

## ORIGINAL ARTICLE

# Agent-based models reveal limits of mark–release–recapture estimates for the rare butterfly, *Bhutanitis thaidina* (Lepidoptera: Papilionidae)

Zhengyang Wang<sup>1,\*</sup> , Yuanheng Li<sup>1,\*</sup> , Anuj Jain<sup>2</sup>  and Naomi E. Pierce<sup>1</sup> 

<sup>1</sup>Department of Organismic and Evolutionary Biology and Museum of Comparative Zoology, Harvard University, Cambridge, MA 02138, USA and <sup>2</sup>Nature Society (Singapore), Singapore, Singapore

**Abstract** Insect diversity and abundance are in drastic decline worldwide, but quantifying insect populations to better conserve them is a difficult task. Mark-release-recapture (MRR) is widely used as an ecological indicator for insect populations, but the accuracy of MRR estimates can vary with factors such as spatial scale, sampling effort and models of inference. We conducted a 3-year MRR study of *B. thaidina* in Yanzigou valley, Mt. Gongga but failed to obtain sufficient data for a robust population estimate. This prompted us to integrate *B. thaidina* life history information to parameterize agent-based models and evaluate the conditions under which successful MRR studies could be conducted. We evaluated: (1) the performance of MRR models under different landscape types, and (2) the influence of experimental design on the accuracy and variance of MRR-based estimates. Our simulations revealed systematic underestimates of true population parameters by MRR models when sampling effort was insufficient. In a total of 2772 simulations, subjective decisions in sampling protocol (e.g., frequency, number of sampling locations, use of spatially explicit models, type of estimands) accounted for nearly half of the variation in estimates. We conclude that MRR-based estimates could be improved with the addition of more field-specific parameters.

**Key words** agent-based simulation; *Aristolochia moupinensis*; habitat; hostplant; landscape; Mt. Gongga; program MARK; spatially explicit capture–recapture (SECR)

## Introduction

Insect diversity and abundance are in drastic decline worldwide (Hallmann *et al.*, 2017; Sánchez-Bayo & Wyckhuys, 2019; van Strien *et al.*, 2019), but quantifying insect populations to better conserve them is a difficult task (Haddad *et al.*, 2008; Didham *et al.*, 2020; Welti *et al.*, 2021). An accurate quantification of popu-

lation size is the first step to effective conservation management (Sutherland, 1996). Field conservationists have traditionally relied on count-based inferences from traps or transects as indicators of relative insect abundance (Gross *et al.*, 2007; Nowicki *et al.*, 2008; Hamm, 2013). Compared with these methods, mark–release–recapture (MRR) requires more costly, long-term fieldwork, but has been widely implemented since Craig (1953) and has been described as the “gold standard” in estimating insect populations (Haddad *et al.*, 2008).

Although statistical models of MRR have been well-studied (Jolly, 1965; Seber, 1965; 1970, 1982, Pollock *et al.*, 1990; Schwarz & Arnason, 1996), fewer studies have been carried out to evaluate the accuracy of population estimates using these methods (see Sandercock, 2006; Lindberg, 2012 for review). The accuracy of a

Correspondence: Zhengyang Wang and Yuanheng Li, Department of Organismic and Evolutionary Biology and Museum of Comparative Zoology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA. Email: zhengyang-wang@g.harvard.edu; yhli@fas.harvard.edu

\*Zhengyang Wang and Yuanheng Li contributed equally to this work.

population estimate is hard to verify since we can seldom count all the individuals in a population (with rare exceptions, see Donkers *et al.*, 2012 on common carp). Validation of MRR estimates, sometimes feasible in birds, fish, and insects, has focused primarily on analyzing subsampled data (Ponchon *et al.*, 2018; Turlure *et al.*, 2018) or comparing MRR with count-based estimation methods (Evans, 2004; Haddad, 2008; Hain *et al.*, 2016). Many validation studies have found that the results of MRR-based estimates vary with factors such as the spatial scale sampled, the effort in sampling, and the downstream analysis tool used (Banks & Brown, 1962; Schneider, 2003; Schtickzelle *et al.*, 2003; Evans, 2004; Rosenberger & Dunham, 2005; Haddad *et al.*, 2008; Pellet *et al.*, 2012). For example, Hain *et al.* (2016) showed that while individual mark-recaptures of fish resulted in underestimates of population sizes compared with visual estimates, batch MRR data (with higher numbers of recaptures) did not. On the other hand, Nowicki *et al.* (2005) showed that MRR sampling schemes to estimate butterfly population size can be simplified while retaining equally efficient estimates. These field-based results suggest that sampling design plays an important role in MRR-based estimates.

Recently, progress has also been made with a set of methods involving “spatially explicit capture–recapture” (SECR) (Efford, 2004; Borchers & Efford, 2008; Efford & Boulanger, 2019). These methods model the home range of animals and their density-dependent detection probabilities. In simulations, SECR has been shown to be more effective than traditional MRR (Efford & Fewster, 2013), but SECR requires more rigid sampling designs involving multiple capture points (Kristensen & Kovach, 2018). Although popular in vertebrate mark-recapture research (Ruiz *et al.*, 2013; Jimenez *et al.*, 2019; Green *et al.*, 2020), it has been less practical to implement MRR surveys in some insect species, particularly those with shorter lifespans that are difficult to trap and mark across multiple encounters. Spatially explicit models of animal movement also underscore the complexity of accounting for animal dispersal across heterogeneous landscapes (Ovaskainen, 2004). Understanding the effect of heterogeneous or fragmented landscapes on population inference is particularly important, as many insects have specific hostplants and microhabitat ranges (Ovaskainen *et al.*, 2008). Appropriate sampling designs such as whether to sample within a given habitat (Dupont *et al.*, 2021) depend upon individual species or populations.

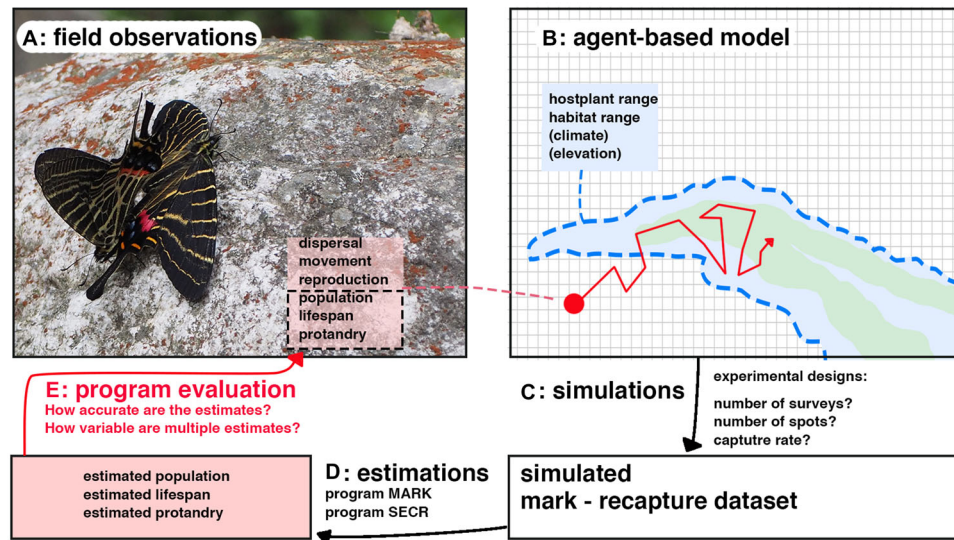
Conservation biologists aiming to estimate insect population size thus need to design a sampling scheme with sufficient power to estimate abundances, but they often don't know their estimand *a priori* to justify a logis-

tic “sufficiency.” Furthermore, what if the insect species that requires assessment is rare and hard to capture? Although MRR is generally recognized as an effective method to measure butterfly population size (Nowicki *et al.*, 2008; Haddad *et al.*, 2008; Pellet *et al.*, 2012; Kral *et al.*, 2018) and has been applied to several endangered, flagship species in relatively isolated habitats (Ferster & Vulinec, 2010; Nowicki *et al.*, 2019), these studies involve locally abundant populations. In practice, many endangered insects are rarely encountered in the wild and are thus understudied. Lack of sufficient capture (or recapture) may preclude accurate population estimates for endangered insects.

Agent-based models—models that specify individual behaviors and interactions across a predefined landscape (Myers, 1976; DeAngelis, 1992; DeAngelis & Mooij, 2005; DeAngelis & Grimm, 2014)—can help address some of these complexities. These models have been increasingly used to analyze large-scale, spatially explicit, individual-process-based biological problems, from epidemiology to insect social structure to human cooperation (Lewis *et al.*, 2014; Eckhoff *et al.*, 2016; Crall *et al.*, 2019). In a conservation management context, agent-based models have been used to model hostplant conditions of endangered insects (Griebeler & Seitz, 2002) and count-based estimates of arthropod population size (Petrovskii *et al.*, 2012, Ahmed & Petrovskii, 2019), but have not been applied to validate MRR-based estimates.

In this study, we wanted to estimate the abundance of the near-threatened swallowtail butterfly, *Bhutanitis thaidina* (family Papilionidae, IUCN Red List, Li 2019). The species is endemic to the mountain ranges of southwestern China (Yi *et al.*, 2011, Gao *et al.*, 2014, Igarashi & Harada, 2015) and is listed as a “Class II protected species.” Its populations are under severe threat of climate change-induced habitat loss (Hu *et al.*, 2019). We performed mark-recapture on a population of *B. thaidina* inside Yanzigou valley, Mt. Gongga, China. Previous telemetry work on a population of golden birdwing butterflies *Troides aeacus* (family Papilionidae) in the same valley showed that individual butterflies disperse along the valley but not across the high-elevation mountains that flank it (Wang *et al.*, 2019), suggesting that similarly large butterflies may remain in relatively restricted habitats where their population parameters are likely to be consistent with MRR model assumptions. A survey of *B. thaidina* populations represents a technical challenge due to the rarity of the butterfly and the inaccessibility of its mountainous terrain.

We first investigated whether a 3-year MRR study of *B. thaidina* within the constraints of our field budget (a total of 141 person-days) was enough to obtain



**Fig. 1** Schematic for using agent-based models to evaluate MRR-based estimates. (A–B) Agent-based models are parameterized with field-based measurements of the life-history and movement parameters of *B. thaidina* and its hostplant (green) and habitat (blue) range. (C) Simulations of MRR sampling strategies (different frequencies and survey spots) on different landscape types. (D) Simulated MRR results are analyzed with a standard pipeline, either in program MARK or spatially explicit SECR. (E) Model estimates are compared with parameters in agent-based models to evaluate the accuracy and variance of different models under different experimental designs.

sufficient data for a robust population estimate. Our unsatisfactory results underscored the intrinsic difficulties in assessing elusive mountainous butterfly populations, and prompted us to integrate *B. thaidina* life history to parameterize agent-based models and evaluate the conditions under which successful MRR studies could be carried out. Specifically we asked: (1) Do estimates from MRR models perform differently under different landscape types (e.g., would MRR conducted in a closed habitat such as a lake produce more accurate estimates than MRR conducted in a heterogeneous habitat, such as that of an elevationally restricted insect)? (2) How much do differences in field sampling design affect the accuracy and variance of MRR-based estimates (e.g., what are the consequences of increasing MRR sampling frequency vs. increasing MRR sampling locations, using traditional MRR-based models vs. spatially explicit models, and estimating population number vs. estimating population protandry and individual lifespan)? Our results regarding the efficiency of MRR-based estimates are widely applicable to all MRR-based studies, especially ones that are similar in range, spatial scale and movement parameters (e.g., many arthropods and small vertebrates). Lastly, we returned to our field-collected MRR data and tried to improve our estimates of *B. thaidina* populations based on general patterns of bias learned from agent-based MRR simulations.

## Materials and methods

See Figs. 1 and S1 for visual summaries of our workflow.

### Study area

Mark-recapture studies were conducted inside Yanzigou glacier valley, Mt. Gongga. The region is topologically characteristic of the “region of extreme relief” (Irving & Hebda, 1993) of the Hengduan mountains—a deep valley flanked by high mountains, providing a restricted environment for butterflies. Vegetation along the valley ranges from alpine grassland (3800 m a.s.l.), *Rhododendron* forest (3200 m a.s.l.), temperate forest (2000–3000 m a.s.l.) to crop fields (1900 m a.s.l.) of a local village. A population of *Bhutanitis thaidina* was discovered in Yanzigou valley in 1985 (Igarashi, 2003). The larvae feed on *Aristolochia moupinensis*, which can be found on edges of disturbed temperate forests inside Yanzigou valley. Another threatened, congeneric species, *Bhutanitis mansfieldi* occurs within the valley in early April (before *B. thaidina*), and its larvae feed on the same hostplant (see Zhang *et al.*, 2019 for life-history comparisons of *Bhutanitis* species). Two other species, *Troides aeacus* and *Byasa* sp., also feed as larvae on *Aristolochia*, and their adults emerge in June and July

(after *B. thaidina*); they are strong flyers and can be seen at elevations lower than those of habitats suitable for *Aristolochia moupinensis* hostplants.

#### Mark-recapture

Mark-recapture of *B. thaidina* was carried out on non-rainy days from May 20 to June 10 in three consecutive years, 2016–2018. Mark-recaptures were performed in a single 30 m × 30 m habitat (2230 m a.s.l.) where vegetation had previously been cleared to build a cow shelter that was abandoned three decades ago. While adults of *B. thaidina* can be observed flying across the temperate forest canopy in this region, they can only be effectively marked and recaptured in this relatively open patch. Each MRR session lasted from 08:00 to 17:30, with two people taking notes and one person catching. Individuals of *B. thaidina* were caught with 5 m long entomological nets, marked on the underside of the left hindwing with a unique number using a waterproof pen, and released. We noted the time at which each individual was captured and its sex and wing condition upon capture (designated as “perfect,” “slightly damaged,” “damaged,” and “severely damaged”). Temperature and humidity were measured every 30 min. We counted all *B. thaidina* individuals flying across the field (but missed by the catcher) and noted their behavioral state in each case. The behaviors were classified as “hill-topping” (traversal of the observation spot from low to high elevation, including male-male fighting), “flying” (non-“hill-topping” traversal), “mating,” “nectaring,” and “resting.” When an individual was observed to fly through the habitat, we measured the time needed to cross the habitat (30 m in width) to estimate the speed of *B. thaidina* movement.

#### Hostplant and habitat transects

Local stakeholders informed us that the habitat range for *B. thaidina* (i.e., the elevational range within which butterflies can be observed in Yanzigou valley) was between 2000 and 2400 m a.s.l. and that their hostplant range (the elevational range within which *Aristolochia moupinensis* can be found) was between 2200 and 2400 m a.s.l. To verify these accounts, in June 2017, we surveyed 10 transects within the hostplant range (2200–2400 m a.s.l.), and 8 transects within the habitat range but outside the hostplant range (2000–2200 m a.s.l.). Each transect was 200 m long and 5 m wide traversing 30 m difference in elevation. All hostplants within the transects were recorded. We also attempted to look for *B. thaidina* during these transect surveys. Since *Aristolochia* host-

plants are more likely to occur in disturbed forests, we also surveyed *Aristolochia moupinensis* along the 15 km road connecting Yanzigou glacier (3800 m a.s.l.) to the village (1900 m a.s.l.).

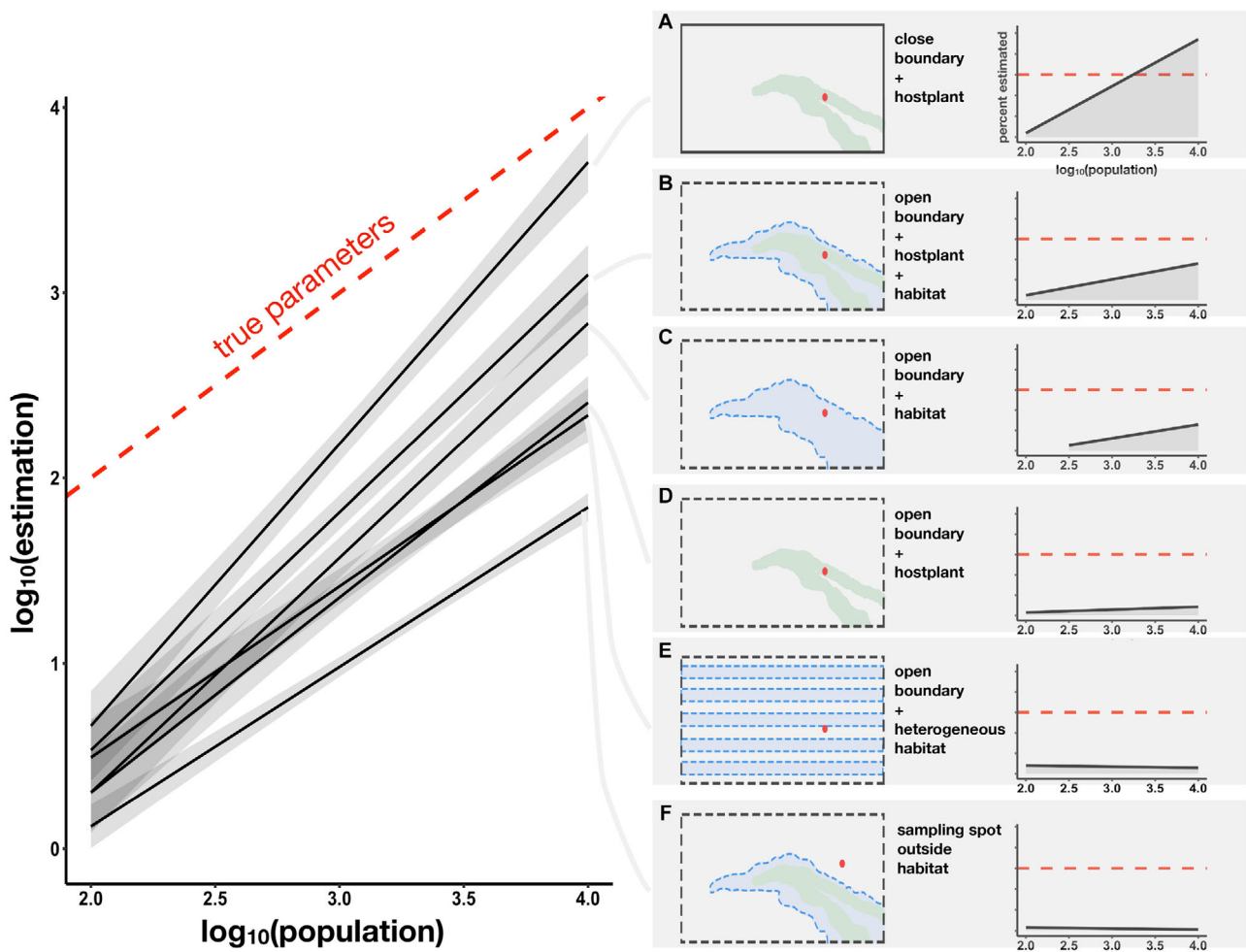
#### Agent-based model

Based on our field measurements, we developed an agent-decision-based, discrete-time model to simulate the MRR sampling procedure of *B. thaidina* in Yanzigou valley (Table S1). The model grid consisted of 221 × 140 cells, each associated with an elevational value projected from the map of Yanzigou valley (27.846 km<sup>2</sup>). Each cell represented a 30 m × 30 m survey spot. The model resolution was set at 15 s per time step.

**Individual movement:** Each individual in the model was characterized by sex (female/male), date of emergence (normal distribution within the first 30 days, SD = 4.5) and life span (normal distribution from 14 to 35 days, SD = 3.2). Individuals only emerged in cells located in elevations within the “hostplant range.” The boundaries of the grid could be either open or closed. In the open-grid setting, any individual could disperse out of the grid freely: if the individual was outside the grid continuously for one hour (240 timesteps), it was removed from the model (treated as dead). In other words, individuals exited the model either at death, or after dispersal outside the grid. At each timestep, the probability of each individual moving to a different cell was determined by a “flight-rest ratio” parameter that changed according to the time of the day. The parameter reflects the change in a butterfly’s activity level during the day. An individual is more likely to move into a cell within the “habitat range” than into a cell not suitable as habitat. The likelihood difference is determined by a “habitat attraction” parameter, which simulates a butterfly’s edge-mediated dispersal out of the habitat (Schultz & Crone, 2001). If the adjacent cell types are the same, then the movement choice is random.

**Population:** The total number of individuals to emerge in the model is set by the “population” parameter. The male and female ratio is decided by the “sex ratio” parameter. When the “protandry” parameter is set above zero, male individuals emerge earlier than females. Both male and female emergences still follow a Gaussian distribution, but the mean emergence time of females is delayed.

**In silico sampling:** Any number of cells in the grid can be designated as sampling cells. In each sampling cell, the frequency of sampling per week can be adjusted by the “sampling effort” parameter. When a cell is being sampled, information about the portion of individuals



**Fig. 2** Program MARK estimation of in silico MRR data of sampling a single spot once per week for 2 months, on six different landscape types performs best on closed-boundary landscapes and habitat-delineated landscapes. The left panel shows the linear regression of estimation results on different landscape types as population parameter of the agent-based model increases. Under the tested population parameters, all estimates are underestimates (below the dotted red line). The expanded panels on the right indicate change in the accuracy of predictions (measured as estimation divided by the population parameter for each population parameter) for each landscape type: (A) a “flat” landscape with closed boundaries; (B) an open landscape with hostplant range and habitat range delineated by empirical data; (C) an open landscape with only habitat range delineated by empirical data; (D) and open landscape with only hostplant range delineated by empirical data; (E) an open landscape with homogenized strips of habitat and nonhabitat; (F) an open landscape with hostplant range and habitat range delineated by empirical data but sampling cell set outside the habitat range. Dotted redlines indicate a 0.3 accuracy threshold. Both bounded landscape and open landscape with habitat and hostplant delineation show significant increases in estimation accuracy as population sizes increase.

(determined by the “capture rate” parameter) that move into the sampling cell is logged (“captured”).

#### Agent-based simulations

We first tested how sensitive MRR estimates are to variations in parameters that are intrinsic to the life-history of *B. thaidina* (see Supplementary Material).

Then, to compare the accuracy of MRR-based population estimates with the true population parameters under different landscape assumptions, we simulated single-location, 2-month-long, once-per-week MRR studies with five population size parameters ( $10^2$ ,  $10^{2.5}$ ,  $10^3$ ,  $10^{3.5}$ ,  $10^4$ ), under six different landscape types (Fig. 2). In the first five landscape settings, the sampling cell is inside the hostplant range: (1) a “flat” landscape with closed boundaries; (2) an open landscape with hostplant

range and habitat range delineated by empirical data; (3) an open landscape with only habitat range delineated empirically; (4) an open landscape with only hostplant range delineated empirically; (5) an open landscape with homogenized strips of habitat and nonhabitat. We also tested (6) an open landscape with hostplant range and habitat range delineated by empirical data, but sampling cells set outside the habitat range. Each simulation was repeated 10 times.

To study the accuracy and variance of mark-recapture estimates under limited conservation resources, we used the grid representing an open landscape with or without hostplant range and habitat range delineated by empirical data (i.e., second and third landscape type in the previous simulations) and simulated three sets of exponentially increasing populations ( $3 \times 10^2$ ,  $3 \times 10^3$ ,  $3 \times 10^4$ ). We studied two strategies of allocating conservation resources: (1) under an “increased sampling spot” strategy, we increased the number of sampling cells from one to seven but kept sampling effort constant at once per week and (2) under an “increased sampling effort” strategy, we increased the sampling frequency from once per week to seven times per week, but kept the number of sampling cells constant at only one per week. Each simulation ran for the time step equivalent of 2 months with 10 repeats. To study the accuracy and variance of temporal estimates based on MRR results, we repeated the above simulations but increased protandry from 0 to 15 days.

### Analysis

Simulated mark-recapture results under various landscapes and sampling schemes were formatted into “encounter history” tables (same-day recaptures were discarded) and fitted with a POPAN model, as implemented in program MARK (Cooch & White, 2006), which is a robust parameterization of the Jolly–Seber model (Schwarz & Arnason, 1996). Models that provide estimates for population size ( $N$ ) had their apparent survival ( $\varphi$ ) and capture probabilities ( $p$ ) set as constant over time, consistent with how individuals in agent-based models are parametrized. We estimated lifespan as  $(1/-\ln[\varphi])$ , adjusted by unit time between each sampling event. When estimating protandry, the population sizes of simulated male and female individuals were analyzed separately. Protandry is calculated as the difference between the time at which the probability of entry into the population ( $pent$  parameter in POPAN model) peaks in male and female groups.

Empirical mark-recapture results from 2016 to 2018 were analyzed with a POPAN model without fixing  $\varphi$

and  $p$  across encounter intervals; estimates for population size ( $N'$ ) were selected from the model with the lowest Akaike Information Criterion adjusted for small sample size (Hurvich & Tsai, 1989). Individual lifespan and population protandry estimates were not included in the analysis for empirical data due to inconsistent temporal sampling.

Simulations with more than two sampling cells were fitted with a spatially explicit capture–recapture model (SECR, as implemented in the “secr” package in R [Efford, 2019]) to estimate population density per hectare. Each sampling cell was modeled as a “multicatch” trap with “half normal” detection function and  $4\sigma$  buffer width. Density estimates in the model were scaled to our simulated grid size to obtain a population estimate.

Each estimation from simulated mark-recapture results under the different parameters by the POPAN and SECR models were evaluated for both their accuracy and variance: (1) for each of the 10 simulations under the same set of parameters, we analyzed the deviation of the average estimated value from the true parameter (“normalized error”) and (2) the variation of these deviations (“normalized variance”). To look at parameters that have the greatest effect on estimation accuracies and variations, we conducted multivariate multiple regression analysis (R core team, 2019) on all sets of parameters (e.g., population, landscape types, sampling strategies), estimation methods (i.e., POPAN vs. SECR model) and estimands (i.e., population, protandry, and lifespan).

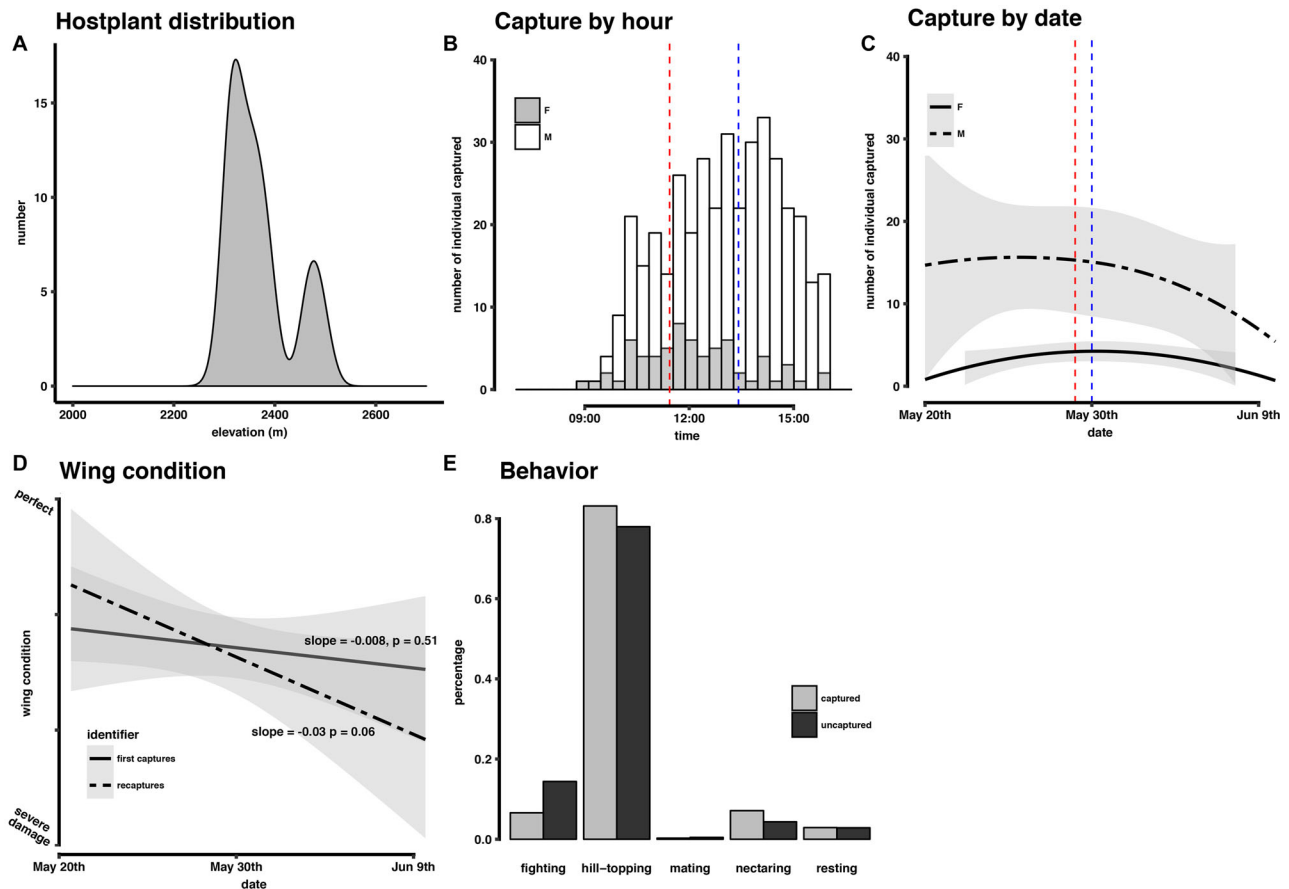
### Adjusting field estimates

To adjust for underestimates of *B. thaidina* population based on field-collected MRR data, we simulated sampling schedules from 2016 to 2018 (Table S2) in silico under different population parameters. The resulting simulated MRR data were used to generate new estimates using the program MARK (see previous section). The prior population parameters that generated population estimates closest to field-based estimates of the year were selected as the “adjusted estimation.”

## Results

### Field measurements

A total of 56 *Aristolochia moupinensis* hostplants were discovered along 10 transects (mean = 5.4, SD = 5.81) between 2200 and 2400 m a.s.l., while no hostplants were seen at low elevation transects from 2000 to 2200 m a.s.l. No *B. thaidina* were seen during a count-based



**Fig. 3** Field measurements used to parametrize agent-based models. (A) Distribution of 105 *Aristolochia moupinensis* hostplants found in transect and road-side surveys, from 1900 to 3800 m a.s.l. Hostplant range is restricted between 2200 and 2500 m a.s.l. (B) The time of the day of 460 capture events as a proxy for the activity level of *B. thaidina* during the day. Blue and red dotted lines indicate the Loess-smoothed activity peak for males and females. (C) The date of the 460 capture events as a proxy for emergence time of *B. thaidina*. Blue and red dotted lines indicate the Loess-smoothed emergence peak for males and females. (D) Deterioration of wing condition across time between first-captured *B. thaidina* and recaptured *B. thaidina* individuals. (E) Observed behavior of captured *B. thaidina* individuals (recorded upon capture) and free-flying *B. thaidina* individuals.

transect survey. Another 49 hostplants were discovered between 2282 and 2505 m a.s.l. along the road connecting Yanzigou glacier (3800 m a.s.l.) to the village (1900 m a.s.l.) (Fig. 3A). Despite the absence of hostplants, both researchers and local stakeholders have observed *B. thaidina* at elevations as low as 2000 m a.s.l., but no higher than 2400 m a.s.l.

We estimated *B. thaidina* flight speed to be 1.86 m/s ( $n = 5$ ,  $SD = 0.429$ ). The fieldwork time devoted to *B. thaidina* mark-recapture each year spanned from 12 to 19 days, with weather-permitted sampling days ranging from 4 to 13 days, nonuniformly distributed throughout the sampling period (Table S2). During mark-recapture experiments, all *B. thaidina* were caught between 08:55 to 15:58, but were most abundant between 12:00 and 14:00

(Fig. 3B). Individuals of *B. thaidina* could be confidently identified as far as 15 m from the observer. When an individual was detected, it was captured 32% of the time in 2017 and 28% of the time in 2018 (relevant data were not recorded in 2016). On average, 117.33 ( $SD = 49.94$ ) individuals were captured each year and the recapture rate was 6.76% ( $SD = 2.27$ ). Among a total of 460 capture events, more than 83% of the butterflies were caught while hill-topping, although we also observed mating, resting, and nectaring behavior in *B. thaidina* (Fig. 3E). We observed no significant differences in the behavior of captured and uncaptured butterflies (modeled as variable in multiple regression:  $t$  value = 1.43,  $P = 0.23$ ). The mean female-male sex ratio observed in the field was 0.17 ( $SD = 0.10$ ). Across 3 years, capture of both

sexes peaked between May 29 and May 30, with no detectable evidence of protandry, here defined as time difference in the observed population peak (Fig. 3C). We did not observe significant deterioration in wing condition over time for either sex (four factor levels,  $Estimate_{\text{male}} = -0.01$ ,  $Estimate_{\text{female}} = -0.04$ , adjusted- $R^2 < 0.01$  for both regressions,  $P_{\text{male}} = 0.25$ ,  $P_{\text{female}} = 0.10$ ), nor could we find significant wing deterioration when considering only recaptured individuals ( $Estimate = -0.03$ , adjusted- $R^2 = 0.018$ ,  $F_{145} = 3.711$ ,  $P = 0.06$ ), suggesting no strong effects of handling during the mark-recapture procedure (Fig. 3D). In 2017 and 2018, 73% of the 52 captured females were mated (sphragis present), but logistic regression showed no significant effect of capture date ( $Estimate = 0.04$ ,  $P = 0.65$ ) or wing condition ( $Estimate = -17.29$ ,  $P = 0.99$ ) in predicting mating status (relevant data were not recorded in 2016).

#### Field MARK estimates

POPAN estimation of individual apparent survival rates fluctuated from unrealistically low in 2016 to 100% in 2018 (Table 2). These fluctuations indicated that our POPAN analysis failed to estimate the parameters and the estimated populations each year are in error (Cooch & White, 2006).

#### Model parameterization

Parameter settings for our agent-based models were based on field estimates and inferences made from available literature (Table S1). See Supplementary Material for justification of each parameter setting. The accuracy of MRR estimates are sensitive to the particular parameters (e.g., flight speed, habitat attraction) we assigned in our model to represent the movement of *B. thaidina* (Fig. S3).

#### Analysis on simulated data

Under all six landscape settings, population estimates of single-location, once-per-week MRR simulation increased significantly with population size (Table S3). Linear regression of population estimates under different population sizes in both closed boundary landscapes (Fig. 2A) and open landscapes with empirically delineated hostplant and habitat ranges (Fig. 2B) have  $R^2$  above 0.5. The accuracy of these estimates (here measured as the proportion of deviation from observed parameters) only increased significantly with increas-

ing population sizes when the boundary was either completely closed (Fig. 2A), or when the boundary was open but the habitat was empirically delineated (Fig. 2B,C). Under single-spot, low-effort sampling, most estimates were below 30% of the simulated population parameter.

Estimation accuracy and variance under different parameter settings are visualized in Figures 4–6, and summarized in Table 1. The model that best explains the absolute deviation from observed parameters for all 2772 simulations had to account for the evaluation model (POPAN or SECR), estimand type (population, lifespan, or protandry), landscape type (with or without habitat delineation), population (30–30 000), sampling effort (once or seven times per week), and number of spots (adjusted- $R^2 = 0.19$ ,  $F_{275} = 11.93$ ,  $P < 0.01$ ), with the first five variables having significant estimates. The model that best explained the deviation (both negative and positive) from the tested parameters has a better fit than that explaining absolute deviation (adjusted- $R^2 = 0.52$ ,  $F_{275} = 50.77$ ,  $P < 0.01$ ). Among the variables with significant estimates, the relative importance of each is as follows: estimand type > evaluation model > landscape type > population > effort. The model that best explains variances in estimates has to account for evaluation model, estimand type, and population (adjusted- $R^2 = 0.096$ ,  $F_{277} = 8.481$ ,  $P < 0.01$ ). Among estimates with temporal estimands (protandry and lifespan), the best model suggests a significant interaction between population and sampling effort ( $P = 0.01$ ).

#### Estimation adjustment

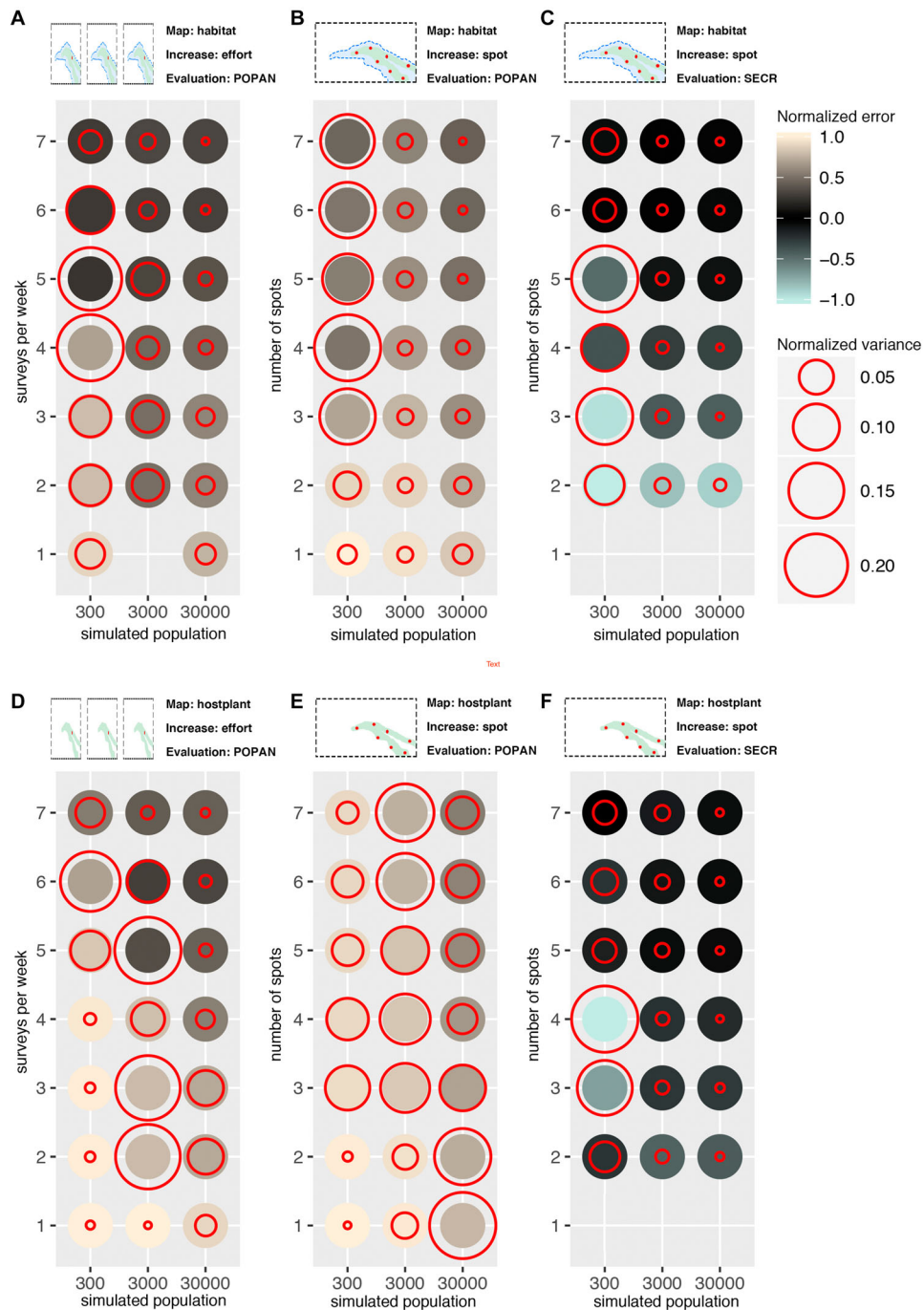
We simulated the mark-recapture schemes from 2016 to 2018 (Table S2) with population sizes from 10 to 10 800. POPAN estimates for in silico MRR data showed a significant linear increase with the population size parameter (adjusted- $R^2 > 0.88$  in all three regressions). In each of the linear regressions, the adjusted populations required to achieve the original estimates from 2016 to 2018 are higher than their prior estimates (Table 2, right side; Fig. S2).

## Discussion

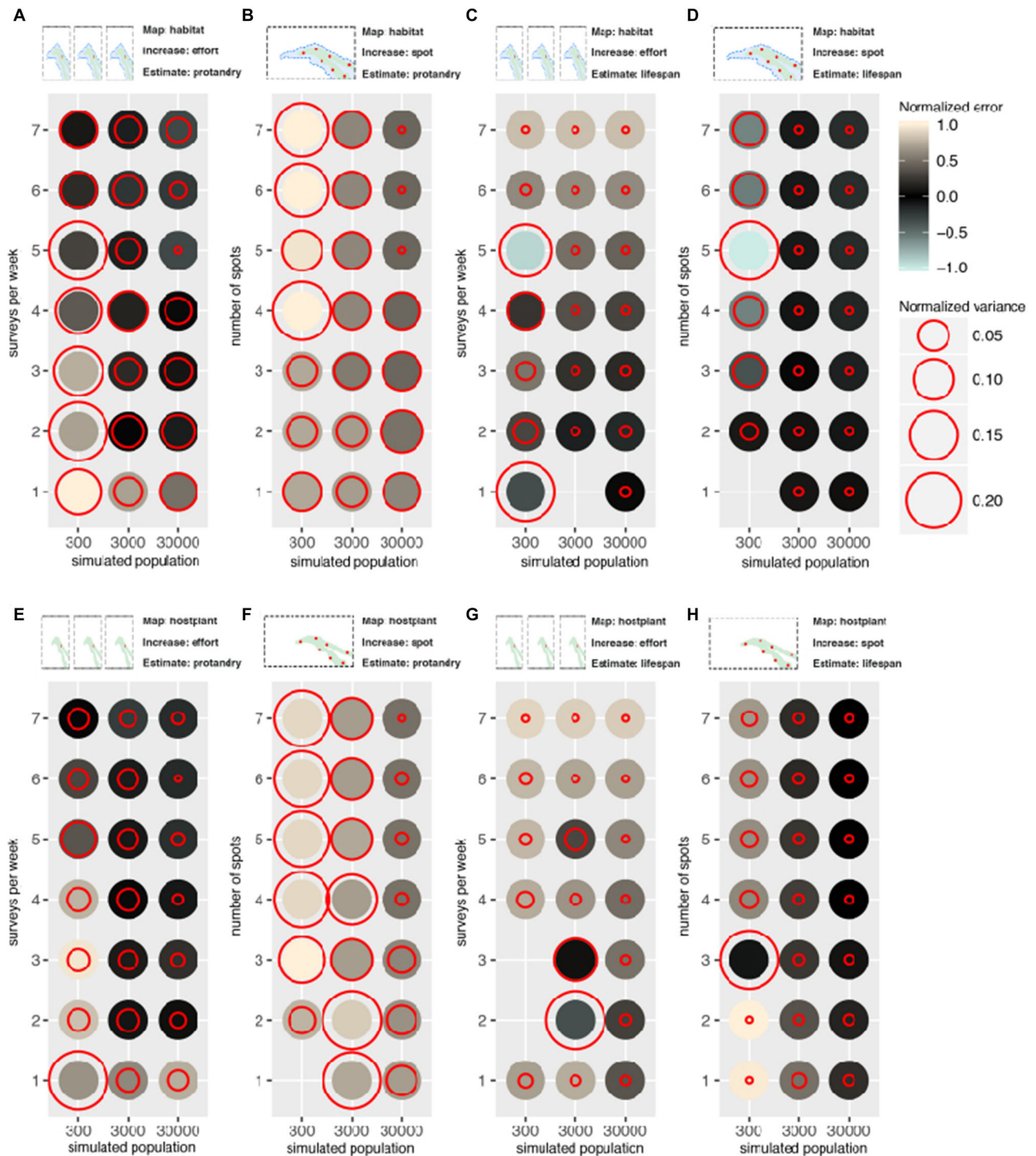
#### Difficulty of *B. thaidina* MRR

Despite our best effort (a total of 141 person-days, 352 marked samples), we did not obtain sufficient data to fit the POPAN model of MRR and failed to provide a robust population estimate. Our low recapture rate (6.76%)

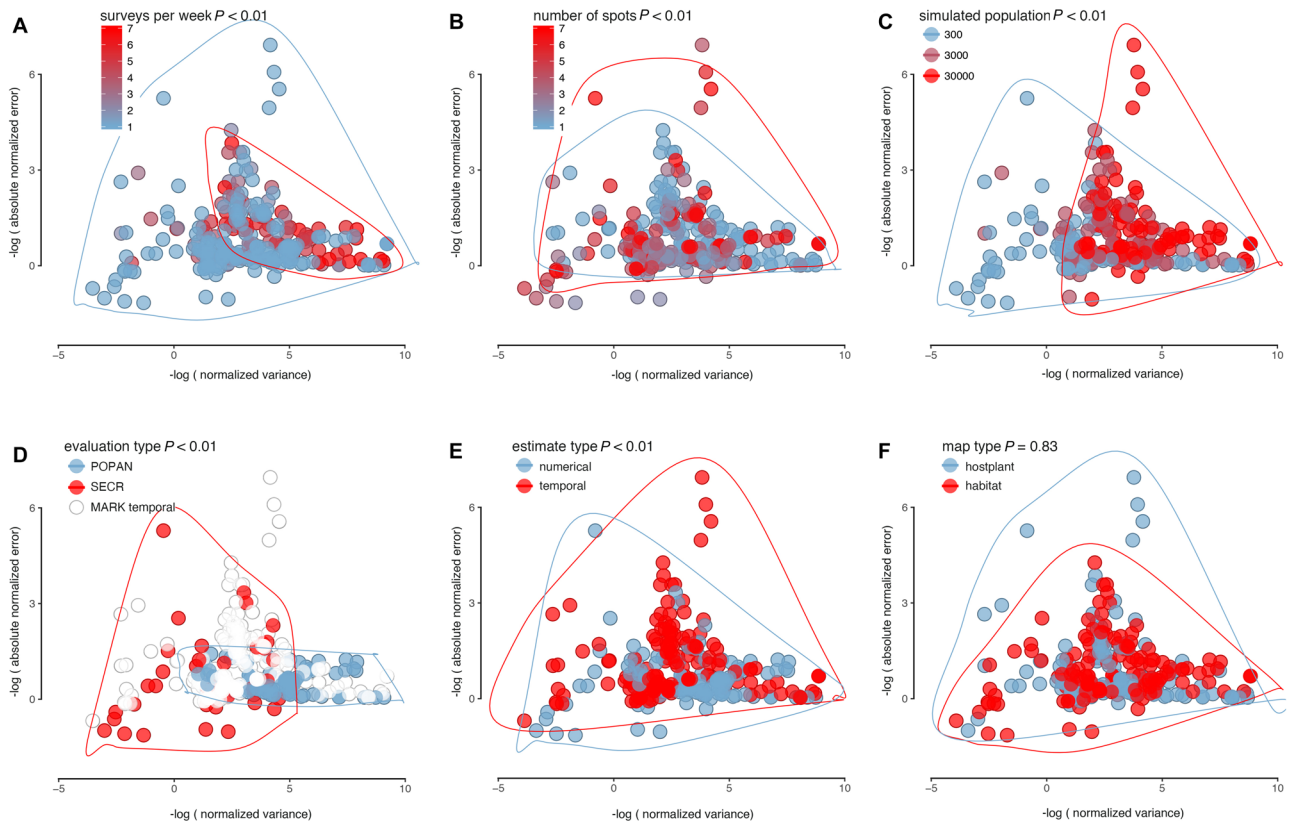




**Fig. 4** Accuracy and variance of MRR-based population estimates, comparing an “increased sampling spot” strategy with an “increased sampling frequency” strategy. (A) Estimation results after increasing sampling frequency up to 7 times per week. (B) Estimation results after increasing to as many as seven sampling spots within the hostplant range. (C) Same sampling scheme with (B) but data were analyzed with a spatially explicit model. (D)–(F) Same experimental design as (A)–(C), but simulation carried out in a landscape with no habitat delineation.



**Fig. 5** Accuracy and variance of MRR-based temporal estimates comparing the “increased sampling spot” and the “increased sampling frequency” strategies. (A)–(B) Estimating protandry with an “increased sampling frequency” strategy and an “increased sampling spot” strategy. (C)–(D) Estimating individual lifespan with an “increased sampling frequency” strategy and an “increased sampling spot” strategy. (E)–(H) Same experimental design and estimands as (A)–(D), but simulation carried out in a landscape with no habitat delineation.



**Fig. 6** Accuracy and variance of in silico MRR-based estimates under 278 model parameter-downstream analysis parameter combinations, each with 10 simulations. Each dot represents the average of 10 simulations using the same set of parameters. In each figure, the same simulation results were classified by their (A) sampling effort, (B) number of observers, (C) simulated population size, (D) type of evaluation, (E) type of estimand, and (F) map type. Dots in the bottom left of each figure represent high variance, low accuracy estimates, while simulation results that are low variance, high accuracy should in theory be in the top right of each figure (but none exist).

reflected the intrinsic difficulty in sampling *B. thaidina* in the field. We were limited to staying at a single patch of cleared forest where capture was possible—a small part of the vast available habitat of this population of *B. thaidina*. Our logistics did not allow us to deploy multiple teams across the mountainous terrain to perform multipoint MRR. While the natural history of *Bhutanitis* is relatively well-known, population estimates of rare insect species with less well-defined habitats requires much higher sampling effort. In our case we turned to agent-based simulations for insights.

#### Agent-based models and natural history

Previously, agent-based models have been used to evaluate count-based population estimates for conservation (Petrovskii *et al.*, 2012; Ahmed & Petrovskii, 2019). In

this study, we simulated MRR results based on as many as 30 000 realistically parameterized individuals, interacting at 15 s timesteps for a duration of 2 months. There are two advantages of using agent-based simulations to evaluate well-established theoretical models: (1) simulations starting with estimands as set parameters allow for quantitative evaluation of estimation errors (and in the case of repeated simulations, a quantitative evaluation of variance). This is especially useful when dealing with insect population estimates, where the true population size in the field is logistically difficult to know. (2) By customizing and adjusting a wide range of parameters in the agent-based models, it is possible to analyze which variables are important sources of variation in MRR-based estimates. While several key factors, from sampling design to analytical models have been associated with variation in MRR-based estimates (Banks & Brown, 1962; Schneider, 2003; Schtickzelle *et al.*, 2003; Evans, 2004;

**Table 1** Multivariate linear regression of the accuracy and variance of estimates under 278 agent-based model parameters-downstream analysis parameter combinations.

	Absolute accuracy				Normalized variance			
	Estimate $\pm$ SE	<i>t</i> Value	Pr(>  <i>t</i>  )	Relative importance	Estimate $\pm$ SE	<i>t</i> Value	Pr(>  <i>t</i>  )	Relative importance
Evaluation model	-0.281 $\pm$ 0.087	-3.253	<0.01	0.052	1.907 $\pm$ 0.596	3.200	<0.01	0.033
Estimand type	-0.143 $\pm$ 0.056	-2.555	<0.01	0.039	-0.719 $\pm$ 0.385	-1.868	0.063	0.006
Landscape type	0.011 $\pm$ 0.049	0.215	0.830	0.000	0.330 $\pm$ 0.340	0.970	0.333	0.003
Population	7.706e-06 $\pm$ 1.833e-06	4.204	<0.01	0.050	-4.124e-05 $\pm$ 1.262e-05	-3.267	<0.01	0.035
Effort	0.063 $\pm$ 0.015	4.243	<0.01	0.048	-0.110 $\pm$ 0.103	-1.074	0.284	0.008
Spot	0.047 $\pm$ 0.014	3.280	<0.01	0.014	0.006 $\pm$ 0.098	0.056	0.955	0.004

**Table 2** Program MARK estimation of *B. thaidina* population in Yanzigou valley and adjusted estimation based on agent-based model simulations. On the left, fluctuations in Phi (survival rates) across the years between extreme high and low values indicate that our POPAN analysis failed to estimate the parameters and the estimated populations each year are in error.

Year	MARK estimation			Estimate adjustment			
	Phi $\pm$ SE	<i>P</i> $\pm$ SE	<i>N</i> $\pm$ SE	Linear fit coefficient	Linear fit coefficient SD	Adjusted <i>R</i> <sup>2</sup>	Adjusted <i>N</i>
2016	0.22 $\pm$ 0.140	0.31 $\pm$ 0.25	187.73 $\pm$ 139.13	0.11	0.005	0.88	1825
2017	1.00 $\pm$ 0.000	0.04 $\pm$ 0.01	845.38 $\pm$ 206.40	0.31	0.007	0.97	2784
2018	0.95 $\pm$ 0.015	0.01 $\pm$ 0.00	1082.00 $\pm$ 351.55	0.47	0.007	0.99	2329

Rosenberger & Dunham, 2005; Haddad *et al.*, 2008; Pellet *et al.*, 2012), it is either infeasible or impossible to repeat experimental conditions in the field for many systems (e.g., controlling for mark-recapture regime and individual behavior but varying the landscape pattern), but easy to implement in agent-based models.

This is not to say that agent-based models eliminate the need for fieldwork components of MRR-based estimates. Rather, our study provides incentives for field conservationists to collect information about life-history parameters such as emergence peak, hostplant distribution, and behavioral data—in our simulations, accurate parameters provide accurate estimates. In fact, one of the most inaccurate estimates in our simulations came from a scenario in which the *in silico* surveyor had chosen to sample outside the habitat (Fig. 2F)—a costly mistake to make for a conservation biologist without sufficient knowledge of a species' habitat. Although we have selected input parameters from both our own empirical study and from published life-history analysis (Table S1), we found that even for a recently IUCN-assessed species such as *B. thaidina*, data for many important param-

eters are lacking. One parameter that is particularly difficult to estimate is the likelihood that individuals will disperse out of their preferred habitat (as measured and modeled by Schultz & Crone, 2001); this is an important parameter to adjust for each population if we were to assess the accuracy of existing MRR estimations, since increasing its value in simulations results in better approximation of true population parameters (Fig. S3C). Similarly, in our study, simulated MRR on individuals with lower flight speeds approximates true population parameters (Fig. S3A), probably because these individuals are less likely to disperse out of the grid. It is important, when applying agent-based models to other organisms, to collect accurate natural history information relevant to the species or population concerned.

Other parameters that could be empirically measured but have not yet been incorporated into our model include the activity state of each individual (e.g., active or inactive state; see Crall *et al.*, 2019) and the decay in wing condition of each individual. Our model also did not take into account any climatic information (see Hu *et al.*, 2019). We included equal sex ratio and equal lifespan for

both sexes as biological parameters in our models (following Zhang *et al.*, 2019's data on primary sex ratio of a captive breed of a closely related *Bhutanitis* species). The male biased effective sex ratio measured in field observations suggests the effects of both protandry and differential survivorship of butterflies (Calabrese, 2012; Sielezniew *et al.*, 2020). We did not differentially parameterize male and female behavior in our models as this could result in sex-based differences in detectability (Pickett *et al.*, 2012; Ehl *et al.*, 2019).

### *Systematic MRR underestimations*

Estimates for simulated MRR data revealed systematic underestimation of the true population parameters, regardless of landscape types, when sampling effort is “insufficient” (Fig. 2). Our study provided a quantitative description of underestimation: when we only sample once per week at a single spot for the entire span of a population's emergence—this doesn't immediately strike us as an “insufficient” sampling scheme—program MARK's population estimate is less than 5% of the true parameter. For a few landscape types, the accuracy of estimation increases with population size (as higher *in silico* population sizes result in larger recapture datasets), but even under these scenarios, the estimate under such a sampling effort was no more than 50% of the true parameter (10 000 individuals). Under more realistic landscape types (open boundary landscapes with habitat and host-plant delineation), our best estimate when sampling once per week is less than 30% of the true population.

### *MRR at different landscapes*

Our simulations show that landscape types and boundary types both have significant effects on the accuracy of MRR-based estimates (Fig. 2; Table S3). Butterflies disperse differently in different landscapes (Jain *et al.*, 2020); landscape heterogeneity and habitat size are known to cause bias in insect population estimates in the field (Alexander *et al.*, 2005; Turlure *et al.*, 2010), but MRR models do not explicitly account for landscape type and boundary effects. Although Kendall (1999) has shown that MRR models should provide unbiased population estimation even in open habitats, this idea has not been tested in simulations. Our simulations show that estimates under a closed landscape more accurately approximate the true population parameters than estimates under landscapes with open boundaries. Within all landscape types with open boundaries, the most realistically parameterized landscape (hostplant and habitat delineation,

with sampling spot inside the habitat; Fig. 2B) outperforms landscape types without habitat specifications or when the simulated sampling spot is outside the habitat range. This model achieved nearly 20% of the true parameter when the population was sufficiently large, despite “insufficient” sampling. This suggests that although MRR models are inherently designed with closed boundaries and lack proper habitat consideration, they may operate satisfactorily in heterogeneous habitats. The difference in estimation accuracy using data obtained from a realistically parameterized landscape versus a homogenized landscape is persistent despite the increase in sampling effort (Figs. 4 and 5, comparing plots on the top with plots on the bottom), but not statistically significant.

### *Effect of research design on population estimates*

When simulations were performed on a fixed landscape type, many other factors in our research design could influence MRR estimates. A common question regarding research design can be formulated as follows: given resources sufficient to support, for example, seven person-days per week, should we allocate one person to sample seven days per week, or conduct once-per week sampling at seven spots in order to obtain the most accurate population estimate? Our results show that subjective decisions in research designs (such as sampling frequency, number of sampling spots, the choice of whether to use spatially explicit models, whether the estimands are population sizes or individual lifespans) account for nearly half of the variation in a total of 2772 simulations (adjusted- $R^2 = 0.46$ ,  $F_{277} = 60.09$ ,  $P < 0.01$ ).

As a rule of thumb, we obtain more accurate, less variable estimates when estimating more abundant populations (or rather, when we obtain more recaptures), and our estimation improves as we increase both our sampling frequency and the number of sampling spots. Specifically, (1) given limited resources for population estimates, it is better to conduct more frequent surveys at a single point than it is to sample multiple points less frequently (although we note that this result is under the parameters fit for our specific *B. thaidina* population, and should be reevaluated in terms of habitat requirement and individual movement parameters of the specific populations or species involved); (2) when we designed our simulated experiments to obtain multipoint recaptures, spatially explicit mark-recapture models indeed provided more accurate and less variable population estimates than MRR models without explicit spatial models; and (3) when estimands are temporal (i.e., individual lifespan or length of protandry), we observed scenarios when

more accurate estimates are achieved through low-frequency sampling. Increasing sampling frequency will result in over-estimation of species lifespan. Note that although we simulated individual life span with normal distribution and fitted them with a “wrong” model in POPAN (constant survival model), accurate temporal estimation is still possible with increased sampling effort.

Even within the limited parameter-space, we explored in this study (no more than seven sampling locations, sampling at most seven days per week), most of this parameter-space yielded high error, high variance estimates (denoted by the light color palette in Figs. 4 and 5) that conservation biologists should ideally avoid in designing field sampling schemes. Even though we used parameters specific to *B. thaidina*, we expect these “pitfalls” to remain in the parameter-space, at least for other insects or invertebrates operating on similar movement scales. This suggests a potential explanation as to why MRR-based insect population estimates in field-based comparative studies tend to report underestimates of population sizes compared with other estimation methods such as absolute counts.

#### *Adjusting field estimates*

We learnt from agent-based simulations that when sampling efforts are insufficient, as in our case in the field, MRR estimates are underestimates. Moreover, since the increase in population estimates with simulated population size is highly predictable regardless of sampling design (adjusted- $R^2 > 0.88$  in all linear regressions), we could reasonably predict the population size that generates each year’s corresponding field estimates if we run simulations of our sampling design across a series of increasing population parameters (Table 2, Fig. S2). These obtained “adjusted estimates” are well above the original MRR estimates.

#### **Conclusions**

Our study combined agent-based modeling, field mark–release–recapture (MRR) methods and natural history surveys to show the large extent to which field-designs and downstream model choices can influence the accuracy and variance of MRR-based population and lifespan estimates. We emphasize the importance of foundational natural history work in improving MRR-based estimates and note several “pitfalls” in sampling design for conservationists using MRR-based estimates of population size. We suggest that agent-based models could also be applied to adjust for existing population estimates.

#### **Acknowledgments**

ZW was supported by a graduate fellowship from Harvard University Department of Organismic and Evolutionary Biology, as well as a Joan Mosenthal DeWind Award in Lepidoptera conservation from the Xerces Society (2017) for studying the spatial ecology of butterflies in Yanzigou valley. ZW and AJ were supported by a Rufford Small Grant (2016) to work on the conservation of *B. thaidina*. Aspects of the modeling approach used in this research were supported by a grant from the Harvard Global Institute (HGI) to YL, ZW and NEP. We thank Huailiang Tang, Zulian Zhou, Qingqing Liu, Yifan Liu, and Yuan Huang for field assistance, and Kadeem Gilbert, Sofia Prado-Irwin, and Jimmy Zeng for commenting on an earlier version of the manuscript.

#### **Author contributions**

ZW and AJ conceived and designed the study, and carried out the field work. YL designed the agent-based models and implemented the simulations. ZW analyzed the data and drafted the first version of the manuscript. The final paper was written by all authors.

#### **References**

- Ahmed, D.A. and Petrovskii, S.V. (2019) Analysing the impact of trap shape and movement behaviour of ground-dwelling arthropods on trap efficiency. *Methods in Ecology and Evolution*, 10, 1246–1264.
- Alexander, C.J., Holland, J.M., Winder, L., Woolley, C. and Perry, J.N. (2005) Performance of sampling strategies in the presence of known spatial patterns. *Annals of Applied Biology*, 146, 361–370.
- Banks, C.J. and Brown, E.S. (1962) A comparison of methods of estimating population density of adult sunn pest, *Eurygaster integriceps* Put (Hemiptera, Scutelleridae) in wheat fields. *Entomologia Experimentalis et Applicata*, 5, 255–260.
- Borchers, D.L. and Efford, M.G. (2008) Spatially explicit maximum likelihood methods for capture-recapture studies. *Biometrics*, 64, 377–385.
- Calabrese, J. (2012) How emergence and death assumptions affect count-based estimates of butterfly abundance and lifespan. *Population Ecology*, 54(3), 431–442.
- Cooch, E.G. and White, G.C. (2006) *Program MARK: A Gentle Introduction* (5th edition). ePublication: <http://www.phidot.org/software/mark/docs/book/>
- Craig, C.C. (1953) On the utilization of marked specimens in estimating populations of flying insects. *Biometrika*, 40(1–2), 170–176.

- Crall, J.D., de Bivort, B.L., Dey, B. and Ford Versypt, A.N. (2019) Social buffering of pesticides in bumblebees: agent-based modeling of the effects of colony size and neonicotinoid exposure on behavior within nests. *Frontiers in Ecology and Evolution*, 7, 51.
- DeAngelis, D.L. (1992) *Individual-Based Models and Approaches in Ecology: Populations, Communities and Ecosystems* (1st edn). Boca Raton, Florida: Chapman and Hall/CRC.
- DeAngelis, D. L. and Grimm, V. (2014) Individual-based models in ecology after four decades. *FI000prime Reports*, 6, 39.
- DeAngelis, D.L. and Mooij, W.M. (2005) Individual-based modeling of ecological and evolutionary processes. *Annual Review of Ecology, Evolution, and Systematics*, 36, 147–168.
- Didham, R.K., Basset, Y., Collins, C.M., Leather, S.R., Littlewood, N.A., Menz, M.H.M., *et al.* (2020) Interpreting insect declines: seven challenges and a way forward. *Insect Conservation and Diversity*, 13, 103–114.
- Donkers, P., Patil, J.G., Wisniewski, C. and Diggle, J.E. (2012) Validation of mark-recapture population estimates for invasive common carp, *Cyprinus carpio*, in Lake Crescent, Tasmania: population estimates for invasive common carp, *Cyprinus carpio*. *Journal of Applied Ichthyology*, 28, 7–14.
- Dupont, G., Royle, J.A., Nawaz, M.A. and Sutherland, C. (2021) Optimal sampling design for spatial capture-recapture. *Ecology*, 102, e03262.
- Eckhoff, P.A., Wenger, E.A., Godfray, H.C.J. and Burt, A. (2016) Impact of mosquito gene drive on malaria elimination in a computational model with explicit spatial and temporal dynamics. *Proceedings of the National Academy of Sciences USA*, 114, E255–E264.
- Efford, M. (2004) Density estimation in live-trapping studies. *Oikos*, 106, 598–610.
- Efford, M. (2019) secr: Spatially explicit capture-recapture models. R package version 4.0.2.
- Efford, M.G. and Fewster, R.M. (2013) Estimating population size by spatially explicit capture-recapture. *Oikos*, 122, 918–928.
- Efford, M.G. and Boulanger, J. (2019) Fast evaluation of study designs for spatially explicit capture-recapture. *Methods in Ecology and Evolution*, 10, 1529–1535.
- Ehl, S., Holzhauer, S.I., Ryrholm, N. and Schmitt, T. (2019) Phenology, mobility and behaviour of the arcto-alpine species *Boloria napaea* in its arctic habitat. *Scientific Reports*, 9, 1–9.
- Evans, T. (2004) Comparing mark-recapture and constant removal protocols for estimating forager population size of the subterranean termite *Coptotermes lacteus* (Isoptera: Rhinotermitidae). *Bulletin of Entomological Research*, 94, 1–9.
- Ferster, B. and Vulinec, K. (2010) Population size and conservation of the last eastern remnants of the regal fritillary, *Speyeria idalia* (Drury) [Lepidoptera, Nymphalidae]: implications for temperate grassland restoration. *Journal of Insect Conservation*, 14, 31–42.
- Gao, K., Li, X.S., Guo, Z.Y. and Zhang, Y.L. (2014) The bioeconomics, habitat requirements and population threats of the butterfly *Bhutanitis thaidina* in Taibai Mountain. *Journal of Insect Conservation*, 18, 29–38.
- Green, A.M., Chynoweth, M.W. and Şekercioğlu, Ç.H. (2020) Spatially explicit capture-recapture through camera trapping: a review of benchmark analyses for wildlife density estimation. *Frontiers in Ecology and Evolution*, 8, 563477.
- Griebeler, E.M. and Seitz, A. (2002) An individual based model for the conservation of the endangered large blue butterfly, *Maculinea arion* (Lepidoptera: Lycaenidae). *Ecological Modelling*, 156, 43–60.
- Gross, K., Kalendra, E.J., Hudgens, B.R. and Haddad, N.M. (2007) Robustness and uncertainty in estimates of butterfly abundance from transect counts. *Population Ecology*, 49, 191–200.
- Haddad, N.M., Hudgens, B., Damiani, C., Gross, K., Kuefler, D. and Pollock, K. (2008) Determining optimal population monitoring for rare butterflies. *Conservation Biology*, 22, 929–940.
- Hain, E.F., Lamphere, B.A., Blum, M.J., McIntyre, P.B., Nelson, S.A. and Gilliam, J.F. (2016) Comparison of visual survey and mark-recapture population estimates of a benthic fish in Hawaii. *Transactions of the American Fisheries Society*, 145, 878–887.
- Hallmann, C.A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Schwan, H. *et al.* (2017) More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PLoS ONE*, 12(10), e0185809.
- Hamm, C.A. (2013) Estimating abundance of the federally endangered Mitchell's satyr butterfly using hierarchical distance sampling. *Insect Conservation and Diversity*, 6, 619–626.
- Hu, S.J., Xing, D.H., Gong, Z.X. and Hu, J.M. (2019) Projecting suitability and climate vulnerability of *Bhutanitis thaidina* (Blanchard) (Lepidoptera: Papilionidae) with conservation implications. *Scientific Reports*, 9, 15384.
- Hurvich, C.M. and Tsai, C.L. (1989) Regression and time series model selection in small samples. *Biometrika*, 76, 297–307.
- Igarashi, S. (2003) Life history of *Bhutanitis mansfieldi* in comparison with those of related species. *Butterflies*, 35, 20–39.
- Igarashi, S. and Harada, M. (2015) *Sequel to "The Life Histories of Asian Butterflies Vols. I & II."* Self-published, Tokyo.
- Irving, E. and Hebda, R. (1993) Concerning the origin and distribution of *Rhododendrons*. *Journal of the American Rhododendron Society*, 47, 139–162.
- Jain, A., Chan, S.K.M., Vlasanek, P. and Webb, E.L. (2020) Impacts of habitat on butterfly dispersal in tropical forests, parks and grassland patches embedded in an urban landscape. *Biotropica*, 52, 404–409.

- Jimenez, J., Chandler, R., Tobajas, J., Descalzo, E., Mateo, R. and Ferreras, P. (2019) Generalized spatial mark-resight models with incomplete identification: an application to red fox density estimates. *Ecology and Evolution*, 9, 4739–4748.
- Jolly, G.M. (1965) Explicit estimates from capture-recapture data with both death and immigration-stochastic model. *Biometrika*, 52(1–2), 225–248.
- Kendall, W.L. (1999) Robustness of closed capture-recapture methods to violations of the closure assumption. *Ecology*, 80, 2517–2525.
- Kral, K., Harmon, J., Limb, R. and Hovick, T. (2018) Improving our science: the evolution of butterfly sampling and surveying methods over time. *Journal of Insect Conservation*, 22, 1–14.
- Kristensen, T.V. and Kovach, A.I. (2018) Spatially explicit abundance estimation of a rare habitat specialist: implications for SECR study design. *Ecosphere*, 9, e02217.
- Lewis, H. M., Vinicius, L., Strods, J., Mace, R. and Migliano, A.B. (2014) High mobility explains demand sharing and enforced cooperation in egalitarian hunter-gatherers. *Nature Communications*, 5, 5789.
- Li, X. (2019) *Bhutanitis thaidina*. The IUCN Red List of Threatened Species 2019, e.T2798A122599287. <https://doi.org/10.2305/IUCN.UK.2019-3.RLTS.T2798A122599287.en>. Downloaded on 13 January 2020.
- Lindberg, M.S. (2012) A review of designs for capture-mark-recapture studies in discrete time. *Journal of Ornithology*, 152(S2), 355–370.
- Myers, J.H. (1976) Distribution and dispersal in populations capable of resource depletion: a simulation model. *Oecologia*, 23, 255–269.
- Nowicki, P., Deoniziak, K., Dziekan śka, I., Kostro-Ambroziak, A., Plazio, E., Rutkowski, R. et al. (2019) What keeps ‘living dead’ alive: demography of a small and isolated population of *Maculinea* (= *Phengaris*) *alcon*. *Journal of Insect Conservation*, 23, 201–210.
- Nowicki, P., Richter, A., Glinka, U., Holzschuh, A., Toelke, U., Henle, K. et al. (2005) Less input same output: simplified approach for population size assessment in Lepidoptera. *Population Ecology*, 47, 203–212.
- Nowicki, P., Settele, J., Henry, P.Y. and Woyciechowski, M. (2008) Butterfly monitoring methods: the ideal and the real world. *Israel Journal of Ecology & Evolution*, 54, 69–88.
- Ovaskainen, O. (2004) Habitat-specific movement parameters estimated using mark-recapture data and a diffusion model. *Ecology*, 85, 242–257.
- Ovaskainen, O., Rekola, H., Meyke, E. and Arjas, E. (2008) Bayesian methods for analyzing movements in heterogeneous landscapes from mark-recapture data. *Ecology*, 89, 542–554.
- Pellet, J., Bried, J.T., Parietti, D., Gander, A., Heer, P.O., Cherix, D. et al. (2012) Monitoring butterfly abundance: beyond pollard walks. *PLoS ONE*, 7, e41396.
- Pickett, E.J., Stockwell, M.P., Pollard, C.J., Garnham, J.I., Clulow, J. and Mahony, M.J. (2012) Estimates of sex ratio require the incorporation of unequal catchability between sexes. *Wildlife Research*, 39, 350–354.
- Petrovskii, S., Bearup, D., Ahmed, D.A. and Blackshaw, R. (2012) Estimating insect population density from trap counts. *Ecological Complexity*, 10, 69–82.
- Pollock, K.H., Nichols, J.D., Brownie, C. and Hines, J.E. (1990) Statistical inference for capture-recapture experiments. *Wildlife Monographs*, 107, 3–97.
- Ponchon, A., Choquet, R., Tornos, J., McCoy, K.D., Tveraa, T. and Boulinier, T. (2018) Survival estimates strongly depend on capture-recapture designs in a disturbed environment inducing dispersal. *Ecography*, 41, 2055–2066.
- R Core Team (2019) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rosenberger, A.E. and Dunham, J.B. (2005) Validation of abundance estimates from mark-recapture and removal techniques for rainbow trout captured by electrofishing in small streams. *North American Journal of Fisheries Management*, 25, 1395–1410.
- Ruiz, D., Rotger, A., Igual, J.M. and Tavecchia, G. (2013) Estimating lizard population density: an empirical comparison between line-transect and capture-recapture methods. *Wildlife Research*, 40, 552–560.
- Sanchez-Bayo, F. and Wyckhuys, K.A. (2019) Worldwide decline of the entomofauna: a review of its drivers. *Biological Conservation*, 232, 8–27.
- Sandercock, B.K. (2006) Estimation of demographic parameters from live-encounter data: a summary review. *Journal of Wildlife Management*, 70, 1504–1520.
- Schneider, C. (2003) The influence of spatial scale on quantifying insect dispersal: an analysis of butterfly data. *Ecological Entomology*, 28, 252–256.
- Schtickzelle, N., Baguette, M. and Boulenge, E.L. (2003) Modelling insect demography from capture-recapture data: comparison between the constrained linear models and the Jolly-Seber analytical method. *The Canadian Entomologist*, 135, 313–323.
- Schultz, C.B. and Crone, E.E. (2001) Edge-mediated dispersal behavior in a prairie butterfly. *Ecology*, 82, 1879–1892.
- Schwarz, C.J. and Arnason, A.N. (1996) A general methodology for the analysis of capture-recapture experiments in open populations. *Biometrics*, 52, 860.
- Seber, G.A.F. (1965) A note on the multiple-recapture census. *Biometrika*, 52(1–2), 249–260.
- Seber, G.A.F. (1970) Estimating time-specific survival and reporting rates for adult birds from band returns. *Biometrika*, 57, 313–318.
- Seber, G.A.F. (1982) *The Estimation of Animal Abundance and Related Parameters*. Edward Arnold, London.



- Sielezniew, M., Kostro-Ambroziak, A. and Korosi, A. (2020) Sexual differences in age-dependent survival and life span of adults in a natural butterfly population. *Scientific Reports*, 10, 10394.
- Sutherland, W.J. (ed.) (1996) *Ecological Census Techniques: A Handbook*. Cambridge University Press, Cambridge, New York.
- Turlure, C., Choutt, J., Van Dyck, H., Baguette, M. and Schtickzelle, N. (2010) Functional habitat area as a reliable proxy for population size: case study using two butterfly species of conservation concern. *Journal of Insect Conservation*, 14, 379–388.
- Turlure, C., Pe'er, G., Baguette, M. and Schtickzelle, N. (2018) A simplified mark-release-recapture protocol to improve the cost effectiveness of repeated population size quantification. *Methods in Ecology and Evolution*, 9, 645–656.
- van Strien, A.J., Van Swaay, C.A., van Strien-van Liempt, W.T., Poot, M.J. and WallisDeVries, M.F. (2019) Over a century of data reveal more than 80% decline in butterflies in the Netherlands. *Biological Conservation*, 234, 116–122.
- Wang, Z., Huang, Y. and Pierce, N.E. (2019) Radio telemetry helps record the dispersal patterns of birdwing butterflies in mountainous habitats: golden birdwing (*Troides aeacus*) as an example. *Journal of Insect Conservation*, 23, 729–738.
- Welti, E.A.R., Joern, A., Ellison, A.M., Lightfoot, D.C., Record, S., Rodenhouse, N. *et al.* (2021) Studies of insect temporal trends must account for the complex sampling histories inherent to many long-term monitoring efforts. *Nature Ecology and Evolution*, 5, 589–591.
- Yi, C.H., He, F.X., He, Q.J. and Wang, L. (2011) The biological characteristics of *Bhutanitis yulongensis* Chou. *Journal of Applied Entomology*, 48, 1505–1508. (in Chinese)
- Zhang, Z., Jiang, Z. and Wang, Z. (2019) Habitat, life history and oviposition choice of the endangered Bhutan Glory swallowtail *Bhutanitis lidderdalii* (Lepidoptera: Papilionidae) in Yunnan, China. *Journal of Insect Conservation*, 23(5–6), 921–931.

Manuscript received February 19, 2021

Final version received June 15, 2021

Accepted June 16, 2021

## Supporting Information

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

**Fig. S1.** Schematic for adjusting MRR-based estimates with agent-based simulations.

**Fig. S2.** Adjusting MRR-based estimation of *B. thaidina* population.

**Fig. S3.** Sensitivity of POPAN population estimates to changes in key life-history parameters in agent-based simulations.

**Table S1.** Parameters for agent-based models used in this study.

**Table S2.** MRR sampling scheme for *B. thaidina* sampling from 2016 to 2018 and results.

**Table S3.** Estimation results based on simulated MRR data under six different landscape types.