Check for updates

Received: 30 March 2022 | Accepted: 10 September 2022

DOI: 10.1111/2041-210X.14014

Methods in Ecology and Evolution BRITISH ECOLOGICAL SOCIETY

APPLICATION

abmR: An R package for agent-based model analysis of large-scale movements across taxa

Benjamin Gochanour[1,2] | **Javier Fernández-López**[3] | **Andrea Contina**[4,5,6]

[1]Corix Plains Institute, University of Oklahoma, Norman, Oklahoma, USA; [2]Oklahoma Biological Survey, University of Oklahoma, Norman, Oklahoma, USA; [3]Department of Biology, University of Massachusetts Boston, Boston, Massachusetts, USA; [4]Department of Integrative Biology, University of Colorado Denver, Denver, Colorado, USA; [5]Department of Microbiology and Plant Biology, Center for Earth Observation and Modeling, University of Oklahoma, Norman, Oklahoma, USA and [6]Department of Integrative Biology, University of Texas at Austin, Austin, Texas, USA

Correspondence
Benjamin Gochanour
Email: gochanour.benjamin@mayo.edu

Funding information
National Natural Science Foundation of China, Grant/Award Number: 81961128002; United States National Science Foundation, Grant/Award Number: DEB 1911955, DGE 1545261 and EF 1840230

Handling Editor: Sarah Goslee

Abstract

1. Agent-based modelling (ABM) shows promise for animal movement studies. However, a robust, open-source and spatially explicit ABM coding platform is currently lacking.
2. We present abmR, an R package for conducting continental-scale ABM simulations across animal taxa. The package features two movement functions, each of which relies on the Ornstein–Uhlenbeck (OU) process.
3. The theoretical background for abmR is discussed and the main functionalities are illustrated using example populations.
4. Potential future additions to this open-source package may include the ability to specify multiple environmental variables or to model interactions between agents. Additionally, updates may offer opportunities for disease ecology and integration with other R movement modelling packages.

KEYWORDS
animal migration, ecology, open-source, R programming, simulations

1 | INTRODUCTION

Animal movement is a complex behavioural trait that affects the survival of populations and species across taxa (Berg, 1983; Dingle, 2014). Long- and short-distance movements can follow predictable environmental constraints, allowing populations to take advantage of seasonal food resources (e.g. migration), or more opportunistic, such as in the case of dispersal behaviours aimed at avoiding predators or finding potential mates (Giuggioli & Bartumeus, 2010). Thus, wild animals make decisions often based on environmental cues that lead to movement patterns characteristic of different populations across the landscape (Dodge et al., 2014; Nathan et al., 2008). However, obtaining a comprehensive understanding of large-scale animal movement behaviour and population occurrence under climate change scenarios or habitat loss has proven to be a challenge (Araujo & Guisan, 2006). Moreover, while the research toolbox in movement ecology studies has seen a considerable expansion over the last two decades due to technological advancements of the tracking devices and molecular markers (Cushman & Lewis, 2010; Williams et al., 2020), the limitation of scaling up individual data to population-level inferences is still a substantial obstacle (Hawkes, 2009; but see Holdo & Roach, 2013). A promising research approach that may overcome the limitations of wildlife movement studies hindered by small sample sizes is represented by computer simulations within an agent-based modelling (ABM) framework (Bridge et al., 2017; Tang & Bennett, 2010).

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2022 The Authors. *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society.

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14014, Wiley Online Library on [09/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The core principle of ABM is to simulate a set of entities, called agents, which are defined by intrinsic properties as well as behavioural rules governing their interactions with the environment (Grimm & Railsback, 2013). That is, agents are described by their inherent attributes while dynamically interacting with external conditions such as the co-occurrence of other agents and/or changing features of their environmental setting. Thus, ABM is used in many areas, including biology, disease risk, social sciences, and economics (Grimm & Railsback, 2013; Klimek et al., 2015; Polhill et al., 2008; Willem et al., 2017) with the unifying goal of investigating and predicting the dynamics of complex systems (Grimm et al., 2005). In particular, wildlife studies have adopted the ABM approach to simulate population growth, reproduction, mortality rate, energy budget, and migration ecology, just to cite a few (Aurbach et al., 2020; Brown & Robinson, 2006; Goldstein et al., 2021; Lustig et al., 2019). However, we currently lack a robust and spatially explicit ABM coding platform for the implementation of large-scale animal movement investigations (but see Chubaty & McIntire, 2021; Thiele et al., 2012). Here, we present a novel ABM framework in R programming language (R Core Team, 2022) for applications in animal behaviour and movement ecology.

2 | PACKAGE OVERVIEW

abmR allows for both computation and visualization of agent movement trajectories through a set of behavioural rules based on environmental parameters. The two movement functions, moveSIM and energySIM, provide the central functionality of the package, allowing the user to run simulations using an Ornstein-Uhlenbeck movement model (Uhlenbeck & Ornstein, 1930; hereafter OU). Additional functions provide a suite of visualization and data summarization tools intended to reduce the effort needed to go from results to presentation-ready figures and tables (Table 1). The package is currently available on the Comprehensive R Archive Network (CRAN) at https://cran.r-project.org/web/packages/abmR/index.html.

The abmR package is built to facilitate customization of movement model parameters within the R environment (Figure 1). Parameters affecting agent behaviour can be obtained from real-world observations (e.g. GPS data) or approximated from the literature and then manually entered as arguments into abmR functions. Thus, if the user knows beforehand the behaviour and the ecological constraints of the agents, such as land-cover preferences (e.g. vegetation composition and structure) or movement direction and average distance travelled per day, they may transfer this knowledge into abmR to study changes in energy consumption, mortality rate or alternative routes across different environmental conditions (e.g. raster layers). Alternatively, for simulations of 'synthetic species' designed to compare relative mortality rates across habitats, for example, the user may develop an analytical framework within abmR without exploring the ecology of a particular organism. In this case, the movement parameters may not be realistic, but selected with the unique goal of comparing a set of scenarios (e.g. movements in predicted future land-use changes) while studying the emerging properties of the system (Yin et al., 2022). Finally, abmR functionalities can be integrated with other approaches, such as a step-selection function (Thurfjell et al., 2014) and/or Bayesian statistics to estimate movement parameters and explore migration patterns or habitat occupancy of a particular set of agents (Cullen et al., 2022; Joo et al., 2020).

Both movement functions used by abmR rely on the same OU approach. The OU process is described by the following stochastic differential equation:

$$dX_t = k(\theta - X_t)dt + \alpha dW_t$$

where W_t is a standard Brownian motion on $t \in [0, \infty)$, k is a positive constant that controls the rate of reversion to the long-term mean, θ, of the OU process (see Blomberg et al., 2020), and $\alpha > 0$ is a constant controlling the volatility of the Brownian motion. In abmR, given the current agent location (x_t, y_t), agent location at the subsequent timestep (x_{t+1}, y_{t+1}) is modelled according to the following equations:

$$x_{t+1} = x_t + \sigma \times Z_x + \phi_x \times (\theta_x - x_t), \quad (1)$$

$$y_{t+1} = y_t + \sigma \times Z_y + \phi_y \times (\theta_y - y_t). \quad (2)$$

Here, σ is a user-specified multiplier on the random terms Z_x and Z_y, two numbers drawn from the Normal (0,1) distribution. In addition, ϕ_x and ϕ_y are movement motivation or attraction strength for the OU process

TABLE 1 Functions contained in the abmR package (v. 1.0.6). For more complete function descriptions, consult the abmR manual.

Function	Usage
moveSIM	Runs agent-based model movement simulations based on environmental data
moveVIZ	Creates a plot or table of moveSIM() results
energySIM	Runs agent-based model movement and energy budget simulations based on environmental data
energyVIZ	Creates a plot or table of energySIM() results
tidy_results	Prints results from moveSIM() or energySIM() in an easier-to-read table
get_ex_data	Downloads data that is used in examples in vignette and documentation
as.species	Creates object of class 'species' for input into moveSIM() or energySIM()

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14014, Wiley Online Library on [09/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

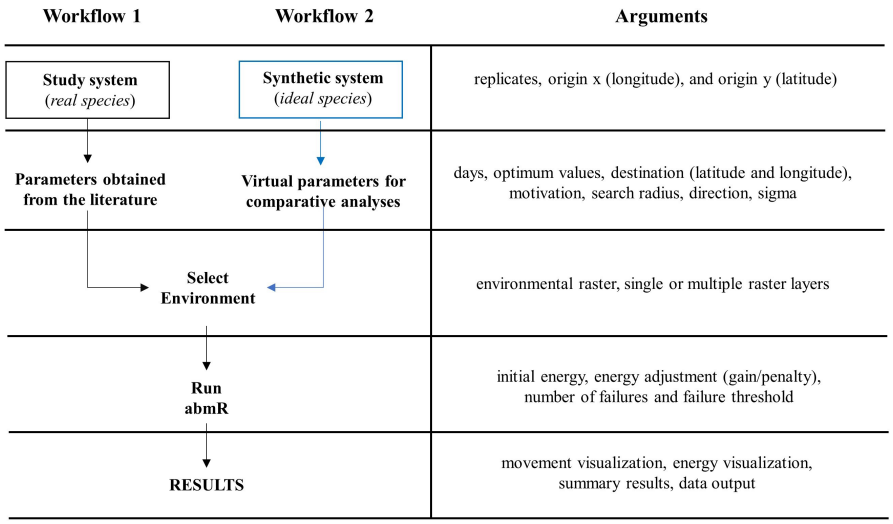


FIGURE 1 abmR workflow from data input to result exploration. The black box and black arrows (workflow 1, left) show the sequence of the analytical steps, starting with the identification of the study system (species or population), followed by the acquisition of the movement parameters from previous observations or the literature. After selecting the environmental conditions (e.g. raster layers), users can run a suite of abmR functions to study movements, changes in energy consumption or mortality rate. The blue box and blue arrows (workflow 2, right) show an alternative analytical workflow in which users can create synthetic species for comparative analyses across predicted ecological and environmental scenarios (e.g. loss of migratory behaviour and habitat fragmentation). The panel to the right of both workflows shows a list of the relative arguments that can be entered in abmR across each step of the analysis.

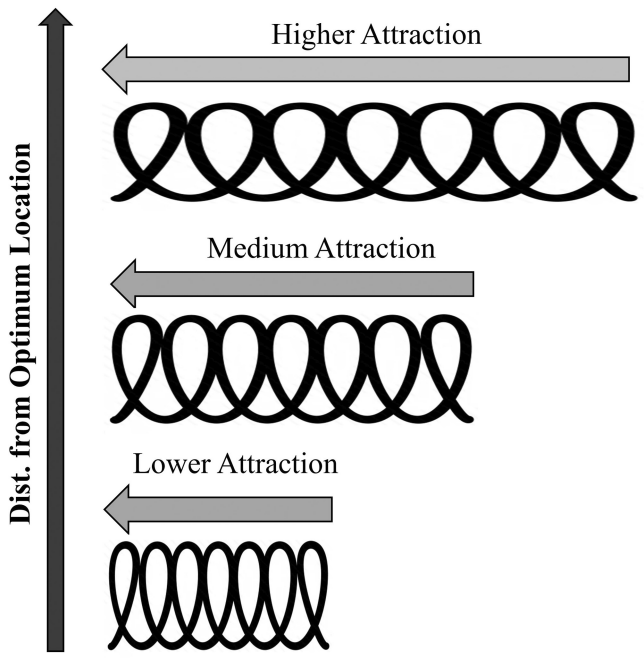


FIGURE 2 The Ornstein–Uhlenbeck (OU) model given in (1) and (2) performs like a spring-coil: Agents further from their target location experience higher attraction (and travel further), while agents closer to their destination experience lesser attraction (and travel less far).

in the longitude and latitude coordinates, respectively, while θ_x and θ_y are optimal *x* (longitude) and *y* (latitude) coordinates, respectively. It is assumed that the origin point (x_1, y_1) is known. The OU model given in (1) and (2) performs similarly to a spring-coil (Figure 2). Greater distance from optimal coordinates θ_x and θ_y acts like a compressed spring to propel distant agents towards θ_x and θ_y. On the other hand, agents closer to θ_x and θ_y will travel a shorter distance on that timestep. However, the length of movement also depends on ϕ_x and ϕ_y, because these motivations serve as a multiplier on $(\theta_x - x_t)$ and $(\theta_y - y_t)$, respectively (Equations 1 and 2).

While the two movement functions are distinct (see below), each follows the same basic two steps. The first, large-scale searching, is illustrated in Figure 3 and follows an elliptical pattern, as opposed to a circular distribution, to account for directional migratory impulses (e.g. genetically controlled migratory behaviours or seasonal Zugunruhe; Merlin & Liedvogel, 2019). This step finds the 'optimum' location for each agent (θ_x and θ_y coordinates from Equations 1 and 2). The coordinates θ_x and θ_y are determined according to an algorithm that selects the location whose observed environmental value has the least possible difference from the user-specified optimal raster value. For moveSIM, this user-specified optimal raster value is supplied directly, while for energySIM it is the average of the lower and upper bounds of a user-specified optimum range. The optimum value or range of values specified depends on the modelling scenario and the type of environmental raster that is used (e.g. vegetation, temperature, etc.).

Agents will move towards the selected optimum location. However, if the attraction strength (ϕ_x and ϕ_y in Equations 1 and 2) is less than 1, agents will have a 'target' location short of the optimal location. Thus, agents find an 'optimum' location within the semi-circular search region and then a 'target' location that lies on the line between the 'current location' and the 'optimum' location. Moreover, agents will move towards this target location with some variance, which is generated by sampling two numbers from the standard normal distribution and multiplying by σ, as specified by

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14014, Wiley Online Library on [09/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

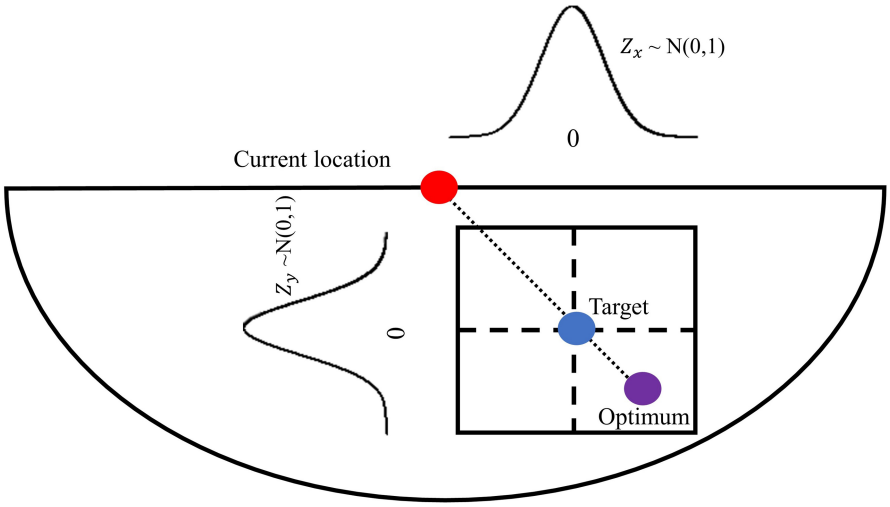


FIGURE 3 Illustration of large-scale searching specified by the OU model of Equations (1) and (2). Agents find an 'optimum' location within the semi-circular search region and then a 'target' location that lies on the line between the 'current location' and the 'optimum' location. If there is a tie between multiple potential 'optimum' cells, one is randomly selected from the list of tied cells to serve as the optimum. Random variance is added by independently sampling Z_x and Z_y from a standard normal distribution. Here, $\sigma = 1$ and $\phi_x = \phi_y < 1$, where σ is the multiplier on the random error and ϕ_x and ϕ_y are the motivations in the x and y directions, respectively. Bounding box represents the most probable samples from the N(0,1) distribution.

TABLE 2 List of arguments used in moveSIM (M), energySIM (E), or as.Species (S). F, false; SD, standard deviation; T, true. In text, these arguments are presented in *italics*. For a more complete list of argument descriptions, see the abmR documentation.

Argument	Function	Usage
Replicates	M, E	# of agents to model
Days	M, E	# of timesteps
modelled_species	M, E	Species object from as.species()
env_rast	M, E	Environmental raster
Optimum	M	Optimal environmental value
optimum_lo	E	Lowest optimum environmental value
optimum_hi	E	Highest optimum environment value
dest_x	M, E	Destination Longitude
dest_y	M, E	Destination Latitude
mot_x	M, E	Motivation (*x* direction)
mot_y	M, E	Motivatitm (*y* direction)
search_radius	M, E	Radius of semi-circular search region (km)
Direction	M, E	Movement direction: N, S, E, W, or R (Random)
Sigma	M, E	Randomness parameter
Mortality	M, E	Incorporate agent mortality? T or F
fail_thresh	M	Deviation from optimum constituting failure
n_failures	M	Allowable # of failures before agent death
init_energy	E	Initial energy
energy_adj	E	Energy gain/penalty vector
single_rast	M, E	Using a single-layer raster? T or F
write_results	M, E	Save results as a .csv? T or F
X	S	Origin longitude
Y	S	Origin latitude

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14014, Wiley Online Library on [09/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the user (in Figure 3, σ is 1). Because the support of the normal distribution consists of all real numbers, large deviations from the 'target' point are possible. However, because the normal distribution has low density at the extreme tails, outcomes are most likely to fall within a certain area of the target, as illustrated in Figure 3. This first step corresponds to the OU model of Equations (1) and (2).

The second step is small-scale searching. Here, agents select the 'best' of the 8 neighbouring cells (queen's case or Moore neighbourhood) after performing step 1, discussed above and in Figure 3. Again, 'best' here means the cell with the environmental raster value closest to the agent's user-defined optimum range. These two steps are then repeated for each timestep until the agent dies or proceeds through all timesteps. Each timestep will use different environmental raster layers. Users may choose to supply a raster stack to simulate changes in environment over time (see for example Section 3.1).

For the moveSIM function, agent death occurs when agents fail to achieve suitable environmental raster values for more than a user-specified number of consecutive timesteps. Here, what constitutes a 'suitable' cell is determined by the optimum value and an allowable deviation proportion, both also specified by the user. For the energySIM function, agent death occurs when energy reaches zero. For both functions, users may choose to disable agent mortality. In the following subsections, we present the differences between moveSIM and energySIM functions and their underlying algorithms.

2.1 | Simulation function: moveSIM

The function moveSIM runs an OU movement simulation based on environmental conditions provided by the user (e.g. raster), optionally including agent mortality and adjusted motivation according to user-specified parameters. The function operates according to the following algorithm. Here, terms in *italics* are moveSIM function arguments (see Table 2).

The following algorithm applies when the argument *direction* is 'N', 'S', 'E', or 'W'. For random movement (*direction* = '*R*') agents simply select a random point from a circle of radius *search_radius* for each timestep (Step 2). Here, let *env_rast* (x_{t+1}, y_{t+1}) be the value of *env_rast* at the point (x_{t+1}, y_{t+1}). The core algorithm shown here assumes that *env_rast* contains no undefined (N/A) grid cells.

1. Specify (x_1, y_1) using *x* and *y* contained in *modelled_species*, set failures = 0.
2. For day *t* in 1:(*days*-1)
   a. Define search area semicircle (radius = *search_radius*) facing *direction* and centered at (x_t, y_t)
   b. Determine (θ_x, θ_y) as location within the search area with *env_rast* value closest to *optimum*.
      (i) If (*dest_x*, *dest_y*) in search area, set (θ_x, θ_y) = (*dest_x*, *dest_y*)
   c. Large scale searching: find $(x_{t+1}, y_{t+1})_0$ according to (1) and (2).
   d. Small-scale searching: set (x_{t+1}, y_{t+1}) as location within eight neighbouring cells (queen's case) of $(x_{t+1}, y_{t+1})_0$ with the value closest to *optimum*.
   Perform (e)–(f) if *mortality* = True.
   a. If observed *env_rast* (x_{t+1}, y_{t+1}) − *optimum* > *optimum***fail_thresh*, set failures = failures + 1. If not, set failures = 0.
   b. If failures > *n_failures* agent dies. End loop.
3. Return dataframe with *days* rows and 2 columns movement track data.
4. Repeat (1)–(3) *replicates* times.

2.2 | Simulation function: energySIM

The function energySIM builds on moveSIM by allowing for dynamic agent energy levels that are affected by the quality of environmental values achieved. These initial user-defined energy levels then serve as a driver of mortality and movement distance per timestep. The energy thresholds range from 0 to 1 and represent proportion deviations from the optimum. These thresholds are arbitrarily divided into 10 even increments (0.1, 0.2, etc.), but users can change the energy gain or penalty for attaining each of the thresholds. It operates according to the following algorithm. Here, terms in *italics* are energySIM function arguments (see Table 2) or calculated variables (e.g. *optimum*, *energy*).

The following algorithm applies when the argument *direction* is 'N', 'S', 'E', or 'W'. For random movement (*direction* = '*R*') agents simply select a random point from a circle of radius *search_radius* for each timestep (Step 3). Here, let *env_rast* (x_{t+1}, y_{t+1}) be the value of *env_rast* at the point (x_{t+1}, y_{t+1}). The core algorithm shown here assumes that *env_rast* contains no undefined (N/A) grid cells.

1. Specify (x_1, y_1) using *x* and *y* contained in *modeled_species*.
2. Compute *optimum* as (*optimum_hi* − *optimum_lo*)/*2 and* set *energy* = *init_energy*
3. For day *t* in 1:(*days*-1)
   a. If *mortality* = True, update *search_radius* as *search_radius* = *search_radius* * *(energy/init_energy)*.
   b. Define search area semicircle(radius = *search_radius*) facing *direction* and centered at (x_t, y_t).
   c. Determine (θ_x, θ_y) as location within the search area with *env_rast* cell value closest to *optimum*.
      (i) If (*dest_x*, *dest_x*) in search area, set (θ_x, θ_y) = (*dest_x*, *dest_y*)
   d. Large-scale searching: find $(x_{t+1}, y_{t+1})_0$ according to (1) and (2).
   e. Small-scale searching: set (x_{t+1}, y_{t+1}) as location within eight neighbouring cells (queen's case) of $(x_{t+1}, y_{t+1})_0$ with the cell value closest to *optimum*.
   f. Update energy level according to deviation of env_rast (x_{t+1}, y_{t+1}) from optimum.
   g. If *mortality* = True and *energy* = 0, the agent dies. End loop.
4. Return dataframe with *days* rows and 2 columns movement track data.
5. Repeat (1)–(4) *replicates* times.

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14014, Wiley Online Library on [09/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

BOX 1 **R code used for performing the simulations presented in Figure 4. First, as.species is called to initialize two populations with different origin locations. Then, energySIM is called to perform a movement simulation for each population; parameters that differ between the two simulations are printed in red, functions in blue, and objects in bold. For argument descriptions, see Table 2 and the package manual.**

Population 1	Population 2
am.pop.1 = as.species(x = −105.7, y = 48.2)	**am.pop.2** = as.species(x = −142.7, y = 63.2)
sim.move <− energySIM(	**simtwo.move <− energySIM(**
replicates = 250,	replicates = 250,
days = 14,	days = 14,
env_rast = as.raster.stack.ndvi.sep,	env_rast = as.raster.stack.ndvi.sep.
search_radius = 150,	search_radius = 800,
sigma = 0.1,	sigma = 0.1,
dest_x = 999,	dest_x = 999,
dest_y = 999,	dest_y = 999,
mot_x = 0.95,	mot_x = 0.8,
mot_y = 0.95,	mot_y = 0.8,
modeled_species = **am.pop.1,**	modeled_species = **am.pop.2,**
optimum_lo = 0.2,	optimum_lo = 0.6,
optimum_h1 = 0.5,	optimum_h1 = 0.8,
init_energy = 100,	init_energy = 70,
direction = "S",	direction = "S",
mortality = F,	mortality = F,
Energy_adj = c(20, 10, 8, 5, 2, 0, −2, −5, −8, −10, −20),	energy_adj = c(20, 10, 8, 5, 2, 0, −2, −5, −8, −10, −20),
write_results = T,	write_results = T,
single_rast = F)	single_rast = F)

3 | EXAMPLE APPLICATIONS

abmR can be used to construct ABM simulations for any desired agent across the globe. In the following examples, we demonstrate the computation capabilities of the energySIM function, although a similar workflow also applies for moveSIM. In the first example, we show how energySIM can be used to compare movements and differential energy allocations of two synthetic populations of 250 agents each (Box 1; Supporting Information S2). In the second example, we replicate the movement pattern of the Painted Bunting *Passerina ciris*, a well-studied migratory songbird occurring in North America and Mexico.

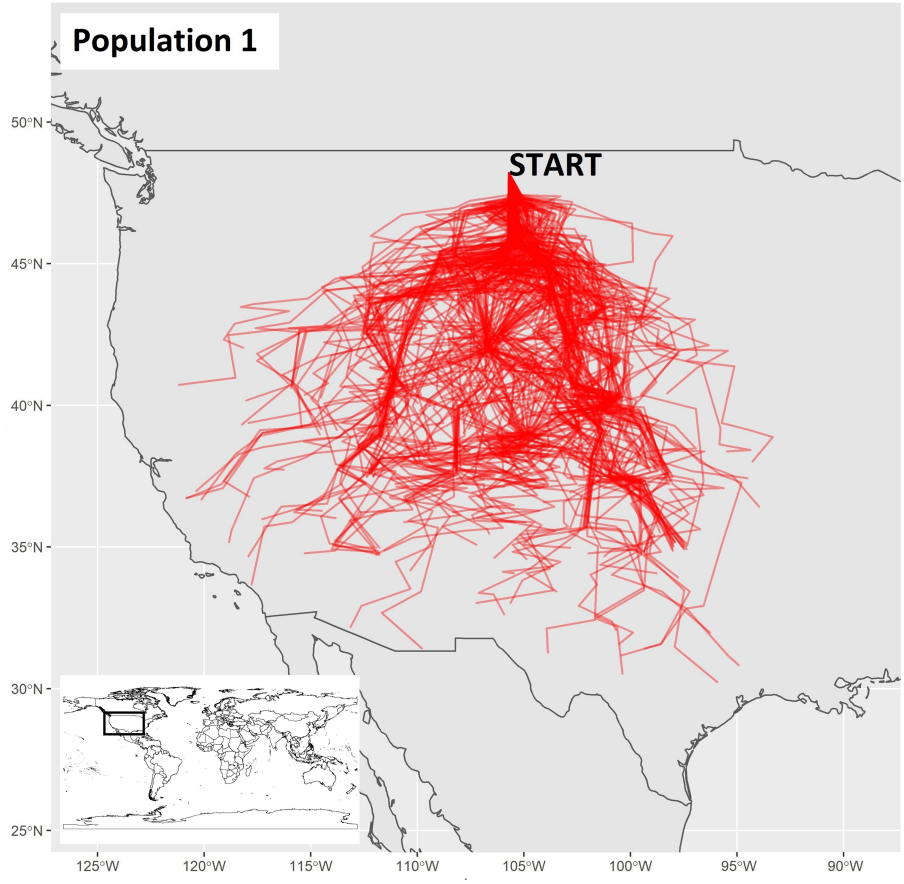
3.1 | Comparisons between populations

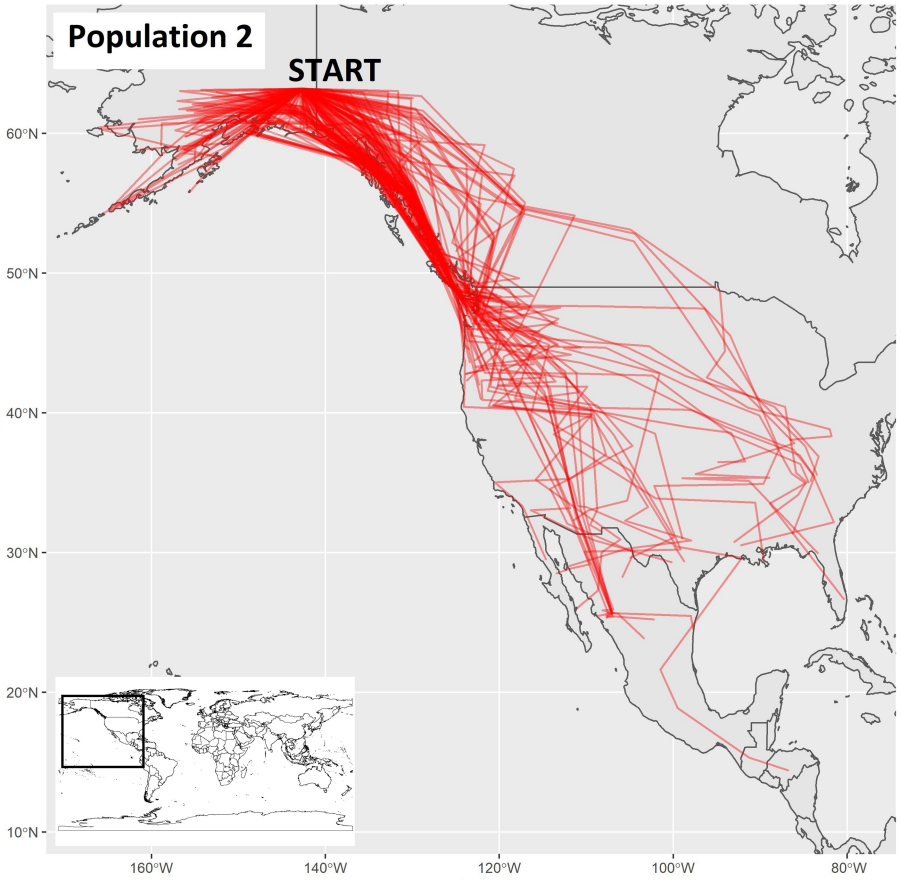
In this example, both synthetic populations are characterized by matching number of replicates and movement timesteps (*days*), equal σ, and the same environmental data provided by a Normalized Difference Vegetation Index (NDVI) raster stack containing 14 days of data between 1–14 September 2019 (Vermote, 2019). Both populations had an unspecified destination (indicated with '999' in the arguments *dest_x* and *dest_y*) and were constrained to move on land. However, Population 1 (P1) agents started their movements from a different point (105.7°W; 48.2°N) situated about 2800 km from the origin of Population 2 (P2) agents (142.7°W; 63.2°N). Additionally, P1 agents had a smaller search radius (150 km) but higher motivation than P2 agents (P1 motivation = 0.95). P1 agents also had different optimum ranges (P1 0.2–0.5; P2 0.6–0.8), and different initial energy units (P1 100; P2 70). These differences in simulation parameterization result in clearly dissimilar movement tracks (Figure 4).

While we can compare the movement tracks visually, Table 3 provides a numerical description of the results. In this simulation, P1 travelled a much smaller average distance (154.6 km) than did P2 (625.5 km). However, P1 travelled more days on average (7.4 days)

FIGURE 4 Movement tracks reveal that Population 1 tended to travel through the Central United States, while Population 2 travelled mostly throughout western Canada, United States, and Mexico. Overall, Population 1 travelled more distance and exhibited more consistent paths near the origin than did Population 2. The movement tracks are natively produced by abmR. Inset world map provided for geographic reference.

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14014, Wiley Online Library on [09/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License





2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14014, Wiley Online Library on [09/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

before stopping than P2 (3.8 days). Additionally, P2 had higher energy consumption than P1; its average remaining energy across all timesteps was 61.2 units compared to 99.7 units for P1. There are several possible reasons for this observed pattern. First, P2 began with a smaller initial energy (70 units) than P1 (100 units). Additionally, P2 had higher optimum NDVI values (0.6–0.8), which might have been less abundant and generally more difficult to reach than those of P1 (0.2–0.5). Finally, because they began in different places, P1 and P2 agents encountered different raster cells along their journey.

Figure 5 visually compares P1 and P2 movement outputs based on longitude and latitude. This is not a native abmR figure, but rather it is produced using the raw data that abmR generates to show the flexible use of the package. In this figure, P1 movements tended to be to the east and south of P2. However, P2 trajectory shows a much wider distribution, with density points extending to the lower values of latitude.

Finally, Figure 6 provides a density surface plot for P2 describing agent energy gains (blue) and losses (red) across the landscape. This surface was created using the inverse distance-weighted interpolation (IDW) function from the R package 'gstat' (Pebesma, 2004). IDW interpolates grid cell values across a surface using a linear combination of observed (sample) points. When interpolating a cell value, the value of the sample points closer to that cell carry a higher weight, while sample points further from that cell carry smaller weight. IDW is discussed in more detail in Wong (2017). The results from Figure 6 match well with what we observe in Figure 4. Movement tracks for P2 tend to follow the blue (energy gain) regions.

3.2 | Migration of the Painted Bunting

The Painted Bunting is a migratory songbird that has been intensively studied over the past two decades due to its steeply declining population trend in the United States and its complex moult-migratory behaviour (Thompson, 1991). This species has been the focus of pioneering light-level geolocator tags that revealed westward post-breeding movements from the southern United States throughout Mexico (Contina et al., 2013), a system for genetic analysis of migratory population connectivity (Battey et al., 2018; Contina et al., 2019) and for candidate genes studies related to the migratory behaviour (Contina et al., 2016). Moreover, Bridge et al. (2016) used an ABM approach to model the post-breeding movements of a Painted Bunting, investigating the association between large-scale vegetation productivity changes and the moult-migratory behaviour. Therefore, this species provides a well-studied migratory system with ample behavioural and population ecology knowledge against which abmR movement predictions can be tested and compared.

We built a basic ABM movement simulation for the Painted Bunting in abmR by using the same key parameters adopted by Bridge et al. (2016) and predicted that a westward migratory movement would emerge throughout the southern U.S. and the coastal regions of Mexico in late summer. Bridge et al. (2016) found that Painted Bunting agents beginning their migratory journey in Oklahoma in late August and September tend to avoid a direct southern migration and show a clear pattern towards southwestern movements targeting high primary productivity areas in northern Mexico and Sinaloa (northwestern Mexico). We used the identical model start location (−98.8°W; 34.8°N) as Bridge et al. (2016), a similar subset of 14 vegetation index raster files (NDVI; Vermote, 2019) representing primary productivity condition during the first 2 weeks of September 2011, and no predefined destination coordinates (indicated with '999' in the arguments *dest_x* and *dest_y*). For the full set of model parameters, see Supporting Information S3. While we note that Bridge et al. (2016) used enhanced vegetation index (EVI) as opposed to NDVI, a multiyear timeframe (2010–2013), and a complex series of model parametrization implemented in ArcGIS, our simulation offers a basic but useful illustration of the predictive capabilities of abmR.

The outcome of 12 abmR Painted Bunting simulations revealed a movement pattern consistent with our predictions based on the results presented by Bridge et al. (2016). Most agents in our simulation (9 out of 12) showed a southwestern movement towards Sinaloa (Mexico) in mid-September (Figure 7), where migratory agents utilize a bloom in primary productivity due to

	Variable	Mean	SD	Median	Min	Max	Range
Population 1	Day	7.4	4	7	1	14	13
	Longitude	−105.5	4.1	−105.2	−121.1	−93.3	27.8
	Latitude	41.6	3.4	41.9	30.2	47.5	17.3
	Energy	99.7	1.3	100	80	100	20
	Delta energy	−0.04	1.5	0	−20	20	40
	Distance	154.6	70	154.4	5.5	316.7	311.1
Population 2	Day	3.8	3.08	3	1	14	13
	Longitude	−126	17.3	−127.1	−166	−80.3	85.7
	Latitude	50.2	10.5	52.2	14.4	63.1	48.7
	Energy	61.2	17.2	60	0	100	100
	Delta energy	−2.3	9.3	−5	−20	20	40
	Distance	625.5	294	649.7	5.5	1445	1439.4

TABLE 3 A numerical comparison of Populations 1 and 2, created by using values across all timesteps for all agents. 'Day' summarizes the timestep variable of the movement tracks. 'Longitude' and 'latitude' summarize the geographical position of agents, while 'energy' summarizes agents' remaining energy. 'Delta energy' corresponds to the change (gain or loss) of energy between each timestep, while 'distance' refers to the distance travelled between each timestep. This table was produced outside of abmR using raw movement data returned by the package.

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14014, Wiley Online Library on [09/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

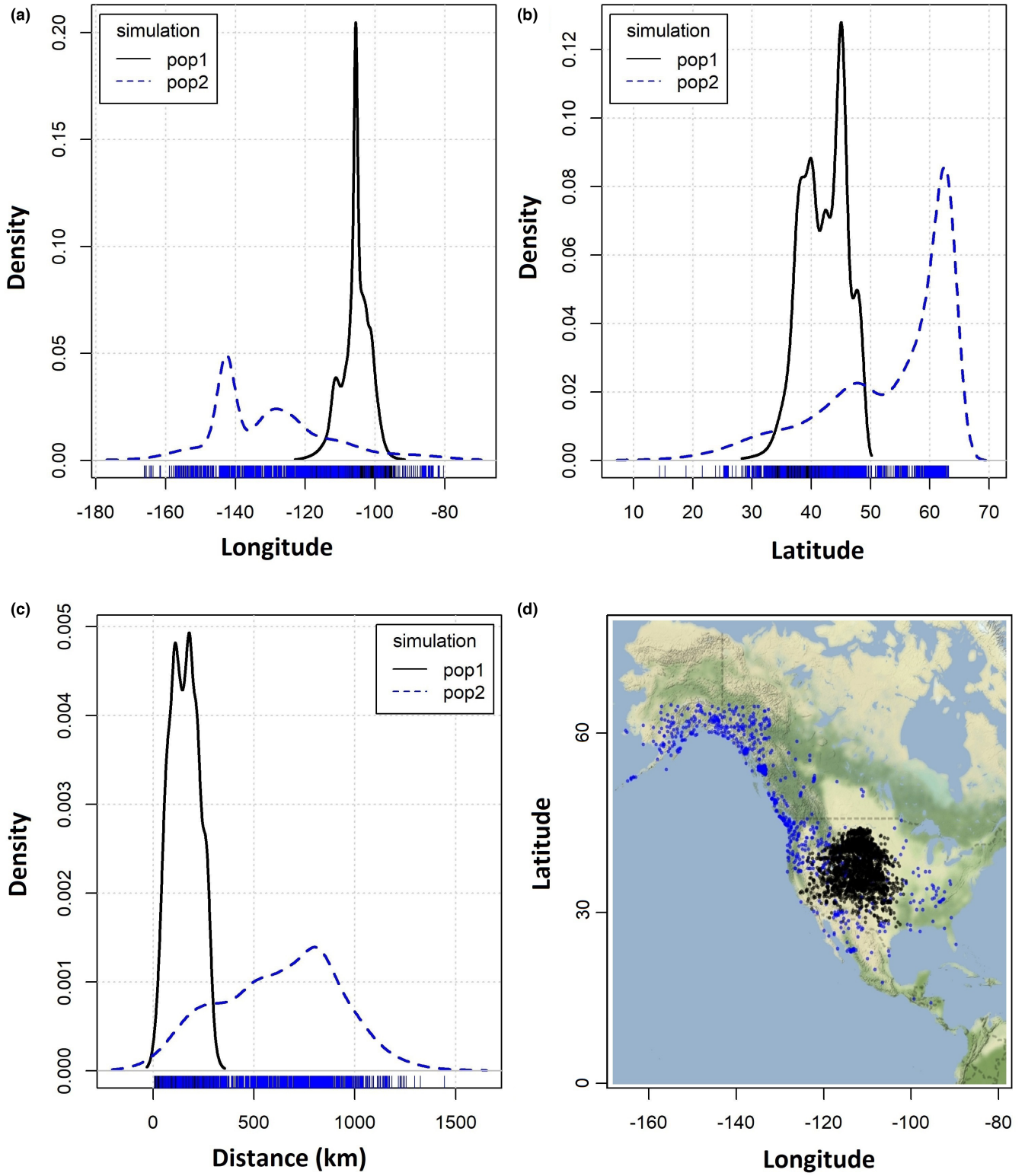


FIGURE 5 Graphical comparisons of Population 1 and Population 2 movements. Panels a and (b) show density plots used to individually compare longitude (panel a) and latitude (panel b) coordinates attained by agents from each population. Panel (c) compares the distance travelled between each timestep, while panel (d) shows geographical position for all agents in each population across all timesteps.

the Monsoonal precipitations (Rohwer et al., 2005). Three agents showed a southeastern movement, following vegetation changes along southeastern United States and the Gulf of Mexico. This result is also in line with sporadic but nonetheless documented variations in migratory strategies in the Painted Bunting where some individuals that breed in south-central United States (e.g. western Oklahoma and Arkansas) show southeastern movements in late summer (Contina et al., 2013).

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14014, Wiley Online Library on [09/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

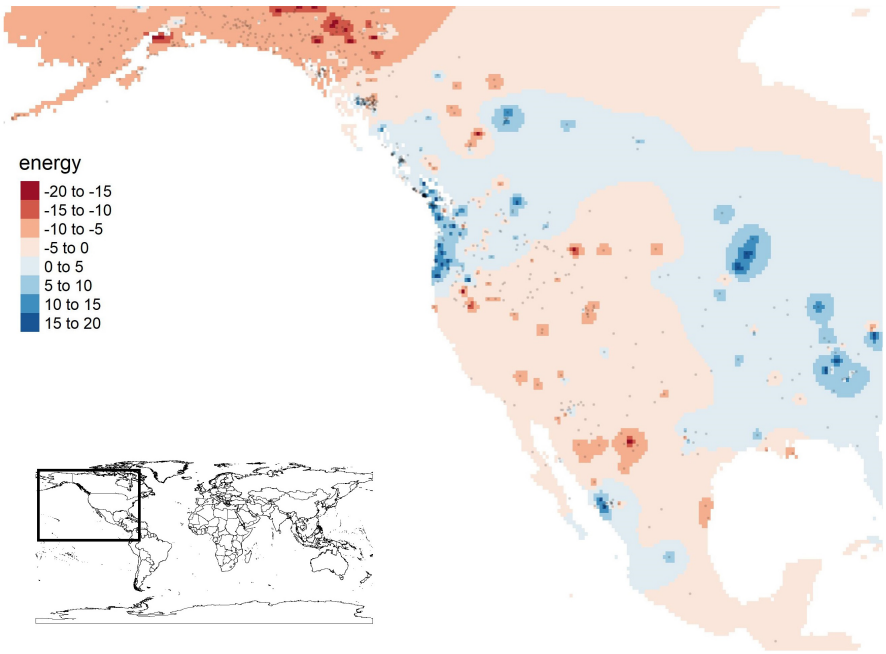


FIGURE 6 Energy gradient plot of Population 2 by timestep. Areas in red reflect energy loss (less suitable environmental values) while areas in blue reflect energy gain (better environmental values). This plot is produced directly by energyVIZ. Inset world map added for geographic reference.

4 | CONCLUSIONS AND FUTURE WORK

abmR provides a novel and efficient programming platform for simulating large-scale movements of species across taxa. We ran most of the initial test simulations on a local machine equipped with an Intel® Core™ i7-5500U CPU—2.40GHz and 8 GB of RAM and obtained results for 100–1000 agents within minutes. The novelty of the software includes the capability of concurrently modelling agent movement trajectories and energy budget. This feature enables a broader exploration of the ecological constraints that shape animal dispersal and/or migration. Moreover, abmR built-in arguments, such as *fail_thresh*, *n_failures*, and *energy_adj*, provide additional flexibility when evaluating mortality scenarios that depend on baseline environmental conditions and energy requirement during prolonged movement bouts (see Table 2 for a full list of arguments affecting mortality).

Over the last decades, spatially explicit simulations, pattern-oriented modelling, approximate Bayesian computing, and ABMs have become more popular in ecological and evolutionary studies (DeAngelis & Grimm, 2014; Gallagher et al., 2021; Railsback et al., 2006; van der Vaart et al., 2015). Analytical platforms, such as InSTREAM, a simulation model approach designed to understand how stream and river salmonid populations respond to habitat alteration (Railsback et al., 2009), or ALMaSS, a predictive modelling tool for answering environmental policy questions regarding the effect of changing landscape structure on threatened animal species (Topping et al., 2003), allow investigation of specific ecological systems using ABM. On the other hand, many programming languages such as Netlogo, R, or Python are widely used to develop custom and more flexible models that can be adapted to address complex ecological or evolutionary research scenarios (Chubaty & McIntire, 2021; Lustig et al., 2019). However, the use of a programming language to develop a flexible ABM from scratch has two important drawbacks. First, it requires advanced programming skills. Second, its reproducibility can be compromised by the idiosyncrasies of the simulation algorithm written by the user. These idiosyncrasies, especially if not well documented, can make it difficult or even impossible for other researchers to replicate findings or adapt code to suit their modelling scenarios. abmR provides a novel framework to perform complex movement simulations through standardized functions and arguments that facilitate model annotation and reproducibility while providing publication-ready visualizations at the end of each run.

While we developed and tested abmR as a movement and energy budget simulation tool, its core software functionalities can be adapted to explore other processes such as disease outbreak scenarios (Dougherty et al., 2018). As an example, pathogen vector movement can be easily simulated within abmR, allowing the study of areas of confluence where disease transmission is more probable (Manore et al., 2015). Moreover, potential future updates will include the ability to specify multiple raster stacks of different movement predictors and interactions between agents. In abmR, each simulation output can be used as the input for the next movement model. However, the option of computing agent interactions affecting movement patterns within the same simulation run is currently missing. This is a clear area of further package development. Additionally, other code expansions might be useful to study plant seed dispersal, density-dependent scenarios, and altitudinal movements.

AUTHOR CONTRIBUTIONS

Benjamin Gochanour led the development of the package and wrote the manuscript; Javier Fernández-López contributed to the package, writing, and critical review of the manuscript; Andrea Contina conceived the manuscript, led its writing, contributed to the package, and performed package testing.

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14014, Wiley Online Library on [09/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

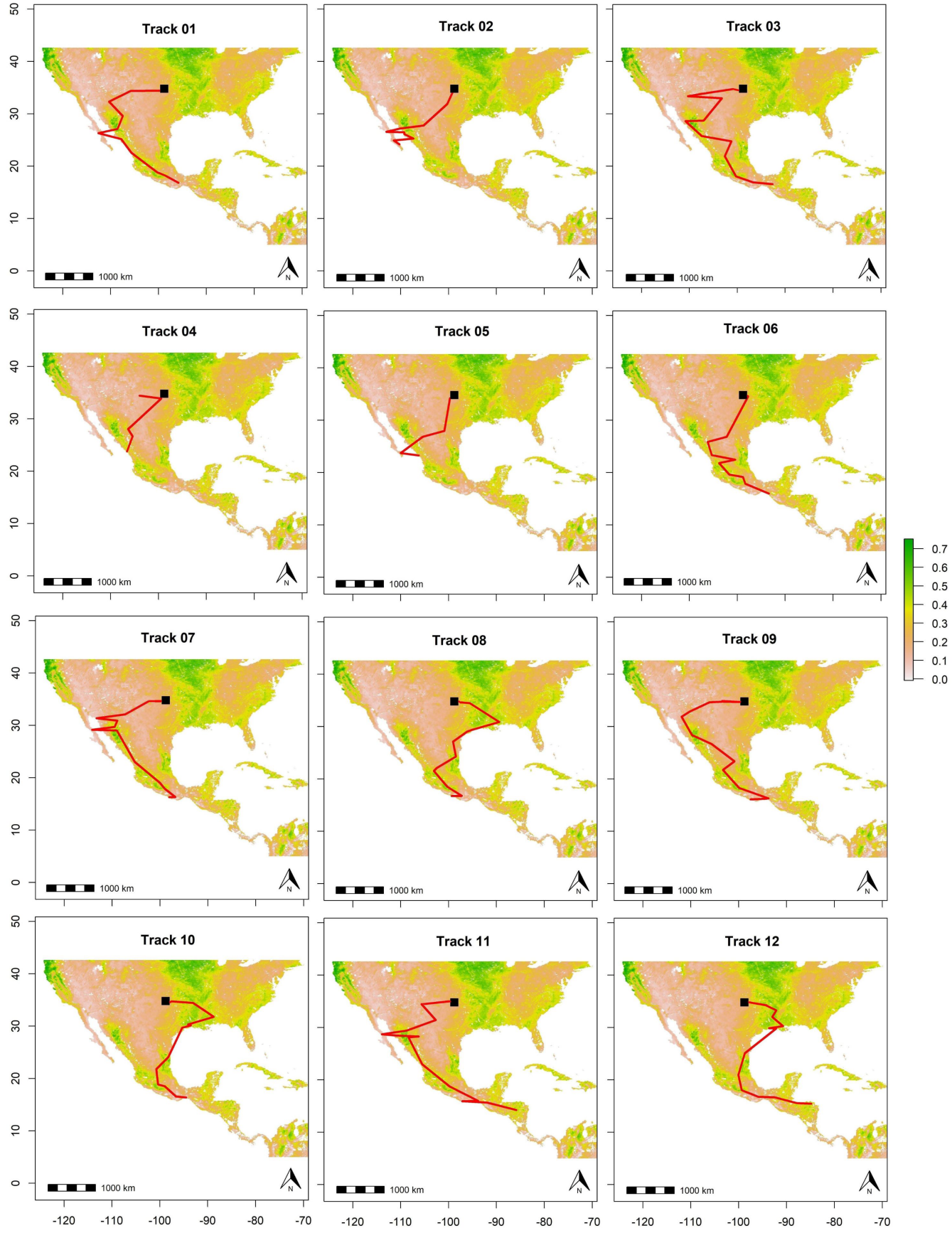


FIGURE 7 Outcome of 12 abmR simulations showing a frequent southwestern migration from the breeding ground in Oklahoma (US) towards wintering grounds in Sinaloa (Mexico) and southern Mexico. The NDVI map in the background is an average of the raster stack object used in abmR, which contained 14 raster layers ranging from September 1 to September 14, 2011.

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14014, Wiley Online Library on [09/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ACKNOWLEDGEMENTS

We thank three anonymous reviewers who provided invaluable feedback on an earlier draft. abmR relies heavily on a large set of R packages, and we thank the R community for providing support and open-source code. We thank Jeff Kelly, Xiangming Xiao, and the Center for Earth Observation and Modelling at the University of Oklahoma for supporting the research effort.

FUNDING INFORMATION

This study was supported in part by research grants from the U.S. National Science Foundation (grants nos. EF 1840230 and DGE 1545261 and DEB 1911955), the National Natural Science Foundation of China (81961128002), and by the Corix Plains Institute.

CONFLICT OF INTEREST

The authors declare that there are no financial or non-financial conflicts of interest.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/2041-210X.14014.

DATA AVAILABILITY STATEMENT

abmR is available on CRAN (https://cran.r-project.org/web/packages/abmR/index.html) and the simulation results for Populations 1 and 2 are available for download from Zenodo at https://doi.org/10.5281/zenodo.7114675 (Gochanour et al., 2022). The environmental raster data set used in the examples is available at https://www.ncei.noaa.gov/data/avhrr-land-normalized-difference-vegetation-index/access/.

ORCID

Benjamin Gochanour https://orcid.org/0000-0002-0950-1739
Javier Fernández-López https://orcid.org/0000-0003-4352-0252
Andrea Contina https://orcid.org/0000-0002-0484-6711

REFERENCES

Araujo, M. B., & Guisan, A. (2006). Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, *33*, 1677–1688.

Aurbach, A., Schmid, B., Liechti, F., Chokani, N., & Abhari, R. (2020). Simulation of broad front bird migration across Western Europe. *Ecological Modelling*, *415*, 108879.

Battey, C. J., Linck, E. B., Epperly, K. L., French, C., Slager, D. L., Sykes Jr., P. W., & Klicka, J. (2018). A migratory divide in the painted bunting (*Passerina ciris*). *The American Naturalist*, *191*(2), 259–268.

Berg, H. C. (1983). *Random walks in biology*. Princeton University Press.

Blomberg, S. P., Rathnayake, S. I., & Moreau, C. M. (2020). Beyond Brownian motion and the Ornstein-Uhlenbeck process: Stochastic diffusion models for the evolution of quantitative characters. *The American Naturalist*, *195*(2), 145–165.

Bridge, E. S., Ross, J. D., Contina, A. J., & Kelly, J. F. (2016). Do molt-migrant songbirds optimize migration routes based on primary productivity? *Behavioral Ecology*, *27*(3), 784–792.

Bridge, E. S., Ross, J. D., Contina, A. J., & Kelly, J. F. (2017). Using agent-based models to scale from individuals to populations. In *Aeroecology* (pp. 259–275). Springer.

Brown, D., & Robinson, D. (2006). Effects of heterogeneity in residential preferences on an agent-based model of urban sprawl. *Ecology and Society*, *11*(1).

Chubaty, A. M., & McIntire, E. J. B. (2021). *SpaDES: Develop and run spatially explicit discrete event simulation models*. R package version 2.0.7. https://CRAN.R-project.org/package=SpaDES

Contina, A., Alcantara, J. L., Bridge, E. S., Ross, J. D., Oakley, W. F., Kelly, J. F., & Ruegg, K. C. (2019). Genetic structure of the painted bunting and its implications for conservation of migratory populations. *Ibis*, *161*(2), 372–386.

Contina, A., Bridge, E. S., & Kelly, J. F. (2016). Exploring novel candidate genes from the mouse genome informatics database: Potential implications for avian migration research. *Integrative Zoology*, *11*(4), 240–249.

Contina, A., Bridge, E. S., Seavy, N. E., Duckles, J. M., & Kelly, J. F. (2013). Using geologgers to investigate bimodal isotope patterns in painted buntings (*Passerina ciris*). *The Auk*, *130*(2), 265–272.

Cullen, J. A., Poli, C. L., Fletcher Jr., R. J., & Valle, D. (2022). Identifying latent behavioural states in animal movement with M4, a nonparametric Bayesian method. *Methods in Ecology and Evolution*, *13*(2), 432–446.

Cushman, S. A., & Lewis, J. S. (2010). Movement behavior explains genetic differentiation in American black bears. *Landscape Ecology*, *25*(10), 1613–1625.

DeAngelis, D. L., & Grimm, V. (2014). Individual-based models in ecology after four decades. *F1000prime Reports*, *6*, 39.

Dingle, H. (2014). *Migration: The biology of life on the move*. Oxford University Press.

Dodge, S., Bohrer, G., Bildstein, K., Davidson, S. C., Weinzierl, R., Bechard, M. J., Barber, D., Kays, R., Brandes, D., Han, J., & Wikelski, M. (2014). Environmental drivers of variability in the movement ecology of Turkey vultures (*Cathartes aura*) in north and South America. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1643), 20130195.

Dougherty, E. R., Seidel, D. P., Carlson, C. J., Spiegel, O., & Getz, W. M. (2018). Going through the motions: Incorporating movement analyses into disease research. *Ecology Letters*, *21*(4), 588–604.

Gallagher, C. A., Chudzinska, M., Larsen-Gray, A., Pollock, C. J., Sells, S. N., White, P. J., & Berger, U. (2021). From theory to practice in pattern-oriented modelling: Identifying and using empirical patterns in predictive models. *Biological Reviews*, *96*(5), 1868–1888.

Giuggioli, L., & Bartumeus, F. (2010). Animal movement, search strategies and behavioural ecology: A cross-disciplinary way forward. *Journal of Animal Ecology*, *79*(4), 906–909.

Gochanour, B., Fernández-López, J., & Contina, A. (2022). Simulation study results of abmR: An R package for agent-based model analysis of large-scale movements across taxa. *bioRxiv*, https://doi.org/10.5281/zenodo.7114675

Goldstein, E., Erinjery, J. J., Martin, G., Kasturiratne, A., Ediriweera, D. S., de Silva, H. J., Diggle, P., Lalloo, D. G., Murray, K. A., & Iwamura, T. (2021). Integrating human behavior and snake ecology with agent-based models to predict snakebite in high risk landscapes. *PLoS Neglected Tropical Diseases*, *15*(1), e0009047.

Grimm, V., & Railsback, S. F. (2013). *Individual-based modeling and ecology*. Princeton University Press.

Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W. M., Railsback, S. F., Thulke, H.-H., Weiner, J., Wiegand, T., & DeAngelis, D. L. (2005). Pattern-oriented modeling of agent-based complex systems: Lessons from ecology. *Science*, *310*(5750), 987–991.

Hawkes, C. (2009). Linking movement behaviour, dispersal and population processes: Is individual variation a key? *Journal of Animal Ecology*, *78*(5), 894–906.

Holdo, R. M., & Roach, R. R. (2013). Inferring animal population distributions from individual tracking data: Theoretical insights and potential pitfalls. *Journal of Animal Ecology*, *82*(1), 175–181.

Joo, R., Boone, M. E., Clay, T. A., Patrick, S. C., Clusella-Trullas, S., & Basille, M. (2020). Navigating through the R packages for movement. *Journal of Animal Ecology*, *89*(1), 248–267.

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14014, Wiley Online Library on [09/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Klimek, P., Poledna, S., Farmer, J. D., & Thurner, S. (2015). To bail-out or to bail-in? Answers from an agent-based model. *Journal of Economic Dynamics and Control*, *50*, 144–154.

Lustig, A., James, A., Anderson, D., & Plank, M. (2019). Pest control at a regional scale: Identifying key criteria using a spatially explicit, agent-based model. *Journal of Applied Ecology*, *56*(7), 1515–1527.

Manore, C. A., Hickmann, K. S., Hyman, J. M., Foppa, I. M., Davis, J. K., Wesson, D. M., & Mores, C. N. (2015). A network-patch methodology for adapting agent-based models for directly transmitted disease to mosquito-borne disease. *Journal of Biological Dynamics*, *9*(1), 52–72.

Merlin, C., & Liedvogel, M. (2019). The genetics and epigenetics of animal migration and orientation: Birds, butterflies and beyond. *Journal of Experimental Biology*, *222*(Suppl_1), jeb191890.

Nathan, R., Getz, W. M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D., & Smouse, P. E. (2008). A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(49), 19052–19059.

Pebesma, E. J. (2004). Multivariable geostatistics in S: The gstat package. *Computers & Geosciences*, *30*, 683–691.

Polhill, J. G., Parker, D., Brown, D., & Grimm, V. (2008). Using the ODD protocol for describing three agent-based social simulation models of land-use change. *Journal of Artificial Societies and Social Simulation*, *11*(2), 3.

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.r-project.org

Railsback, S. F., Harvey, B. C., Jackson, S. K., & Lamberson, R. H. (2009). InSTREAM: The individual-based stream trout research and environmental assessment model. Gen. Tech. Rep. PSW-GTR-218. US Department of Agriculture, Forest Service, Pacific Southwest Research Station.

Railsback, S. F., Lytinen, S. L., & Jackson, S. K. (2006). Agent-based simulation platforms: Review and development recommendations. *SIMULATION*, *82*(9), 609–623.

Rohwer, S., Butler, L. K., Froehlich, D. R., Greenberg, R., & Marra, P. P. (2005). Ecology and demography of east–west differences in molt scheduling of neotropical migrant passerines. In R. Greenberg & P. P. Marra (Eds.), *Birds of two worlds: The ecology and evolution of migration* (pp. 87–105). Johns Hopkins University Press.

Tang, W., & Bennett, D. A. (2010). Agent-based modeling of animal movement: A review. *Geography Compass*, *4*(7), 682–700.

Thiele, J. C., Kurth, W., & Grimm, V. (2012). RNetLogo: An R package for running and exploring individual-based models implemented in NetLogo. *Methods in Ecology and Evolution*, *3*(3), 480–483.

Thompson, C. W. (1991). Is the painted bunting actually two species? Problems determining species limits between allopatric populations. *The Condor*, *93*(4), 987–1000.

Thurfjell, H., Ciuti, S., & Boyce, M. S. (2014). Applications of step-selection functions in ecology and conservation. *Movement Ecology*, *2*(1), 1–12.

Topping, C. J., Hansen, T. S., Jensen, T. S., Jepsen, J. U., Nikolajsen, F., & Odderskær, P. (2003). ALMaSS, an agent-based model for animals in temperate European landscapes. *Ecological Modelling*, *167*(1–2), 65–82.

Uhlenbeck, G. E., & Ornstein, L. S. (1930). On the theory of the Brownian motion. *Physical Review*, *36*(5), 823–841.

van der Vaart, E., Beaumont, M. A., Johnston, A. S., & Sibly, R. M. (2015). Calibration and evaluation of individual-based models using approximate Bayesian computation. *Ecological Modelling*, *312*, 182–190.

Vermote, E., & NOAA CDR Program. (2019). *NOAA climate data record (CDR) of AVHRR normalized difference vegetation index (NDVI), version 5*. NOAA National Centers for Environmental Information. https://doi.org/10.7289/V5ZG6QH9

Willem, L., Verelst, F., Bilcke, J., Hens, N., & Beutels, P. (2017). Lessons from a decade of individual-based models for infectious disease transmission: A systematic review (2006–2015). *BMC Infectious Diseases*, *17*(1), 1–16.

Williams, H. J., Taylor, L. A., Benhamou, S., Bijleveld, A. I., Clay, T. A., de Grissac, S., & Börger, L. (2020). Optimizing the use of biologgers for movement ecology research. *Journal of Animal Ecology*, *89*(1), 186–206.

Wong, D. W. (2017). Interpolation: Inverse-distance weighting. *The International Encyclopedia of Geography*, 1–7.

Yin, S., Xu, Y., de Jong, M. C., Huisman, M. R., Contina, A., Prins, H. H., Huang, Z. Y., & de Boer, W. F. (2022). Habitat loss exacerbates pathogen spread: An agent-based model of avian influenza infection in migratory waterfowl. *PLoS Computational Biology*, *18*(8), e1009577. https://doi.org/10.1371/journal.pcbi.1009577

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Gochanour, B., Fernández-López, J., & Contina, A. (2022). abmR: An R package for agent-based model analysis of large-scale movements across taxa. *Methods in Ecology and Evolution*, *00*, 1–13. https://doi.org/10.1111/2041-210X.14014

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14014, Wiley Online Library on [09/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License