton, M. Musiani, and M. S. Boyce. 2014. Habitat selection during ungulate dispersal and exploratory movement at broad and fine scale with implications for conservation management. *Movement Ecology* 2:15.
- Knaus, P., S. Antoniazza, S. Wechsler, J. Guélat, M. Kéry, N. Strebel, and T. Sattler. 2018. Schweizer Brutvogelatlas 2013–2016: Verbreitung und Bestandsentwicklung der Vögel in der Schweiz und im Fürstentum Liechtenstein. Schweizerische Vogelwarte, Sempach, Switzerland.
- Kosmala, M., P. Miller, S. Ferreira, P. Funston, D. Keet, and C. Packer. 2016. Estimating wildlife disease dynamics in complex systems using an approximate Bayesian computation framework. *Ecological Applications* 26:295–308.
- Marchand, P., M. Boenke, and D. M. Green. 2017. A stochastic movement model reproduces patterns of site fidelity and long-distance dispersal in a population of fowler's toads (*Anaxyrus fowleri*). *Ecological Modelling* 360:63–69.
- McRae, B. H., B. G. Dickson, T. H. Keitt, and V. B. Shah. 2008. Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology* 89:2712–2724.
- Michel, V. T., M. V. Jiménez-Franco, B. Naef-Daenzer, and M. U. Grüebler. 2016. Intraguild predator drives forest edge avoidance of a mesopredator. *Ecosphere* 7:e01229.
- Michel, V. T., B. Naef-Daenzer, H. Keil, and M. U. Grüebler. 2017. Reproductive consequences of farmland heterogeneity in little owls (*Athene noctua*). *Oecologia* 183:1019–1029.
- Naef-Daenzer, B., D. Früh, M. Stalder, P. Wetli, and E. Weise. 2005. Miniaturization (0.2 g) and evaluation of attachment techniques of telemetry transmitters. *Journal of Experimental Biology* 208:4063–4068.
- Naef-Daenzer, B., F. Korner-Nievergelt, W. Fiedler, and M. U. Grüebler. 2017. Bias in ring-recovery studies: causes of mortality of little owls (*Athene noctua*) and implications for population assessment. *Journal of Avian Biology* 48:266–274.
- Nafus, M. G., T. C. Esque, R. C. Averill-Murray, K. E. Nussear, and R. R. Swaisgood. 2017. Habitat drives dispersal and survival of translocated juvenile desert tortoises. *Journal of Applied Ecology* 54:430–438.
- Nathan, R., W. M. Getz, E. Revilla, M. Holyoak, R. Kadmon, D. Saltz, and P. E. Smouse. 2008. A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences* 105:19052.
- Nunes, M. A., and D. Prangle. 2015. abctools: An R package for tuning Approximate Bayesian Computation analyses. <https://CRAN.R-project.org/package=abctools>.
- Pe'er, G., and S. Kramer-Schadt. 2008. Incorporating the perceptual range of animals into connectivity models. *Ecological Modelling* 213:73–85.
- Pe'er, G., K. Henle, C. Dislich, and K. Frank. 2011. Breaking functional connectivity into components: A novel approach using an individual-based model, and first outcomes. *PLoS ONE* 6:e22355.
- Pelletier, D., M. Clark, M. G. Anderson, B. Rayfield, M. A. Wulder, and J. A. Cardille. 2014. Applying circuit theory for corridor expansion and management at regional scales: tiling, pinch points, and omnidirectional connectivity. *PLoS ONE* 9:e84135.
- Prangle, D. 2018. Summary statistics. Pages 125–152 in S. A. Sisson, F. Yanan and M. A. Beaumont, editors. *Handbook of approximate Bayesian computation*. Chapman and Hall/CRC, Boca Raton, Florida, USA.
- R Core Team. 2018. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [www.R-project.org](http://www.R-project.org)
- Rabinowitz, A., and K. A. Zeller. 2010. A range-wide model of landscape connectivity and conservation for the jaguar (*Panthera onca*). *Biological Conservation* 143:939–945.
- Raynal, L., J.-M. Marin, P. Pudlo, M. Ribatet, C. P. Robert, and A. Estoup. 2017. ABC random forests for Bayesian parameter inference. *arXiv preprint arXiv:1605.05537*
- Saulnier, E., O. Gascuel, and S. Alizon. 2017. Inferring epidemiological parameters from phylogenies using regression-ABC: A comparative study. *PLoS Computational Biology* 13:1–31.
- Signer, J., and O. Ovaskainen. 2017. Detecting the influence of environmental covariates on animal movement: a semivariance approach. *Methods in Ecology and Evolution* 8:561–570.
- Stevens, V. M., C. Verkenne, S. Vandewoestijne, R. A. Wesselingh, and M. Baguette. 2006. Gene flow and functional connectivity in the natterjack toad. *Molecular Ecology* 15:2333–2344.
- Tavaré, S., D. J. Balding, R. C. Griffiths, and P. Donnelly. 1997. Inferring coalescence times from DNA sequence data. *Genetics* 145:505–518.
- Taylor, P. D., L. Fahrig, K. Henein, and G. Merriam. 1993. Connectivity is a vital element of landscape structure. *Oikos* 68:571–573.
- Van Nieuwenhuysse, D., J.-C. Génot, and D. H. Johnson. 2008. The little owl. *Conservation, ecology and behavior of Athene noctua*. Cambridge University Press, New York, New York, USA.
- Van Winkle, W. 1975. Comparison of several probabilistic home-range models. *Journal of Wildlife Management* 39:118–123.
- Vogelwarte Radolfzell. 2016. Ringfunde nestjunger beringter Steinkäuze seit 2005. Technical Report. Max Planck Institute for Ornithology, Radolfzell, Germany.
- Wegmann, D., C. Leuenberger, and L. Excoffier. 2009. Efficient approximate Bayesian computation coupled with Markov chain Monte Carlo without likelihood. *Genetics* 182:1207–1218.
- With, K. A. 1997. The application of neutral landscape models in conservation biology. *Conservation Biology* 11:1069–1080.
- Wood, S. 2017. *Generalized additive models: an introduction with R*. Second edition. Chapman and Hall/CRC, New York, New York, USA.
- Wright, M. N., and A. Ziegler. 2017. ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *Journal of Statistical Software* 77:1–17.
- Zhang, J., T. E. Dennis, T. J. Landers, E. Bell, and G. L. W. Perry. 2017. Linking individual based and statistical inferential models in movement ecology: a case study with black petrels (*Procellaria parkinsoni*). *Ecological Modelling* 360:425–436.

## SUPPORTING INFORMATION

Additional supporting information may be found online at: <http://onlinelibrary.wiley.com/doi/10.1002/eap.1873/full>

## DATA AVAILABILITY

The analysis of the model and all data manipulations were done in R v. 3.3 (R Core Team 2018). The IBM was implemented in C++ and interfaced to the LOSim R package, which is available at <https://doi.org/10.5281/zenodo.2557419>. Telemetry data and all code to reproduce the analysis is available in a figshare repository at <https://doi.org/10.6084/m9.figshare.7673669.v1>.