Detecting dispersal: A spatial dynamic occupancy model to reliably quantify connectivity across heterogeneous conservation landscapes

Divya Vasudev [a,b,c], Varun R. Goswami [a,b,d,e,*], Madan K. Oli [d,e]

[a] Conservation Initiatives, Guwahati 781022, Assam, India
[b] Centre for Wildlife Studies, Bengaluru 560042, Karnataka, India
[c] Wildlife Conservation Trust, Mumbai 400021, Maharashtra, India
[d] School of Natural Resources and the Environment, University of Florida, Gainesville, FL 32611, USA
[e] Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, FL 32611, USA

ARTICLE INFO

ABSTRACT

Connectivity shapes species distribution, spatial population dynamics and genetic structure, and is critical for conservation. It is imperative to reliably identify factors that limit connectivity across heterogeneous, fragmented landscapes. Viewing connectivity as the persistence of movement in space from source to destination—along potential corridors simulated from correlated random walks—we adapt the dynamic occupancy modeling framework to present an approach to reliably quantify connectivity. This approach allowed us to estimate the probability of spatial persistence of movement along potential corridors, while accounting for imperfect detection of animal use of the matrix. We use simulations to test our model, and data from sign surveys of Asian elephant *Elephas maximus* space use in Garo Hills, Northeast India, as a practical application of the method. Distance to forests and ruggedness predominantly shaped elephant connectivity patterns. Negative effects of distance to forests were intensified at the onset of dispersal, and in the more disturbed part of our landscape. We mapped locations critical to maintaining connectivity in our study landscape. We demonstrate that the widely used occupancy modeling approach, when combined with appropriate field data collection, permits explicit assessment of matrix impacts on connectivity while accounting for imperfect detection of animal movement. In so doing, we highlight the value of the approach in enabling inference on *where* animals move, as well as *why*. Obtaining a reliable understanding of factors shaping connectivity is central to understanding and predicting species responses to environmental change, thereby facilitating effective long-term conservation in fragmented landscapes.

1. Introduction

Connectivity, which refers to the movement of individuals or genes among populations or habitats (Clobert et al., 2001), has major implications for species viability and evolution. Connectivity impacts adaptability to novel habitats, demography, colonization of unoccupied habitat, species turnover, and inter-species interactions, among other effects (MacArthur and Wilson, 1967; Orrock et al., 2003; Fletcher et al., 2016). Most endangered species are faced with habitat fragmentation (Crooks et al., 2017). Hence, connectivity is prioritized by conservation programs in an effort to mitigate negative consequences of habitat loss and fragmentation, land-use dynamics, and climate change, on biodiversity and species persistence.

Our ability to observe animal dispersal—which shapes patterns of connectivity—and to model connectivity based on empirical observations, has vastly increased over the last two decades (McRae et al., 2008; Cagnacci et al., 2010). Methodological advances facilitate insights into the impacts of the matrix (e.g., McRae et al., 2008) and animal behavior (e.g., Vasudev and Fletcher, 2015) on connectivity patterns. In particular, GPS and satellite telemetry has drastically changed our ability to observe animal dispersal (Cagnacci et al., 2010). Consequently, not only is our understanding of processes that shape connectivity patterns improved, we are also able to predict future connectivity in the face of environmental change (e.g., Revilla and Wiegand, 2008), and provide conservation solutions for endangered species in fragmented landscapes (e.g., Vasudev and Fletcher, 2015).

---

* Corresponding author at: Conservation Initiatives, Shillong 793001, Meghalaya, India.
  E-mail address: varunr.goswami@gmail.com (V.R. Goswami).

Notwithstanding new methodological approaches, dispersal continues to be difficult to observe as it is an infrequent event that occurs at large spatial scales (Lima and Zollner, 1996). Telemetry, which offers a fine-scale view of long-distance animal movement, is expensive and requires invasive capture processes; while non-invasive approaches are subject to biases from partial observability, or the possibility that a disperser moves through a location but goes undetected. In this sense, it is a non-trivial issue to distinguish areas that are unused for dispersal (i. e., true absence of dispersal) from areas that are used but where dispersers move through undetected (i.e., non-detection of dispersal). Neglecting to account for such imperfect detection of dispersal can lead to underestimates of dispersal rates between populations, and to biased inference on the influence of landscape features on dispersal patterns (see Yackulic et al., 2013). These biases could further translate into inaccurate predictions of connectivity patterns, misleading demarcations of 'critical' corridors and stepping-stones, and consequently, inefficient connectivity conservation investments.

Imperfect detection impacts our observation of a number of ecological patterns and processes, and analytical approaches that account for imperfect detection has allowed reliable inference on species demography (Ozgul et al., 2009) and richness (Boulinier et al., 1998), and human–wildlife conflict (Goswami et al., 2015). These approaches are broadly encompassed by the capture–recapture and occupancy modeling frameworks, which rely on replicated surveys to determine the probability of detecting the animal or ecological event of interest, given its presence (Williams et al., 2002; MacKenzie et al., 2018).

These methods have been used previously to incorporate imperfect detection of dispersal events. Multistate capture–recapture models provide pair-wise population dispersal rates in metapopulations with marked individuals (e.g., Ozgul et al., 2009). Occupancy models with auto-logistic terms have been used to make inferences on connectivity (e.g., Yackulic et al., 2012) whereby a dispersal (and colonization) event is inferred if an unoccupied site is occupied at a subsequent time. These models, as applied till date, account for imperfect detection of *species presence in patches*.

These are a class of methods that focus on sampling from *source* and *destination patches* (of dispersers) to make indirect inference on the impact of the intervening matrix on dispersal (e.g., using genetic assessments: Broquet et al., 2006; capture–recapture: Ozgul et al., 2009). These methods can be contrasted with other approaches that make inferences on dispersal through fine-scale information on individual dispersers in the matrix (e.g., radio-telemetry: Revilla and Wiegand, 2008; movement follows: Vasudev and Fletcher, 2015). Notwithstanding the utility of both classes of methods, an approach that addresses imperfect detection of disperser movement through fine-scale sampling in the matrix, can add substantially to our ability to directly infer matrix impacts on species movement, understand dispersal patterns in heterogeneous landscapes, and implement connectivity conservation interventions.

Ultimately, there has been slow progress on understanding *why* dispersal is limited; in fact, there are few species in heterogeneous landscapes for which we have clearly identified major limitations or threats to connectivity (Vasudev et al., 2015). These issues are particularly pertinent for wide-ranging species with large habitat requirements, such as the endangered Asian elephant *Elephas maximus*. Most landscapes that house Asian elephants are under serious human impact (Leimgruber et al., 2003), fragmenting existing habitats, and making connectivity critical to the species (Goswami and Vasudev, 2017).

We propose an adaptation of dynamic occupancy models (MacKenzie et al., 2003; MacKenzie et al., 2018) to identify factors limiting elephant connectivity, while: (a) accounting for imperfect detection of animal movement; and (b) sampling at a fine-scale within the matrix to allow direct inference on the impacts of matrix characteristics on dispersal. Our data are from a heterogeneous, multiple-use landscape in Garo Hills, in the state of Meghalaya, Northeast India. We test our approach using

simulations, and highlight how this widely applicable approach can provide reliable and comparable connectivity assessments, and importantly, better inference on the impact of landscapes elements on connectivity.

2. Methods

2.1. Study area

Our study landscape in Garo Hills (25°08′–25°23′N; 90°37′–90°58′E), Northeast India, comprises a mosaic of protected areas and other government-managed forests in a matrix of community-managed forests, agriculture and human habitation (see Goswami et al., 2014). Of the government-managed forests in the landscape, Balphakram National Park along with Siju Wildlife Sanctuary and Rewak Reserve Forest forms one block of contiguous protected habitat (hereinafter, 'Balphakram') that is physically separated from the Baghmara Reserve Forest to the southwest (hereinafter, 'Baghmara') by other land uses. Earlier investigations have shown that these protected forest patches serve as primary habitat for elephants, while the community-managed forests and agricultural land uses play a subsidiary role of supporting elephant movement and secondary habitat use (Goswami et al., 2014). Following from these findings, here we assess factors shaping connectivity between the Balphakram and Baghmara patches of elephant habitat. Our landscape is therefore a two-patch system and our sampling space is the matrix that lies *between* the two focal patches.

2.2. Quantifying connectivity

We measure connectivity as the probability of successful dispersal, given initiation of dispersal (Vasudev et al., 2015), which broadly aligns with some commonly used measures of connectivity (Table S1). We posit that this definition, when scaled as a probability, is quantifiable and comparable across contexts. We condition the probability of successful dispersal on initiation of dispersal because once dispersal is initiated, landscape-related factors largely shape connectivity rather than within-population factors that may influence decisions of individuals to emigrate (Bowler and Benton, 2005). This rationale is used in other connectivity modeling approaches, and does not preclude examining edge effects, or movement impediments at the onset or extremely early stages of dispersal. We restrict our inference to *potential connectivity*—or the probability of an area being used for connectivity—rather than *actual* or *realized connectivity*, which represents the actual use of an area for connectivity. We do so as the realization of connectivity would be impacted by both landscape-scale drivers, and within-population factors such as population density.

Connectivity is cumulative, in that it is achieved through the successful completion of movement that initiates at a source patch, persists along a particular dispersal path or corridor, and concludes at a destination patch. Therefore, connectivity can be measured as the probability of animal movement persisting across space, from source to destination patches.

The matrix can serve as both secondary habitat and movement corridors (Driscoll et al., 2013), but on-ground differentiation between 'use' and 'movement' is difficult. Here, we do not make this distinction; instead, we simply consider the spatial persistence of animal 'use' of the matrix, from the source patch to destination patches, as connectivity.

2.3. Spatial dynamic occupancy model

Occupancy models assess spatial variation in species distribution (MacKenzie et al., 2018). They have also been applied at finer scales to assess spatial patterns of animal 'use' across heterogeneous landscapes (e.g., Goswami et al., 2014).

We quantify connectivity through a simple extension of dynamic occupancy models (MacKenzie et al., 2003, 2018). This approach

considers the occupancy state of sampling units to change from one season to the next, due to colonization or local extinction (MacKenzie et al., 2018). In the connectivity context, we consider sampling units to be subdivisions of the matrix corresponding to sections of potential dispersal corridors, some of which are surveyed.

Whereas past applications of dynamic occupancy models have modeled the persistence of animal occupancy (or use) over *time* (i.e.,

across seasons) (e.g., Yackulic et al., 2012), here we use the same conceptual and analytical framework to model persistence of animal use across *space* (Fig. 1a). We consider each 'step' along a spatial sequence of sampling units that originates at the source, and extends continuously within the matrix to end at the destination, to be equivalent to a season in traditional dynamic occupancy models. We model the persistence of animal use across space—or steps—along these predetermined spatial
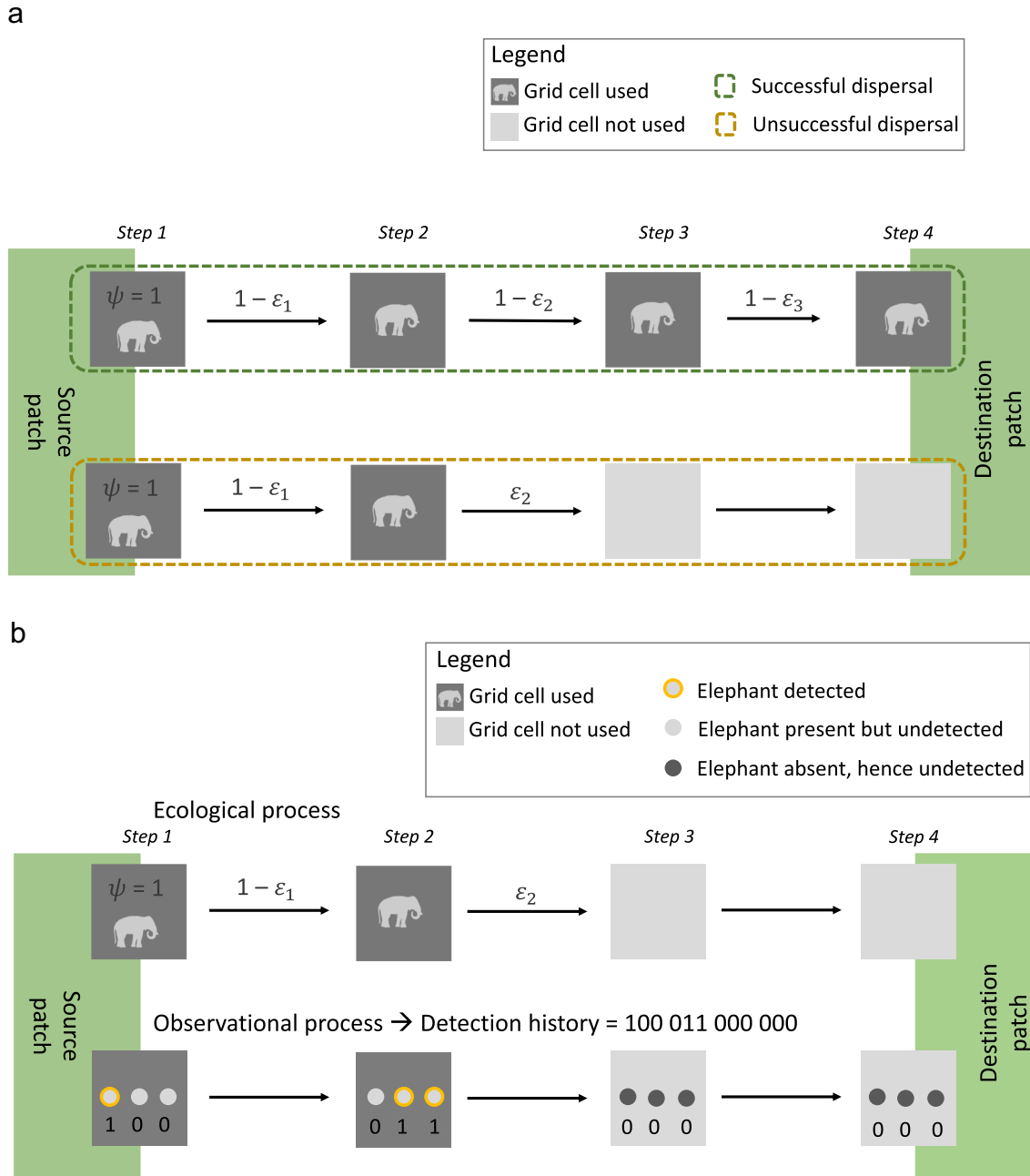

Fig. 1. (a) Schematic representation of change in 'use' of grid cells along a hypothetical potential corridor. Potential corridors are sequences of 'steps' or sampling units (grey squares) extending from the source to the destination patch (shown in green). We condition the model on use in the first step, which is at the edge of the source patch, that is, $\psi_1 = 1$. Subsequently, used cells can become unused with an extinction probability ε and continue to be used with a persistence probability of $1 - \varepsilon$. Successful dispersal can occur when 'use' of grid cells persists from the source till the destination, all along the potential corridor (marked as green dashed rectangle). In contrast, when 'use' does not persist, this precludes successful dispersal (marked as yellow dashed rectangle. (b) We show the observational process in our model for one example potential corridor. For this hypothetical corridor, step 1 and 2 are used, while step 3 and 4 are not used. We show an example detection history considering three replicates per grid cell. In step 1, elephant signs are detected in replicate 1, but not detected in replicates 2 and 3, leading to a detection history of 100; and so forth. (c) We show three example potential corridors in the study area connecting habitat patches, Baghmara and Balphakram, overlaid on the sampled grid network. The three example corridors each represent a scenario where: (i) use persists along the corridor (successful dispersal); (ii) use does not persist along the entire length of the corridor (unsuccessful dispersal); and (iii) use does not persist past the edge of the source patch (edge effects). The inset shows the location of the study area in India. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
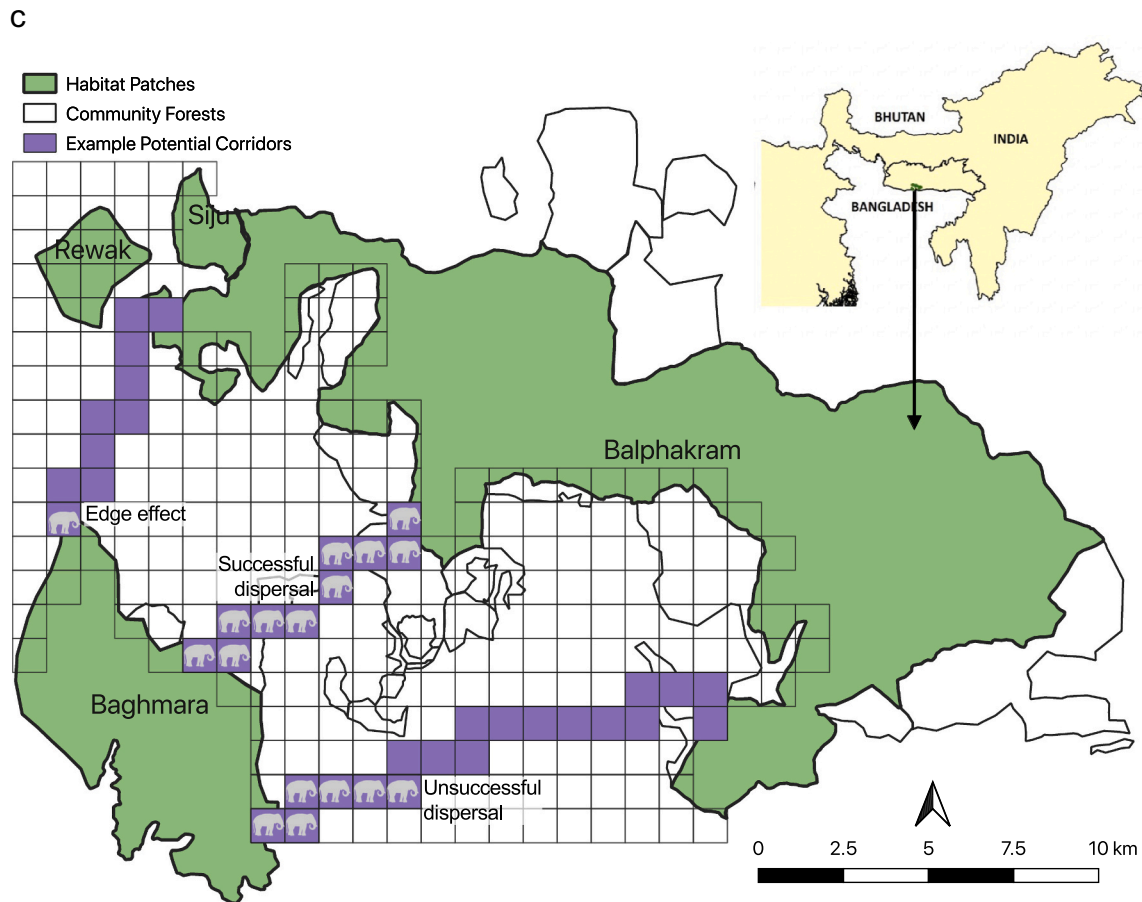
c

Fig. 1. (*continued*).

sequences of sampling units, henceforth, 'potential corridors' (Fig. 1b). Thus, in our model, we use (a) steps in place of temporal units (seasons) of traditional dynamic occupancy models, and (b) potential corridors (representing a sequence of steps) as the unit of inference (Fig. 1; Table S2). Along each corridor, animals can only proceed by passing through the identified steps in a sequential manner.

Dynamic occupancy models allow us to then assess the impact of different covariates on parameters of the model, namely (a) the persistence (or 1 – extinction) of animal use along potential corridors, and (b) colonization of movement routes, representing a funneling effect whereby corridors attract dispersers from their immediate neighborhood. For each potential corridor, we can also calculate the *probability of successful dispersal* as the probability of animal use persisting through all of its steps from source to destination.

$$\varphi_s = \psi_s^1 \times \prod_{i=1}^{n-1}\left(1 - \varepsilon_s^i\right) \tag{1}$$

where, φ_s is the probability of successful dispersal along potential corridor s, and ε_s^i is the probability of extinction from movement step i to step $i + 1$, along n steps of the corridor. ψ_s^1 is the probability of the animal using the first step of the potential corridor, and when this quantity is equal to 1, φ_s becomes the probability of successful dispersal, conditional on initiation of dispersal along potential corridor s. We note that the probability of successful dispersal corresponds to at least one animal starting at the source and successfully reaching the destination; it does not take into consideration the number of animals that have emigrated from the source patch, nor does it include colonizers of corridors, even if such individuals end up reaching the destination.

Through this model, patches that are farther apart, or potential

corridors that are more circuitous, automatically get penalized through an increased number of movement steps (see Eq. (1)). That is, longer corridors entail a greater number of steps, and hence are subjected to more extinction parameters. This is in consonance with established impacts of distance on successful dispersal (MacArthur and Wilson, 1967). Further, factors that lead to the preclusion of successful dispersal, or positively influence extinction probability ε_i^s, can be identified as limits to dispersal. By increasing ε_i^s, these factors would reduce the probability of successful dispersal φ_s in Eq. (1). Thus, a short corridor with high extinction probability may turn out to be less successful than a more circuitous corridor with low probabilities of extinction; this aligns with our understanding of isolation by resistance (McRae et al., 2008). Taken together, both length of the corridor (through increased number of steps) and matrix impermeability (through higher extinction probability) impact the probability of successful dispersal.

Key assumptions of the spatial dynamic occupancy model are analogous to those of the standard dynamic occupancy model (MacKenzie et al., 2003). (1) The use state of a sampling unit does not change (or changes randomly) over replicated surveys within a step. The use of sampling units change from one step to the next through the colonization and extinction processes. (2) Heterogeneity in colonization and extinction probabilities from one step to the next, across potential corridors, is modeled using appropriate covariates. (3) Detection probability is either constant across space, or if spatially heterogeneous, is modeled appropriately. (4) Species use of grid cells along potential corridors is Markovian, the validity of which can be tested using the non-Markovian parameterization of the model. (5) There are no false positives in species detection, arising potentially from species misidentification. Our model differs from previous applications of dynamic occupancy models to the study of connectivity (e.g., Yackulic et al.,

2012) in (a) directly modeling the persistence of movement across space, and (b) sampling fine-scale animal use of the *matrix*. This allows us to directly model impacts of landscape characteristics on the probability of successful dispersal.

2.4. Potential corridors

We overlaid a 1 km × 1 km grid network over the study area using Quantum GIS (QGIS.org., 2019). We generated potential corridors based on simulated correlated random walks, a model that is used extensively across taxa to model animal movement (Bovet and Benhamou, 1988), modified to include some restrictions described below. We modeled both Balphakram—the larger and more undisturbed habitat fragment in the landscape—as the source with Baghmara as the destination, and vice versa. We generated 100 potential corridors for each source, restricting the corridor length to 50 km.

To generate potential corridors, we followed the following steps. First, we identified 'exit points' at the boundary of the source patch at a resolution of 100 m. For each random walk, we chose an exit point and start direction from a uniform distribution. Step lengths were sampled from a Weibull distribution with a mean of 100 m. We chose turning angles from a wrapped Cauchy distribution, with a mean of 0, and a shape parameter ρ randomly chosen from a uniform distribution, for each route. We thus generated both relatively straight-line pathways as well as those that were more tortuous. We then generated a random walk with the above-described values. We discarded paths that did not enter the destination patch before accumulating a total path length of 50 km, and only included paths that lead from the source to the destination. Thus, all paths had their initial step at the edge of the source and their last step at the edge of the destination patch. Animals could leave from any part of the perimeter of the source, and enter the destination at any point along its edge.

We superimposed the 100 generated walks onto the 1 km × 1 km grid network, and identified the sequence of grid cells that encompassed each random walk. Each grid cell represented a sampling unit or a 'step'. This sequence of grid cells now represented the configuration or sequence of 'steps' along potential corridors (see Fig. 1b for example corridors). We thus had 100 potential corridors for each of the two sources in our landscape; we note that these 100 are a sample of a near-infinite number of potential corridors in the landscape. These and subsequent analyses were conducted in R (R Development Core Team, 2019).

Potential corridors had a maximum of 30 steps in our model. These corridors are representative of 1-km wide movement swathes that stretch from the source patch to the destination patch, encompassing one or more *potential* movement paths of elephants. Since occupancy models are currently designed for sampling units to have the same number of seasons, we artificially extended potential corridors to accommodate extra steps at their last grid cell where necessary, such that all potential corridors had the same number of steps. This extension was only used when estimating parameters within the occupancy modeling framework; we set extinction of these 'dummy' steps to 0 while calculating probabilities of successful dispersal for each potential corridor.

2.5. Simulation

Dynamic occupancy models have been previously tested through simulation (MacKenzie et al., 2003). We therefore focused our simulations to test the robustness of this approach in estimating the impact of covariates on extinction probability ε_i^s, our parameter of primary interest. We considered a simple context of two matrix types: an *inhospitable* matrix where persistence of movement was low, compared to a *hospitable* matrix. Thus, in our simulation, there was a single categorical factor influencing extinction probability. In real-world applications, there may be multiple such categorical or continuous covariates that lead to spatial variation in extinction probability. Each grid cell was randomly assigned to either one of these two matrix types. We ran 100 simulations for multiple scenarios of matrix permeability (Table S3).

We simulated animal movement along potential corridors generated as described above, based on randomly chosen values of ε for the hospitable and inhospitable matrix (Table S3). We denoted a grid cell as used if an animal moved through it in their simulated walks along at least one potential corridor. We simulated detection histories based on a randomly chosen detection probability. We then estimated model parameters based on simulated detection histories, using the spatial dynamic occupancy model. The simulations of animal pathways were independent of the fitting process.

For each simulation, we noted the mean squared error between the estimated extinction parameter $\widehat{\varepsilon}$ and the true value of the parameter, and that of the ratio of extinction rates of the two matrix types based on their true and estimated values. To test for biases arising due to the number of simulated potential corridors, and the number of dispersal events, we varied these randomly for each simulation (Table S3).

2.6. Field data collection

A significant part of our study landscape was remote and difficult to access. Hence, we used spatial replication for our survey (MacKenzie et al., 2018) where each spatial replicate constituted a 300 m walk in a pre-determined direction. We walked between 0 and 13 spatial replicates in each grid cell. Unsampled replicates were included as missing data. The field data were collected as part of a larger research project assessing the conservation value of the landscape; further details on the sampling design can be found in Goswami et al. (2014).

Between January 2011 and February 2012, we invested a total walk effort of approximately 540 km, and recorded 2225 detections of elephant presence based on recent signs, primarily their dung (96.4%). There was no temporal ordering to our sampling (i.e., we did not sample source cells first and destination cells last, or vice versa). We recorded the following ground-based covariates: land use, which included the categories forest, *jhum* fallow (i.e., areas undergoing varying stages of successional regeneration following the abandonment of slash-and-burn shifting cultivation), monoculture plantation and human habitation; and, the presence of water bodies.

2.7. Occupancy analyses

For each grid cell, we recorded the following remotely-sensed covariates: Euclidean distance to the source patch; distance to forests, which was the shortest distance between the grid cell centroid and any forest in the landscape; ruggedness index, which is a measure of the variation in elevation across the grid cell (Riley et al., 1999); and mean village density. We used QGIS and ArcGIS (Environmental Systems Research Institute, Redlands, USA) to calculate these covariates.

We estimated the following parameters in our extension of the standard dynamic occupancy model (see MacKenzie et al., 2003): (1) ψ_s^1, which is the occupancy in the first step for movement route s; (2) ε_s^i, which is the extinction probability (1 – persistence) for step i in corridor s, (3) γ_s^i, which is the colonization probability for step i of corridor s, and (4) p, which is the probability of detecting animal use of a grid cell in a replicate, given that the grid cell has been used. As we define dispersal success conditional on emigration, and we know that elephants use the source patch (Goswami et al., 2014), we set ψ_s^1, or the probability of use in the first step, as 1.

Following Goswami et al. (2014), we first identified the most appropriate model structure for detection probability p, while using the most general model for other parameters (i.e., we modeled ε and γ as functions of additive effects of all considered covariates). We modeled p as a function of (a) ruggedness index, (b) land use as categorized above, and (c) additive models of the above two covariates. However, models incorporating land use failed to converge. We chose the most appropriate model based on Akaike's Information Criterion, corrected for

small sample sizes (AIC$_c$; Burnham and Anderson, 2002).

We fixed p to the model chosen above, and allowed extinction probability ε and colonization probability γ to vary as a function of covariates of interest. We modeled ε as a function of distance from source fragment, distance to forests, ruggedness index, village density and the presence of water bodies, and additive effects of the above. Additionally, we assessed *disperser fatigue*, that is, if persistence decreased as the disperser progressed farther along a particular movement corridor or if limits to dispersal impact extinction more at later stages of dispersal. We did this by including the additive and interactive impacts of step identity, a numerical identifier that ranged from 1 to 29 and denoted where the grid cell was located along a sequence of steps in a potential corridor. We also modeled the effect of distance to forest on ε_s^1 to assess if the *onset* of dispersal is impacted by the presence of community forests (forests outside the sources and destinations in our landscape), that is, if community forests mitigated edge effects of less permeable matrix types.

We expected the presence of water bodies, distance to forests, distance to protected areas (which excluded community forests), and the contrast between the village density of a grid cell and that of its immediate neighborhood (henceforth relative village density), to potentially impact colonization probability γ. We included additive models of the above-listed covariates. Throughout, we checked for multicollinearity before including more than a single covariate in models.

The assumption of spatial dependency in species use of the landscape, that is, the dependency of ψ_s^{i+1}, or the use of potential corridor s at movement step $i + 1$, on ψ_s^i is fundamental to our model. To test the validity of this assumption, we included a model wherein we fixed extinction probability ε to $1 - \gamma$ (1 – colonization probability). In setting $\varepsilon = 1 - \gamma$, the probability of use of a potential corridor s at movement step $i + 1$, calculated as:

$$\psi_s^{i+1} = \psi_s^i \times (1 - \varepsilon) + \left(1 - \psi_s^i\right) \times \gamma$$

reduces to:

$$\psi_s^{i+1} = \gamma = 1 - \varepsilon$$

Thus, in this model parameterization, species use of grid cells along potential corridors becomes non-Markovian (MacKenzie et al., 2003). We compared this model to one where extinction and colonization probabilities were spatially invariant but estimated separately, that is, based on Markovian dynamics. We used program MARK (White and Burnham, 1999) implemented in R through the RMark library (Laake and Rexstad, 2007) to run the dynamic occupancy models. We modeled dispersal from Balphakram, and from Baghmara, separately, as two distinct model sets, as factors limiting dispersal may differ between these two contexts.

From the chosen model, we noted the impact of covariates on ε (1 – persistence) of movement across space within the matrix. We estimated the probability of successful dispersal for each potential corridor φ_s as per Eq. (1), using the above extinction probabilities. We mapped the above-calculated success probability for each grid cell; if a grid cell was part of more than one potential corridor, we assigned it the maximum value of the estimated probability of successful dispersal across corridors. Finally, as forest edges become increasingly hard, and with the fencing of certain forests (Goswami and Vasudev, 2017), it is sometimes important to identify potential 'entry' and 'exit' points from and to primary habitats that are critical to maintain connectivity. To identify these, we quantified an Edge Importance Index by simply adding the probability of successful dispersal across all potential corridors that a particular grid cell either commences or ends in.

3. Results

3.1. Simulation

Simulations suggest that the model was able to estimate extinction probability with more precision at relatively low 'true' values. We were able to estimate relative extinction parameter estimates between the *hospitable* and *inhospitable* matrix (that is, the coefficients linked to the extinction parameter) with more precision when true parameter values were not very low (Fig. S1). The estimators were robust to changes in detection probability, the proportion of the landscape that comprised *hospitable* matrix, number of simulated potential corridors, and number of 'used' corridors (results not shown). The number of potential corridors impacted representation of grid cells (Fig. S2), but did not have substantial effects on uncertainty around parameter estimates, except for very small numbers of potential corridors (Fig. S3).

3.2. Detecting animal movement

Detection probability was consistently estimated to be <1, supporting the claim that partial observability is relevant to recording animal movement. The overall probability of detecting animal movement (or use of a site for movement) per sampling replicate within a grid cell p_s^i, as estimated by the best supported model where detection probability was spatially invariant (i.e., constant), was: (a) 0.53 (95% CI = 0.50–0.54) when Balphakram was modeled as the source; and (b) 0.58 (95% CI = 0.57–0.59), when Baghmara was modeled as the source. However, irrespective of the source, there was substantially greater support for a model that included effects of ruggedness on detection probability (Table S4). Detection probability was negatively influenced by ruggedness, that is, we are able to detect animal use of a sampling unit, given use of the unit, better on more even terrain (Tables S5 and S6; Fig. S4).

3.3. Spatial variation in movement

There was little support for the model where $\varepsilon = 1 - \gamma$; the difference in AIC$_c$ between this model and a Markovian model where the two parameters were spatially invariant (i.e., the intercept-only $\varepsilon[.]$, $\gamma[.]$ model) was 2412.95 and 625.95 for when Balphakram and Baghmara were the source, respectively. This lends strong support to our assumption of spatial dependence in use of grid cells along potential corridors.

When Balphakram was the source, one model received clear support with ΔAIC$_c$ > 2 for all other models (Table S5; Burnham and Anderson, 2002). For movement in the opposite direction with Baghmara as the source, three models were within ΔAIC$_c$ of 2 (Table S5). We present results from the top model in the first case, and from the most general model among the three models for which ΔAIC$_c$ < 2, when Baghmara was the source. Results were qualitatively similar for both model sets, with the same model receiving support irrespective of the source of dispersal.

Extinction probability varied based on an interactive effect of step identity and distance to forests, and an additive effect of ruggedness (Tables S5 and S6, Fig. 2). Extinction probability increased with distance to forests (including both government- and community-managed forests), nearing 1 as corridors extended to 4 km away from the closest forest (Fig. 2). This effect was stronger when elephants dispersed out of Baghmara, the more disturbed habitat patch (Table S6). The increase in extinction probability with greater distance to forests, was intensified at later stages of dispersal when Balphakram was the source; or viewed another way, when the destination (Baghmara) was a more disturbed habitat patch (Fig. 2). The effect was the opposite when Baghmara was the source. Furthermore, the impact of distance to forests in increasing extinction probability (or conversely, reducing the persistence of movement) was intensified substantially at the onset of dispersal (Fig. 2). Extinction probability also increased with ruggedness, though
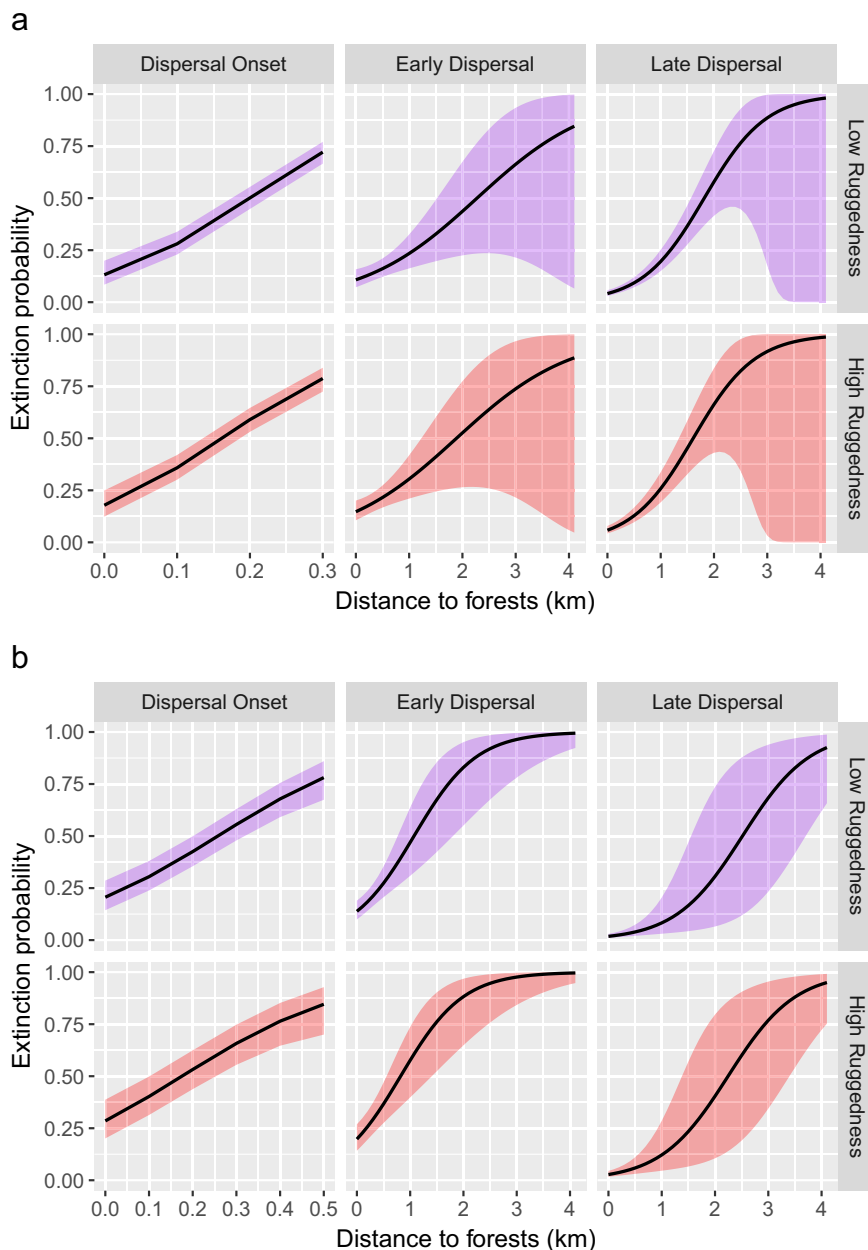
Fig. 2. Extinction probability when dispersing from (a) Balphakram or (b) Baghmara, varied as an interactive function of distance to (government- and community-managed) forest and stage of dispersal, wherein extinction probability increased as the animal moved farther away from forests, and this relationship intensified at later stages of dispersal when originating from Balphakram, but early stages when originating from Baghmara. The onset of dispersal was strongly influenced by distance from forests (shown as the first column). Extinction probability was also weakly but positively influenced by ruggedness.

this relationship was relatively weak (Table S6, Fig. 2).

Colonization probability varied with relative village density (the difference between village density of a grid cell, and its neighborhood; Fig. S5). When grid cells had village densities much lower than their neighborhood, colonization probability was high; it dipped when village density at a site was comparable to that in its neighborhood (Fig. S5). Colonization probability was higher closer to forests. Contrary to expectations, the presence of water bodies decreased the probability of colonization of a grid cell (Fig. S5). Estimates of the impact of the above covariates on model parameters are provided in Table S6.

3.4. Probability of successful dispersal and edge importance index

The probability of successful dispersal averaged across all simulated potential corridors was 0.03 ± 0.004 (mean $\pm$ SE; probability of successful dispersal from Balphakram $= 0.04 \pm 0.007$, and from Baghmara $= 0.02 \pm 0.004$, Fig. S6). We show the probability of successful dispersal across grid cells in Fig. 3a. The Edge Importance Index of grid cells

adjoining the two focal patches, which was the cumulative probability of successful dispersal across all corridors passing through those 'entry' or 'exit' points, is shown in Fig. 3b. The grid cells with high probability of successful dispersal loosely form a wider movement zone, or corridor, between Balphakram and Baghmara (Fig. 3a).

4. Discussion

Dispersal is an infrequent event that plays out over large landscapes. Clearly, the issue of partial observability is pertinent for unbiased inference. Detection probability was consistently <1 in our study (Fig. S4) and it varied with the same ecological covariate (namely ruggedness) as our parameter of interest: extinction probability of movement across space. This evidence lends credence to the need to account for models that deal with imperfect detection for reliable inference on connectivity and limits to successful dispersal (see Yackulic et al., 2013). By adapting the rapidly developing occupancy modeling framework (MacKenzie et al., 2018), we demonstrate a method that is
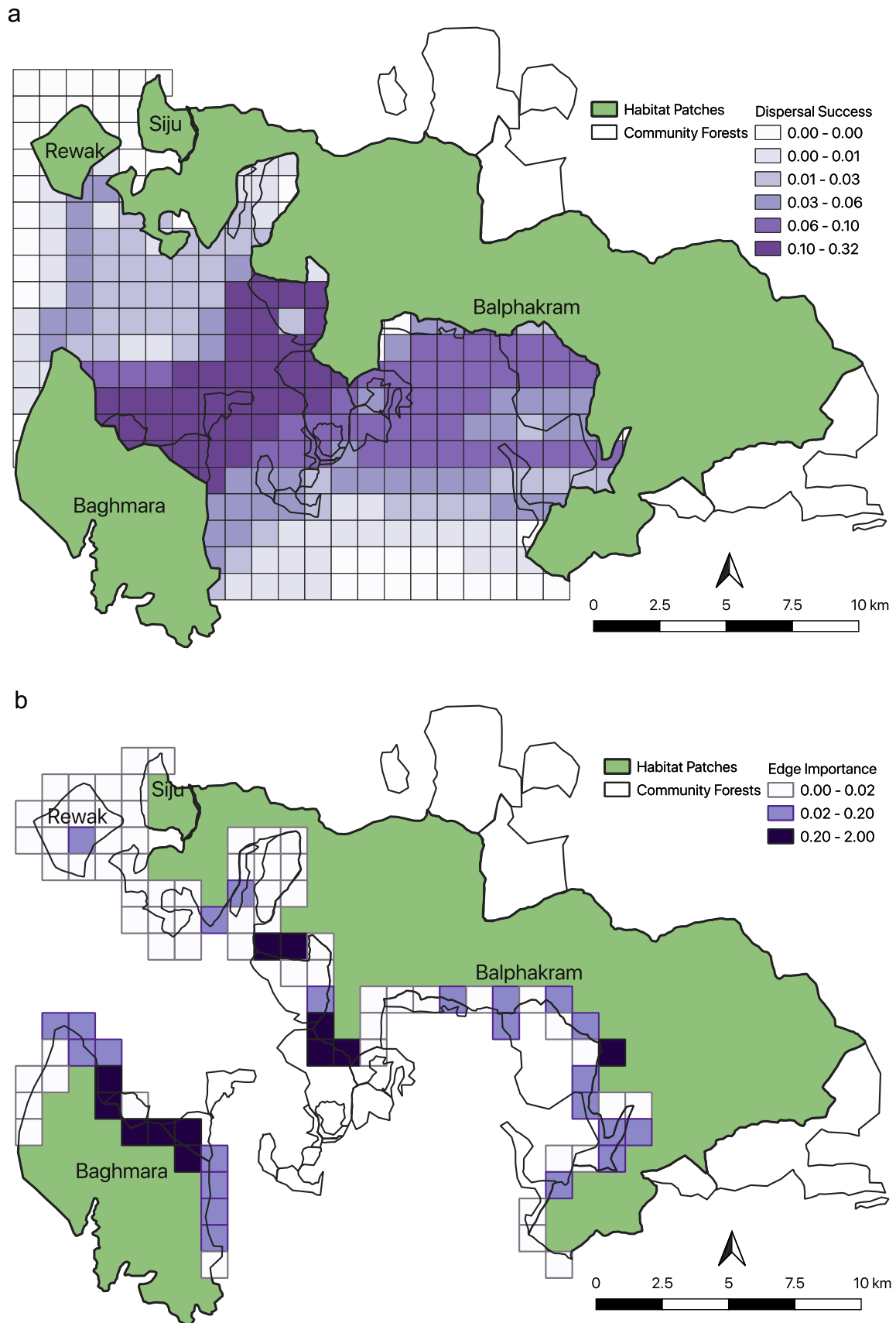
Fig. 3. (a) Probability of successful dispersal for grid cells between two protected area patches in the Baghmara-Balphakram landscape, Meghalaya, Northeast India, showing their contribution to connectivity through potential corridors, and (b) Edge Importance Index of grid cells adjoining the habitat patches, showing their contribution to connectivity as 'entry' and 'exit' dispersal points.

non-invasive, widely accessible to conservationists, applicable to a broad range of taxa and landscapes, and ultimately, an approach that has much potential to spur better understanding of limits to successful dispersal in heterogeneous landscapes.

4.1. Probability of successful dispersal as a metric for landscape connectivity

There are several metrics to describe connectivity. Still there is a lack of clarity on the relationship between such metrics, the robustness and comparability across contexts, and their relation to ecological processes (Calabrese and Fagan, 2004). Developments in scientific decision-making approaches, such as structured decision modeling and adaptive management (Martin et al., 2009), have seldom been considered in the field of connectivity conservation. We suggest that one of the reasons for this lag is that there are few metrics that translate easily into the quantifiable objective functions required for these approaches.

We posit that the metric we use, namely the probability of successful dispersal, presents advantages as a quantifiable metric that describes connectivity, while aligning with other commonly used metrics and definitions (Table S1). As the metric is scaled as a probability, it is comparable across taxa, landscapes, and time. Consequently, it can be adapted to objective functions within systematic decision-making approaches. Additionally, this metric is rooted in ecological theory and its biological significance is clear (Vasudev et al., 2015). We provide a method to identify limits to dispersal (Fig. 2), which can also feed into these models to inform conservation.

While we identify (a) grid cells, and a 'movement zone' or corridor, with high probability of successful dispersal, and (b) critical 'entry' and 'exit' points (Fig. 3), we emphasize that these can change over time. The critical 'entry' and 'exit' points were not restricted to the 'movement zone' (Fig. 3a and b), emphasizing the need for caution in fencing forests or otherwise obstructing animal movement in and out of habitats. Following from Vasudev et al. (2015), we suggest that these predictions of *where* animals move are not as important as inference on *why* they move through certain places and not others.

We clarify that the probability of successful dispersal relates to the probability of *at least* one animal dispersing, that is, it does not make any assumptions on the number of animals dispersing. This is true for most applications of presence/absence data, or detection/non-detection data used in occupancy models, where a site that is occupied, may be occupied either by a single animal or more than one animal. Since it does not make this assumption, the model can be used for both gregarious species moving in groups as well as solitary dispersers. That said, an extension of the dynamic occupancy model does allow multiple states of species occurrence to be combined with transitions arising due to colonization and extinction processes (MacKenzie et al., 2009); these states can be used to represent low-intensity versus high-intensity space use (Goswami et al., 2014). Adapting the multistate dynamic occupancy model (MacKenzie et al., 2009) to quantify connectivity might allow potential corridors to be distinguished based on their frequency of use. However, certain corridors that are infrequently used may still be critical to maintaining metapopulation persistence, and should be prioritized for conservation.

4.2. Limits to dispersal

Even though elephants are wide ranging and frequent non-forested habitat, our findings show that connectivity had little chance of success farther than 4 km from forests. This highlights the need to protect forested corridors between populations. The importance of proximity to forests increased closer to Baghmara, a habitat patch that adjoins an urban center, has a major road traversing through it, and as a reserve forest, is accorded a lower level of legal protection than the more remote Balphakram National Park. It is conceivable therefore that elephants should have a higher threat perception outside forests nearer to Baghmara than Balphakram. Therefore, the effect of distance to forests, interacting with the length of corridors, is likely to be a manifestation of elephants' response to a perceived risky and unsuitable matrix. This is further substantiated by our finding that during the onset of dispersal, extinction probability of elephant movement increased with distance to forests (Fig. 2). In real-world contexts, geographic isolation of populations and habitats is likely to be accompanied by an increased hostility of the matrix; our results indicate that these two factors occurring in unison, is likely to have a compounded and disproportionately strong negative impact on connectivity.

Lands between protected areas can serve as both secondary habitat and movement conduits (Driscoll et al., 2013). Goswami et al. (2014) found that the value of the matrix in this landscape as habitat depended on distance to government-managed forests and village density. Here, we show that the use of the same lands for *movement* (during the same time period) is impacted by ruggedness and the distance to forests (inclusive of both government- and community-managed forests). While we did not distinguish between these two roles of the matrix during data collection, the distinct analytical approaches used in Goswami et al. (2014) and here, throw to light the different constraints to elephant use of the landscape as secondary habitat, and their use of it for movement, respectively.

Village density did impact movement, but by exerting a funneling effect on potential corridors. Water bodies also exerted the same effect, but contrary to expectation, they decreased the probability of new dispersers using a corridor. Garo Hills largely harbors wet vegetation types, and hence water in itself may not be a limiting factor for animal space use or movement. Rather, water bodies here may be serving as a proxy for an underlying and unobserved factor, such as the increased activity of humans around water bodies. We note that colonization in our model could emerge from an animal entering a corridor from the matrix, or from grid cells featuring in more than one potential corridor.

4.3. Recommendations for future studies

We present a non-invasive, feasible approach to studying a burning question in connectivity research and conservation: what limits dispersal? The first step to implementing this approach would be to identify source and destination patches. These can be protected areas, potential habitats for viable or breeding populations, or polygons of high habitat suitability. The matrix interspersing these source habitats would be the target survey area. We recommend subdividing this space into sampling units (e.g., grid cells) of appropriate size. We chose grid cells of size 1 km^2 based on a scale that is: (a) ecologically appropriate, as the study species likely perceives the landscape and takes movement decisions at this scale; (b) relevant for conservation, as corridors are often this width or more; and (c) methodologically appropriate, as local spatial homogeneity is maintained within grid cells in the landscape. The spatial dynamic occupancy model can provide insights into limits to connectivity at this scale, but is not designed to assess finer-scale (i.e., within grid cell) movement decisions.

The sampling units should be surveyed within an occupancy framework (MacKenzie et al., 2018), recording species presence and appropriate covariates that may restrict or facilitate animal movement. We note that estimates of ε are dependent on the size of the grid cell: the larger the grid cell, the fewer the steps in each corridor, and the higher will be the value of ε Thus, we recommend making inference on factors impacting ε, and on the probability of successful dispersal φ_s, for the entirety of movement from source to destination. There is a trade-off between two major, and opposing, considerations that go into determining the number of potential corridors. One, that the majority of sampling units that may play a role in connectivity need to be represented in at least one potential corridor. Second, we would want to limit representation of a single sampling unit in multiple potential corridors. It is unavoidable that certain sampling units, simply by virtue of their geographic location, will be represented more frequently than others.

However, we would want to limit such over-representation. Our chosen number of 100 potential corridors provides a nice balance between these two criteria (Figs. S2 and S3).

5. Conclusion

The importance of separating ecological processes from sampling processes has been widely recognized in the study of animal populations and communities (e.g., Yackulic et al., 2013), and the utility of inference thus obtained for informed conservation (Martin et al., 2009). Connectivity modeling and conservation lags far behind in this regard, despite the relevance of partial observability issues for the large-scale infrequent processes shaping connectivity patterns. Further, there is a critical knowledge gap on limits to successful dispersal in real-world conservation landscapes (Vasudev et al., 2015). Obtaining reliable inference on these drivers of connectivity can greatly enhance our understanding of how animals move across landscapes. In turn, this will improve our ability to understand impacts of ecological and anthropogenic factors on species connectivity and landscape-level persistence, and predict their responses to environmental change. Ultimately, this will enable us to better conserve species in heterogeneous, human-dominated landscapes.

CRediT authorship contribution statement

Divya Vasudev: Conceptualization, Formal analysis, Funding acquisition, Methodology, Software, Visualization, Writing - original draft, Writing - review & editing. **Varun R. Goswami:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Visualization, Writing - original draft, Writing - review & editing. **Madan K. Oli:** Funding acquisition, Methodology, Supervision, Validation, Writing - review & editing.

Declaration of competing interest

The authors have no conflict of interest to report.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.biocon.2020.108874.

References

Boulinier, T., Nichols, J.D., Sauer, J.R., Hines, J.E., Pollock, K.H., 1998. Estimating species richness: the importance of heterogeneity in species detectability. Ecology 79, 1018–1028.

Bovet, P., Benhamou, S., 1988. Spatial analysis of animals' movements using a correlated random walk model. J. Theor. Biol. 131, 419–433.

Bowler, D.E., Benton, T.G., 2005. Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. Biol. Rev. 80, 205–225.

Broquet, T., Ray, N., Petit, E., Fryxell, J.M., Burel, F., 2006. Genetic isolation by distance and landscape connectivity in the American marten (*Martes americana*). Landsc. Ecol. 21, 877–889.

Burnham, K., Anderson, D.R., 2002. Model Selection and Multimodel Inference: A Practical Information-theoretic Approach. Springer-Verlag, New York, USA.

Cagnacci, F., Boitani, L., Powell, R.A., Boyce, M.S., 2010. Animal ecology meets GPS-based radiotelemetry: a perfect storm of opportunities and challenges. Philos. Trans. R. Soc. B-Biol. Sci. 365, 2157–2162.

Calabrese, J., Fagan, W., 2004. A comparison-shopper's guide to connectivity metrics. Front. Ecol. Environ. 2, 529–536.

Clobert, J., Danchin, E., Dhondt, A.A., Nichols, J.D., 2001. Dispersal. OUP, Oxford.

Crooks, K.R., Burdett, C.L., Theobald, D.M., King, S.R.B., Di Marco, M., Rondinini, C., Boitanig, L., 2017. Quantification of habitat fragmentation reveals extinction risk in terrestrial mammals. Proc. Natl. Acad. Sci. U.S.A. 114, 7635–7640.

R Development Core Team, 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Driscoll, D.A., Banks, S.C., Barton, P.S., Lindenmayer, D.B., Smith, A.L., 2013. Conceptual domain of the matrix in fragmented landscapes. Trends Ecol. Evol. 28, 605–613.

Fletcher, R.J., Burrell, N.S., Reichert, B.E., Vasudev, D., Austin, J.D., 2016. Divergent perspectives on landscape connectivity reveal consistent effects from genes to communities. Curr. Landsc. Ecol. Rep. 1, 67–79.

Goswami, V.R., Vasudev, D., 2017. Triage of conservation needs: the juxtaposition of conflict mitigation and connectivity considerations in heterogeneous, human-dominated landscapes. Front. Ecol. Evol. 4, 144. https://doi.org/10.3389/fevo.2016.00144.

Goswami, V.R., Sridhara, S., Medhi, K., Williams, A.C., Chellam, R., Nichols, J.D., Oli, M. K., 2014. Community-managed forests and wildlife-friendly agriculture play a subsidiary but not substitutive role to protected areas for the endangered Asian elephant. Biol. Conserv. 177, 74–81.

Goswami, V.R., Medhi, K., Nichols, J.D., Oli, M.K., 2015. Mechanistic understanding of human–wildlife conflict through a novel application of dynamic occupancy models. Conserv. Biol. 29, 1100–1110.

Laake, J.L., Rexstad, E.A., 2007. RMark – an alternative approach to building linear models in MARK. In: Cooch, E.G., White, G.C. (Eds.), Program MARK: A Gentle Introduction. http://www.phidot.org/software/mark/docs/book.

Leimgruber, P., Gagnon, J.B., Wemmer, C., Kelly, D.S., Songer, M.A., Selig, E.R., 2003. Fragmentation of Asia's remaining wildlands: implications for Asian elephant conservation. Anim. Conserv. 6, 347–359.

Lima, S.L., Zollner, P.A., 1996. Towards a behavioral ecology of ecological landscapes. Trends Ecol. Evol. 11, 131–135.

MacArthur, R.H., Wilson, E.O., 1967. The Theory of Island Biogeography. Princeton University Press, Princeton, USA.

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

MacKenzie, D.I., Nichols, J.D., Seamans, M.E., Gutierrez, R.J., 2009. Modeling species occurrence dynamics with multiple states and imperfect detection. Ecology 90, 823–835.

MacKenzie, D.I., Nichols, J.D., Royle, J.A., Pollock, K.H., Bailey, L.L., Hines, J.E., 2018. Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence, Second edn. Elsevier, London, UK.

Martin, J., Runge, M.C., Nichols, J.D., Lubow, B.C., Kendall, W.L., 2009. Structured decision making as a conceptual framework to identify thresholds for conservation and management. Ecol. Appl. 19, 1079–1090.

McRae, B.H., Dickson, B.G., Keitt, T.H., Shah, V.B., 2008. Using circuit theory to model connectivity in ecology, evolution, and conservation. Ecology 89, 2712–2724.

Orrock, J.L., Danielson, B.J., Burns, M.J., Levey, D.J., 2003. Spatial ecology of predator–prey interactions: corridors and patch shape influence seed predation. Ecology 84, 2589–2599.

Ozgul, A., Oli, M.K., Armitage, K.B., Blumstein, D.T., Van Vuren, D.H., 2009. Influence of local demography on asymptotic and transient dynamics of a yellow-bellied marmot metapopulation. Am. Nat. 173, 517–530.

QGIS.org., 2019. QGIS Geographic Information System: Open Source Geospatial Foundation Project. Retrieved from. http://qgis.org.

Revilla, E., Wiegand, T., 2008. Individual movement behavior, matrix heterogeneity, and the dynamics of spatially structured populations. Proc. Natl. Acad. Sci. U.S.A. 105, 19120–19125.

Riley, S.J., DeGloria, S.D., Elliot, R., 1999. A terrain ruggedness index that quantifies topographic heterogeneity. Int. J. Therm. Sci. 5, 23–27.

Vasudev, D., Fletcher, R.J., 2015. Incorporating movement behavior into conservation prioritization in fragmented landscapes: an example of western hoolock gibbons in Garo Hills, India. Biol. Conserv. 181, 124–132.

Vasudev, D., Fletcher, R.J., Goswami, V.R., Krishnadas, M., 2015. From dispersal constraints to landscape connectivity: lessons from species distribution modeling. Ecography 38, 967–978.

White, G.C., Burnham, K.P., 1999. Program MARK: survival estimation from populations of marked animals. Bird Study 46 (Supplement), 120–138.

Williams, B.K., Nichols, J.D., Conroy, M.J., 2002. Analysis and Management of Animal Populations. Academic Press, San Diego, California, USA.

Yackulic, C.B., Reid, J., Davis, R., Hines, J.E., Nichols, J.D., Forsman, E., 2012. Neighborhood and habitat effects on vital rates: expansion of the Barred Owl in the Oregon Coast Ranges. Ecology 93, 1953–1966.

Yackulic, C.B., Chandler, R., Zipkin, E.F., Royle, J.A., Nichols, J.D., Campbell Grant, E.H., Veran, S., 2013. Presence-only modelling using MAXENT: when can we trust the inferences? Methods Ecol. Evol. 4, 236–243.