

1 Accounting for imperfect detection in species with sessile life cycle stages: a case study of  
2 bumble bee colonies

3 David T. Iles<sup>1,2</sup>, Genevieve Pugesek<sup>1\*</sup>, Natalie Z. Kerr<sup>1</sup>, Nicholas N. Dorian<sup>1</sup>, and Elizabeth E.  
4 Crone<sup>1</sup>

5 <sup>1</sup>Department of Biology, Tufts University, Medford, MA, USA

6 <sup>2</sup>Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA, USA

7 \*Corresponding author:

8 Phone: (208) 610-5290; Email: [genevieve.pugesek@tufts.edu](mailto:genevieve.pugesek@tufts.edu)

9 G.P. ORCID: 0000-0001-9517-1328

10

11 **Acknowledgements**

12 We thank the Trustees of Reservations and Appleton Farms for providing access to their  
13 properties on which this study was conducted. We thank Russ Hopping for invaluable assistance  
14 with study design. We also thank Annika Greenleaf, Max McCarthy, Moshe Steyn, Erin  
15 Wampole, and our sniffer-dogs-in-training, Indy and Molly, for assistance with field work. This  
16 work was supported by the US National Science Foundation (DEB1354022) and the US  
17 Strategic Environmental Research and Development Program (SERDP, RC-2119).

18 **Abstract**

- 19 1. For bumble bees, colonies (not individual workers) are the functional unit of the population.  
20 Estimates of colony density are thus critical for understanding population distribution and  
21 trends of this important pollinator group. Yet, surveys of bumble bee colonies and other taxa  
22 with sessile life cycle states rarely account for imperfect detection.
- 23 2. Here we demonstrate the use of mark-recapture methods to estimate the density of bumble  
24 bee colonies across the landscape using standardized survey protocols.
- 25 3. We found that the probability of detecting colonies in standardized surveys varied  
26 considerably across space, through time, and among colonies.
- 27 4. Using simulations, we also show that imperfect detection can obscure true variation in  
28 density among plots, or generate spurious variation in counts even when all plots have the  
29 same density. In both cases, we show that mark-recapture can be used to generate unbiased  
30 estimates of density, with relatively low search effort compared to conventional survey  
31 methods for bumble bee colonies.
- 32 5. Our study illustrates the advantages of mark-recapture for optimizing survey protocols for  
33 species with cryptic and sessile life cycle stages, which will be a valuable tool in ongoing  
34 studies of pollinator nesting ecology.

35

36 **Key-words** *Bombus*; closed population model; mark-recapture; monitoring; nesting habitat;  
37 occupancy

## 38 **Introduction**

39           Measuring the size of natural populations is a primary goal of conservation monitoring  
40 and is often a prerequisite for management action or wildlife policy decisions (Williams et al.  
41 2002). Yet, estimating abundance and the factors that influence it can be challenging when  
42 organisms are difficult to detect. Accordingly, study designs and analyses that account for  
43 imperfect detection have a rich tradition in wildlife research (White and Burnham 1999;  
44 Williams et al. 2002; Kéry and Schaub 2012). Failure to account for imperfect detection of  
45 cryptic organisms can introduce both uncertainty and bias into population estimates, potentially  
46 leading to erroneous conclusions about population status and habitat requirements (Gu and  
47 Swihart 2004; Kéry and Schmidt 2008).

48           In contrast to studies of mobile vertebrates, imperfect detection is rarely accounted for in  
49 population estimates for both invertebrates and taxa that are sessile or have a sessile life-cycle  
50 stage, including plants and ground-nesting animals (Kellner and Swihart 2014; Berberich et al.  
51 2016). Kellner and Swihart (2014) found that only 9.0% of invertebrate and 1.4% of plant  
52 population studies accounted for imperfect detection, compared to the mean of 23% across all  
53 taxa. This discrepancy potentially stems from the incorrect assumption that detection probability  
54 is inherently high in less mobile organisms. Yet, this assumption may be false if sessile  
55 organisms are rare, inconspicuous, or logistically difficult to survey. For example, in a long-term  
56 study of a 4.5 ha plot in Kansas, Slade et al. (2003) estimated that fewer than 4% of existing  
57 Mead's milkweed plants were discovered in years without spring fires, and only 18% were  
58 discovered in years with spring burning. In a study of Baltimore Checkerspot butterfly  
59 demography, Brown et al. (2017) estimated the density of caterpillar aggregations on the  
60 landscape using mark-resight surveys. The detection probability ranged from 0.32 to 0.56 on

61 native *Chelone glabra* host plants, and from 0.17 to 0.25 on exotic *Plantago lanceolata* plants  
62 (L.M. Browne and E.E. Crone, pers. comm, from analyses in Brown et al. 2017). Surprisingly,  
63 Berberich et al. (2016) found that observers conducting one hour surveys for red wood ant nests  
64 in 60m x 60m plots failed to detect up to 40% of ant nests larger than 50cm. Thus, even for  
65 studies of sessile organisms, detection can be highly imperfect and systematically biased.

66 Bumble bees are an important group of pollinators with a sessile life-cycle stage. After a  
67 queen bumble bee has established a colony and produced her first cohort or two of workers, she  
68 remains in the nest until the colony expires (Goulson 2010). Since colonies are the functional  
69 unit of the population for social insects such as bumble bees, studies of nesting habitat are  
70 particularly valuable for conservation planning. Yet, colonies are difficult to find, so relatively  
71 few studies have investigated the correlates of nest density compared to the numerous studies of  
72 habitat preference by foraging workers (but see examples in Harder 1986; Osborne et al. 2008;  
73 Waters et al. 2010; Lye et al. 2012; O'Connor et al. 2012, 2017). This disparity is problematic  
74 because workers are highly mobile, e.g., workers often forage up to 1 km from their colony  
75 (Greenleaf et al. 2007) and may be attracted to areas rich in floral resources. This decoupling of  
76 foraging and nesting sites potentially obscures the landscape drivers of population performance  
77 (Heard et al. 2007). Indeed, Herrmann et al. (2007) reported that bumble bee colony densities  
78 were uncorrelated with local worker densities across an agricultural landscape in Germany.

79 A variety of methods have been used to estimate the density of bumble bee nests across  
80 landscapes. These methods have included “free searches” in which observers haphazardly search  
81 particular habitat types for nests (Harder 1986; O'Connor et al. 2012; Rao and Skyrms 2013;  
82 O'Connor et al. 2017), canine-assisted searches (Waters et al. 2010; O'Connor et al. 2012), and  
83 concentrated stationary observation of small plots by individual researchers or distributed citizen

84 science networks (Osborne et al. 2008; Lye et al. 2012). As an alternative to ground-based  
85 searches, molecular (i.e., genetic) analysis of foraging workers has also been used to infer nest  
86 density at large spatial scales (Darvill et al. 2004; Goulson et al. 2010). Other studies have  
87 inferred the relative density of nests across space based on the prevalence of spring prospecting  
88 behavior by newly emerged queens (Svensson et al. 2000; Kells and Goulson 2003; O'Connor et  
89 al. 2017). Combined, these studies have reported considerable variation in nest densities (range:  
90 0.1 to 50.1 nests·ha<sup>-1</sup>; Table 2 in Appendix 1), potentially owing to ecologically relevant  
91 differences across species, habitat types, and landscape configurations. However, O'Connor et al.  
92 (2012) also reported a 20-fold difference in the number of nests detected between fixed and free  
93 searches, indicating an extreme degree of variation in nest detection among survey strategies.  
94 Detection probability also often varies through time, across space, and between individuals  
95 (Anderson 2001; Kéry and Schaub 2012). Therefore, unaccounted differences in detection  
96 probability within and among studies could contribute to observed variation in nest density,  
97 limiting the ability to generalize across studies and resolve the true environmental correlates of  
98 bumble bee population abundance.

99         Here, our objective was to generate unbiased estimates of bumble bee nest density using  
100 mark-recapture methods while simultaneously examining the factors that influence imperfect  
101 detection of nests. We first discuss the general application of closed-population modelling  
102 approaches to estimate the abundance of cryptic sessile organisms. We then review a classic  
103 catalogue of model structures that can be used to correct for systematic bias in detection  
104 probability, along with their specific relevance to our study of bumble bee nest density. We fit  
105 these models to empirical data and examine the consequences of imperfect nest detection in a  
106 field setting. Finally, we conduct a simulation to illustrate the (spurious) variability induced into

107 count data by imperfect detection, and to demonstrate that this approach can be usefully applied  
108 in cases where a large number of sites are only visited twice. Our study emphasizes the utility of  
109 mark-recapture approaches for examining the ecological correlates of nest density for social  
110 insects, and outlines a strategy for surveying bumble bee colonies with imperfect detection.

### 111 *Mark-recapture for estimating abundance with imperfect detection*

112 Closed population models can be used to estimate abundance while accounting for  
113 imperfect detection when the processes of birth, mortality, and movement do not alter the  
114 number of individuals in plots over the course of a study (i.e., when the population is  
115 “demographically closed” across the sampling period). For sessile organisms, this assumption is  
116 satisfied when birth and death processes are unlikely to occur across the sampling period.  
117 Individuals can also be censored if they are known to have died during the sampling period.

118 Otis et al. (1978) outlined a classic catalogue of model structures that can be used to  
119 examine drivers of variation in detection probability for closed populations. We adopt this  
120 framework for our analysis of bumble bee nests to demonstrate how these models can be applied  
121 to studies of sessile organisms and to link our specific study system to a well-defined body of  
122 mark-recapture research. Mathematical descriptions of each model are presented in Table 1.

123 The simplest model,  $M_0$ , estimates a single detection probability that is common across  
124 all nests and sampling occasions. In reality, individuals may differ in their probability of being  
125 detected, a phenomenon known as individual heterogeneity. Consequently, the model  $M_H$   
126 estimates both a mean and variance in detection probabilities across nests (i.e., individual  
127 random effects). In our study of bumble bee nests, this could be due to unmeasured differences in  
128 the worker activity level or the location of nests that increase the probability particular nests will  
129 be detected. Similarly, the model  $M_T$  estimates a mean detection probability across nests and a

130 variance in detection associated with visits (i.e., a temporal random effect). For bumble bees, this  
131 is could be due to temporal variation in colony size, ambient temperature (which potentially  
132 affects activity level), or seasonal changes in vegetation within plots that alter the probability  
133 nests will be detected on each visit. The model  $M_b$  described by Otis et al. (1978) accounts for a  
134 discrete behavioral change in organisms that affects their individual detection in subsequent  
135 marking occasions, commonly referred to as “trap-shyness” or “trap-happiness”. In studies of  
136 sessile organisms, rather than behavioral changes of the study organisms themselves, this  
137 response can plausibly occur if the vegetation surrounding focal organisms becomes trampled by  
138 researchers or if researchers remember the location of individuals (in this case, nests) they have  
139 previously located. Either of these scenarios would result in different detection probabilities for  
140 initial and subsequent capture events in studies of sessile organisms.

141         Although not explicitly described by Otis et al. (1978), fixed effects of explicit covariates  
142 can also be incorporated to examine the drivers of variation in detection probability (Kéry and  
143 Schaub 2012). These can include age or size of the organism, habitat covariates, or explicit  
144 temporal covariates (e.g., to examine temporal trends in detection). In our study, four of the eight  
145 models included explicit covariates affecting detection of bumble bee nests. Towards this goal,  
146 we constructed a model that included different detection probabilities for each survey plot  
147 ( $M_{plot}$ ; see *Study Site* in *Methods* section below), and three models that included effects of  
148 vegetation height above the nest ( $M_{veg}$ ), hour of the day at which plots were surveyed ( $M_{hour}$ ),  
149 and ambient air temperature during the survey ( $M_{temp}$ ). Our objective was to demonstrate how  
150 these effects could be incorporated to generate deeper insights into the processes influencing nest  
151 abundance surveys, rather than to exhaustively examine the diverse suite of (potentially  
152 interacting) factors that influence nest detection.

153 **Methods**

154 ***Study Site***

155 We conducted searches for bumble bee nests in three survey plots located at Appleton  
156 Farms (42.65°N, 70.86°W) in Ipswich, Massachusetts. Each plot was approximately 3000 m<sup>2</sup> in  
157 area and not used for agriculture, although each plot is mowed annually to prevent succession.  
158 Two plots were adjacent to one another, while the third plot was located approximately 1,000  
159 meters away. Primary vegetation cover in each plot consisted of a variety of grasses, sedges  
160 (*Carex* spp.), perennial forbs (e.g., *Plantago lanceolata*, *Linaria vulgaris*, *Lotus corniculatus*,  
161 *Solidago* spp., *Asclepias* spp.), and ericaceous shrubs (e.g., *Vaccinium angustifolium*, *Rubus*  
162 spp.). Each plot was bordered by a hedgerow of trees or forest, and the surrounding landscape  
163 was mixed agriculture (pasture and hay fields) and natural areas (forest and wet meadows).

164 ***Data Collection***

165 Each plot was visited repeatedly in July and August of 2017, after bumble bee colonies  
166 had produced several cohorts of workers, between 6:30 am and 7:30 pm, though the majority of  
167 nest surveys took place in the morning. Searches were conducted independently by eight  
168 different investigators, several of whom surveyed each plot multiple times. Each plot was  
169 surveyed for two hours between 12-16 times, for a total effort of 24-32 survey hours per plot.  
170 During each survey, searchers moved slowly through the plot looking for bumble bee activity  
171 that might indicate the presence of nest (i.e., workers quickly descending to the ground, slowly  
172 ascending from the ground, or conducting circular navigation flight behavior). Upon locating a  
173 bumble bee nest, searchers placed an inconspicuous, numbered identifier next to the nest  
174 entrance and recorded the nest location, identity, and whether the nest had been previously  
175 located either by themselves or by other searchers. For each located nest, we recorded the height



176 of the tallest vegetation immediately above the nest entrance, which could affect searcher ability  
177 to detect bumble bee movement near nest entrances. We extracted hourly ambient air  
178 temperature measurements during each survey from the nearest weather station (approx. 9.5 km  
179 away); ambient air temperature could affect bumble bee metabolism and nest activity, which  
180 could influence our ability to detect nests. Finally, we also recorded the time of day of each  
181 survey, as bumble bee activity varies throughout the day (Kwon and Saeed 2003).

## 182 *Statistical Methods*

183 The foundation of mark-recapture approaches is the encounter histories of individuals (in this  
184 case bumble bee colonies) that are generated from repeated surveys of plots. Closed population  
185 models assume the abundance of individuals within each plot does not change across sampling  
186 periods; thus, variation in the number of individuals detected across repeated surveys is caused  
187 entirely by observation error. The goal of closed population models is to estimate detection  
188 probability ( $p$ ) of individuals along with the spatial, temporal, and/or individual-level factors that  
189 influence it. Once the factors that influence detection probability are estimated, the observed  
190 count of individuals can be corrected to generate an unbiased estimate of the true number of  
191 individuals present at a site.

192 Successful detection of the individual  $i$  during each survey  $t$  occurs with some  
193 probability  $p_{i,t}$  and the probability of not detecting the individual is  $1 - p_{i,t}$ . An encounter  
194 history of '0110' implies the individual was detected on the second and third survey of the plot  
195 and not detected on the first and fourth survey. The entire encounter history '0110' for individual  
196  $i$  therefore occurs with probability  $(1 - p_{i,1}) \times p_{i,2} \times p_{i,3} \times (1 - p_{i,4})$ . The actual encounter  
197 history (i.e., the observed sequence of 0's and 1's) is assumed to arise from a series of Bernoulli

198 trials. Additionally, a link function can be used to model the influence of covariates on detection  
199 probability, analogous to a logistic regression.

200 We fit the series of closed population models described in Table 1 using the empirical nest  
201 encounter histories generated by our repeated surveys of plots. We fit models using Bayesian  
202 methods, outlined by Kéry and Schaub (2012, ch 6); though we note that such models can also  
203 be fit in a frequentist framework using maximum likelihood approaches. Bayesian analysis  
204 allows for random effects to be easily incorporated, for nests to be right-censored part way  
205 through the study (e.g., if a nest was known to have failed, which violates an assumption of  
206 closed population models), and for uncertainty in parameter estimates to be easily propagated to  
207 the model output. The Bayesian implementation of closed population models uses an additional  
208 technique called “data augmentation” to estimate the true number of individuals in a plot, based  
209 on the number of nests actually detected and their estimated detection probabilities (see Kéry and  
210 Schaub 2012, ch. 6; Royle and Dorazio 2008, section 5.6 for further discussion of this technique;  
211 also see our implementation of this approach in Supplementary Material 1).

### 212 *Simulation Study*

213 The methods described above demonstrate the application of a mark-recapture approach  
214 to the study of sessile insect life cycle stages when detection is highly imperfect. However, to  
215 further illustrate the value of this approach, we conducted a simulation to illustrate a second (and  
216 converse) problem of failing to account for imperfect detection: imperfect detection generates  
217 spurious sampling variability in counts, even when no variation exists. This simulation was also  
218 used to examine whether the models we used would converge with only two visits (the minimum  
219 number of visits required for mark-recapture) if a larger number of plots had been surveyed. We  
220 simulated 40 hypothetical plots, each containing exactly 5 bumble bee nests. We used a fixed

221 detection probability of 0.30 for all nests (approximately equal to the mean detection probability  
222 we estimated; see Results section and Fig. 1a). We then simulated the number of nests observed  
223 on a single visit to each plot by drawing from a binomial distribution with 5 trials for each plot  
224 (i.e., one for each nest). We used this simulation to evaluate 1) the degree of spurious variability  
225 introduced into counts simply by imperfect detection, and 2) whether the closed population  
226 mark-recapture approach could correct for this spurious variability with only two visits to each  
227 plot.

228 To facilitate the straightforward application of this analytical toolkit to other studies of  
229 sessile insect life cycle stages, we include full encounter history data, covariate data, and  
230 commented R code for Bayesian analysis and simulations as Supplementary Materials 1-9. Data  
231 management and simulations were conducted in R version 3.4.4 and Bayesian analysis was  
232 conducted in JAGS using the *jagsUI* library in R (Kellner 2018).

233

## 234 **Results**

235 We located 18 bumble bee nests across the three survey plots (10, 5, and 3 nests in each  
236 plot, respectively). All nests were constructed by *Bombus impatiens*, except for one that was  
237 constructed by *B. bimaculatus*. We used all nests for subsequent analysis. The number of nests  
238 located on single visits to each of the three plots ranged from 0 to 6, 0 to 3, and 0 to 3 for each  
239 plot, respectively. The three plots were searched 11, 16, and 14 times by at least 6 different  
240 observers.

### 241 ***Model results – patterns in nest detection***

242 The mean detection probability of nests based on the intercept-only model ( $M_0$ ) was 0.30  
243 (95% credible interval [CRI] = 0.24 - 0.36; Fig. 1a). A model that included heterogeneity in  $p$

244 across nests ( $M_H$ ) suggested that nests differed in their individual detection probabilities, with a  
245 median posterior estimate for  $p$  of 0.26 (95% CRI = 0.11 to 0.37) and standard deviation (on the  
246 logit scale) of 0.79 (95% CRI = 0.14 to 1.84; Fig. 1b). Similarly, a model that included  
247 heterogeneity in  $p$  across visits ( $M_T$ ) indicated that detection probability varied through time with  
248 a median  $p$  of 0.27 (95% CRI = 0.20 to 0.38) and standard deviation (on the logit scale) of 0.67  
249 (95% CRI = 0.10 to 1.29; Fig. 1b). We note that there was substantial uncertainty associated with  
250 estimates of both individual and temporal random effects, as is common for random effect  
251 models fit to relatively sparse data (Kéry and Schaub 2012).

252         There was weak evidence for different detection probabilities between the first and  
253 subsequent capture occasions (Fig. 1c). Thus, nests were not more likely to be detected after their  
254 initial discovery.

255         A model including different detection probabilities for nests within each survey plot  
256 ( $M_{plot}$ ) indicated nest detectability varied systematically across plots (Fig. 1d). Under this  
257 model, median estimates of detection probabilities in each plot were 0.36 (95% CRI = 0.27 to  
258 0.45), 0.23 (95% CRI = 0.14 to 0.33), and 0.26 (95% CRI = 0.15 to 0.41). The probability that  
259 detection probability was greater for nests in plot 1 than plot 2 was 0.98 (calculated directly from  
260 posterior probability distributions).

261         We then constructed a series of models to examine effects of specific covariates on  $p$ .  
262 Height of vegetation above the nest entrance did not have a strong effect on  $p$  (standardized  
263 effect of vegetation height from  $M_{veg} = 0.02$ ; 95% CRI = -0.02 to 0.06; Fig. 1e). However,  $p$   
264 declined strongly throughout the day during sampling times (standardized effect of hour from  
265  $M_{hour} = -0.68$ ; 95% CRI = -1.25 to -0.25; Fig. 1e) and was negatively correlated with ambient

266 air temperature during the survey (standardized effect of air temperature from  $M_{temp} = -0.38$ ;  
267 95% CRI = -0.69 to -0.09; Fig. 1e).

### 268 ***Model results – nest abundance***

269 Across all eight models, median estimates of nest abundance were in close agreement.  
270 All models estimated approximately 10, 5, and 3 nests in each of the three plots, respectively  
271 (Fig. 2). The corresponding median estimate of nest density in each plot was therefore 33.3, 16.7,  
272 and 10 nests·ha<sup>-1</sup>. Consequently, on single surveys of each plot, we located approximately 0-60%  
273 of the nests in plots 1 and 2, and 0-100% of the nests in plot 3.

274 However, even with our large number of repeated searches (11-16 per survey plot), there  
275 was a high probability that undiscovered nests remained in each plot at the end of our study (Fig.  
276 2; note range of credible intervals). For example, while the median estimate of total nest  
277 abundance from model  $M_0$  was 18 (equal to the number of nests we located), the 95% credible  
278 interval was 18 to 20, and the probability that the true abundance was greater than 18 was 0.25.  
279 Notably, the credible intervals for estimates from model  $M_H$  were wide relative to other models.  
280 This reflects two important features of individual heterogeneity: 1) a fraction of nests have  
281 extremely low detection probabilities and it is difficult to estimate how many remained  
282 undetected, and 2) the existing amount of heterogeneity is difficult to estimate, especially with  
283 low sample sizes (see Fig. 1b).

284 Given that our counts of bumble bee nests were subject to substantial observation error,  
285 we performed an additional analysis to illustrate how imperfect detection can obscure large  
286 differences in the density of colonies among plots. We randomly selected nest counts from  
287 single surveys to each plot, calculated the resulting rank order of nest densities, and compared  
288 them to our estimates from mark-recapture models. We repeated this process 1000 times. This

289 analysis revealed that based on single visits to each field, the incorrect rank-order of density  
290 between plots arises 70% of the time despite large differences in the relative density of nests in  
291 each plot. Strikingly, the incorrect rank order between plots 1 and 3 arises 17% of the time,  
292 despite a three-fold difference in estimated nest density between these two plots (33.3 vs 10  
293 nests·ha<sup>-1</sup>). Next, to quantify the effort needed to reliably estimate differences in nest density  
294 between plots, we sequentially re-fit model  $M_0$  for different numbers of visits. With our small  
295 number of plots and so few nests initially detected, the model would not converge with only 2  
296 visits to each plot. This also occurred when models were fit with maximum likelihood in  
297 program MARK instead of using Bayesian methods. The model converged with 3 visits to each  
298 plot, but uncertainty associated with abundance estimates was extremely large (Fig. 3). After 5  
299 visits to each plot, clear differences in abundance between plots 1 and 3 were apparent. As  
300 expected, uncertainty in estimates continued to decline as the number of surveys increased.

### 301 *Simulation results – correcting spurious variation in nest abundance*

302 Our final simulation illustrates the sampling variability induced into count data even  
303 when detection rates are fixed at 0.3. We then used this simulation to examine whether this  
304 spurious source of variability could be corrected by visiting each plot only twice if a larger  
305 number of plots were visited. Despite the true presence of exactly 5 nests in each of the 40  
306 simulated plots (Fig. 4; solid black line), the number of nests detected per plot ranged from 0 to 4  
307 with a mean of 1.51 and standard deviation across plots of 0.95 (Fig. 4, dashed line and dots).  
308 Thus, based on a single visit to each plot, there is considerable (but spurious) variation in nest  
309 abundance across the 40 plots. We then simulated a second visit to each plot, fit mark-recapture  
310 model  $M_0$  to the resulting encounter histories, and estimated the number of nests in each plot  
311 while accounting for imperfect detection. The model successfully converged with only 2 visits to

312 each plot and the credible intervals (Fig. 4; gray ribbons) overlapped the true abundance in every  
313 plot. The model  $M_0$  therefore correctly indicates weak evidence for variation in abundance  
314 among the 40 simulated plots, despite the high degree of variation in nest counts from a single  
315 survey. Thus, for larger sample sizes of plots, two to three visits to each may be sufficient in  
316 estimating nest densities.

317

## 318 **Discussion**

319 Our study is the first to apply mark-recapture methods to estimate the density of bumble bee  
320 nests, which represent a critical and understudied life cycle stage for this important pollinator  
321 group. In our study, single surveys of bumble bee nest abundance were subject to considerable  
322 bias and observation error, owing to imperfect detection. On average, we only detected 30% of  
323 existing nests on each 2-hour survey of a 0.3 ha plot, and we show that the low detection  
324 probability on single surveys can introduce substantial and spurious variation into counts (Fig.  
325 4). Thus, in order to understand the nesting ecology and monitoring requirements of bumble  
326 bees, imperfect detection of nests must be properly accounted for.

327 Our estimate of detection probability is well within the range of reported rates for surveys of  
328 other sessile organisms: 0.01 to 1.0 in plants (Chen et al. 2013; Kellner and Swihart 2014), 0.17  
329 to 0.60 for insect nests (Berberich et al. 2016; Brown et al. 2017), and 0.09 to 0.93 for patches of  
330 freshwater mussels (Reid 2016). These studies use mark-recapture approaches to improve  
331 estimates of density or occupancy on the landscape for sessile organisms that are not perfectly  
332 detectable, and our study adds bumble bees to this list of taxa.

333 The range of bumble bee nest densities we detected are comparable to those reported in  
334 Osborne et al. (2008), who used intensive fixed searches by citizen science volunteers to count

335 nests in UK gardens and countryside habitats. Of our three plots, the highest density we detected  
336 was 33.3 nests·ha<sup>-1</sup>, similar to hedgerow (29.5 nests·ha<sup>-1</sup>), garden (35.8 nests·ha<sup>-1</sup>), and fence line  
337 (37.2 nests·ha<sup>-1</sup>) habitats reported in Osborne et al. (2008). Notably, Cumber (1953) is the only  
338 other study to report higher nest densities than these; his estimate of 48.6 nests·ha<sup>-1</sup> was based on  
339 intensive free searches of a refuse dump in England. Conversely, our lowest density plot  
340 contained 10 nests·ha<sup>-1</sup>, which is similar to the density of 10.9 nests·ha<sup>-1</sup> reported in Harder  
341 (1986) who intensively surveyed a successional field in Ontario, Canada. This estimate is also  
342 similar to the lowest densities in Osborne (10.8 nests·ha<sup>-1</sup> in woodland and 11.4 nests·ha<sup>-1</sup> in  
343 short grassland habitat). Therefore, our three plots seem to have captured the range of nest  
344 densities observed in other studies, if we restrict these studies to those with intensive search  
345 effort and extremely high detection probabilities.

346 Other studies have reported far lower nest densities than those in our study or those in  
347 Osborne et al. (2008), Cumber (1953), and Harder (1986). However, comparisons among studies  
348 are ultimately hampered by differences in survey efforts, and thus, differences in detection error.  
349 For example, low-intensity free searches by researchers or volunteers produced estimates of nest  
350 density ranging from 1.4 to 3.6 nests·ha<sup>-1</sup> similar to the range of nest densities discovered by  
351 bumble bee “sniffer dogs” (O’Connor et al. 2012, 2017). Both of these studies acknowledge that  
352 detection error is likely substantial for these methods. Molecular studies also typically yield  
353 estimates of nest density that are 1-2 orders of magnitude lower than intensive ground-based  
354 searches (range: 0.13 to 1.9 nests·ha<sup>-1</sup>; Supplement 2). Several molecular studies have used ad-  
355 hoc approaches to account for imperfect detection, but these approaches likely under-estimate  
356 the true nest density (Goulson et al. 2010). Molecular methods also integrate habitat quality over  
357 larger spatial extents than ground-based surveys, and likely incorporate areas that are unsuitable



358 for nesting (e.g., water bodies). Formal mark-recapture approaches are necessary to understand  
359 the degree to which variation in nest density between studies is driven by ecologically relevant  
360 factors (e.g., variation in habitat quality, differences in spatial scale at which studies occur)  
361 versus unresolved differences in imperfect detection.

362 In addition to estimating overall probability of nest detection, we found that nest detection  
363 declined when surveys were conducted later in the day and in warmer temperatures (Fig. 1e).  
364 Based on estimates from model  $M_{hour}$ , mean detection probability during 6 am surveys was 0.40  
365 (95% CRI = 0.31 to 0.50), but was only 0.05 (95% CRI = 0.01 to 0.17) for surveys initiated at 6  
366 pm. This result is consistent with Kwon and Saeed (2003) who found that colony traffic and  
367 foraging activity of *Bombus terrestris* declined throughout the day and when temperatures were  
368 warmer. Similarly, Couvillon et al. (2010) reported that workers of all sizes conducted fewer  
369 foraging trips in warmer temperatures. Thus, the results of our mark-recapture estimate of  
370 detection probability are broadly consistent with prior knowledge of bumble bee foraging  
371 ecology, and suggest that variation in forager behavior is a likely driver of differences in  
372 detectability.

373 Although we did not measure them in our study, other factors could also influence colony  
374 activity, and in turn, the probability that nests are detected on a given survey. For example, larger  
375 colonies have higher traffic at nest entrances (Kwon and Saeed 2003) and are therefore likely to  
376 be more detectable. Colony size, in turn, depends on floral resources available throughout the  
377 season (Williams et al. 2012; Crone and Williams 2016). Thus, detection of nests may differ  
378 between high and low quality habitat owing to systematic differences in colony size. Here, we  
379 found evidence for systematic differences in nest detection across our three survey plots. Plot 1  
380 had the highest nest density (33.3 nests·ha<sup>-1</sup>), and simultaneously, detection was highest for nests

381 in this plot (Fig. 1d). Mark-recapture approaches can correct for these spatio-temporal biases in  
382 detection probability, which in turn, improve comparisons of abundance within and between  
383 studies.

384 Population monitoring schemes will often have multiple objectives, in addition to generating  
385 reliable estimates of population size (e.g., sampling large numbers of individuals to assess body  
386 condition, biochemistry, or disease status; monitoring behavior of individual animals in different  
387 environmental contexts). For rare or highly cryptic species, intensive surveys over a small area  
388 may be required to generate reliable estimates of local abundance, but yield few individuals for  
389 further detailed study. In the case of bumble bees, intensive fixed searches of small plots are  
390 considered sufficient to detect all existing nests and therefore reliably estimate nest density  
391 (Osborne *et al.*, 2008; O'Connor *et al.*, 2012). However, intensive fixed searches are inefficient  
392 when a large number of nests are desired (e.g., for examining microsite characteristics associated  
393 with nests, or for studying workers at nest entrances) and are logistically challenging to  
394 implement over large areas without distributed citizen science networks (Osborne *et al.* 2008;  
395 Lye *et al.* 2012). We were able to locate a high number of nests, generate precise estimates of  
396 abundance with only 5-6 surveys in each of our 3 plots (Fig. 3a), and produce relatively unbiased  
397 estimates with an even smaller number of surveys at a larger number of sites (Fig 4). This  
398 equates to much lower search effort than typical intensive fixed searches, while simultaneously  
399 locating nests at a comparable rate to low-intensity free searches (O'Connor *et al.* 2012). Our  
400 study illustrates the advantage of mark-recapture for optimizing survey protocols for cryptic and  
401 sessile organisms. Further research in this area will be valuable in illuminating the ecological  
402 drivers of pollinator nesting ecology, a critical but understudied subject.

403

404 **Authors' Contributions**

405 All authors conceived of research ideas, designed methodology, and collected field data.  
406 DI and EC led statistical analysis and writing of the manuscript. All authors contributed critically  
407 to the drafts and gave final approval for publication.

408

409 **Data Accessibility**

410 All data and code associated with these analyses will be archived in Dryad Digital  
411 Repository upon acceptance of this manuscript.

412

413 **Conflict of Interest**

414 The authors declare they have no conflict of interest.

415

416 **Compliance with Ethical Standards**

417 The authors declare that they have complied with ethical standards.

418

419 **References**

420 Anderson DR (2001) The need to get the basics right in wildlife field studies. *Wildl Soc Bull*

421 (1973-2006) 29:1294–1297

422 Berberich GM, Dormann CF, Klimetzek D, Berberich MB, Sanders NJ, Ellison AM (2016)

423 Detection probabilities for sessile organisms. *Ecosphere* 7:e01546.

424 <https://doi.org/10.1002/ecs2.1546>

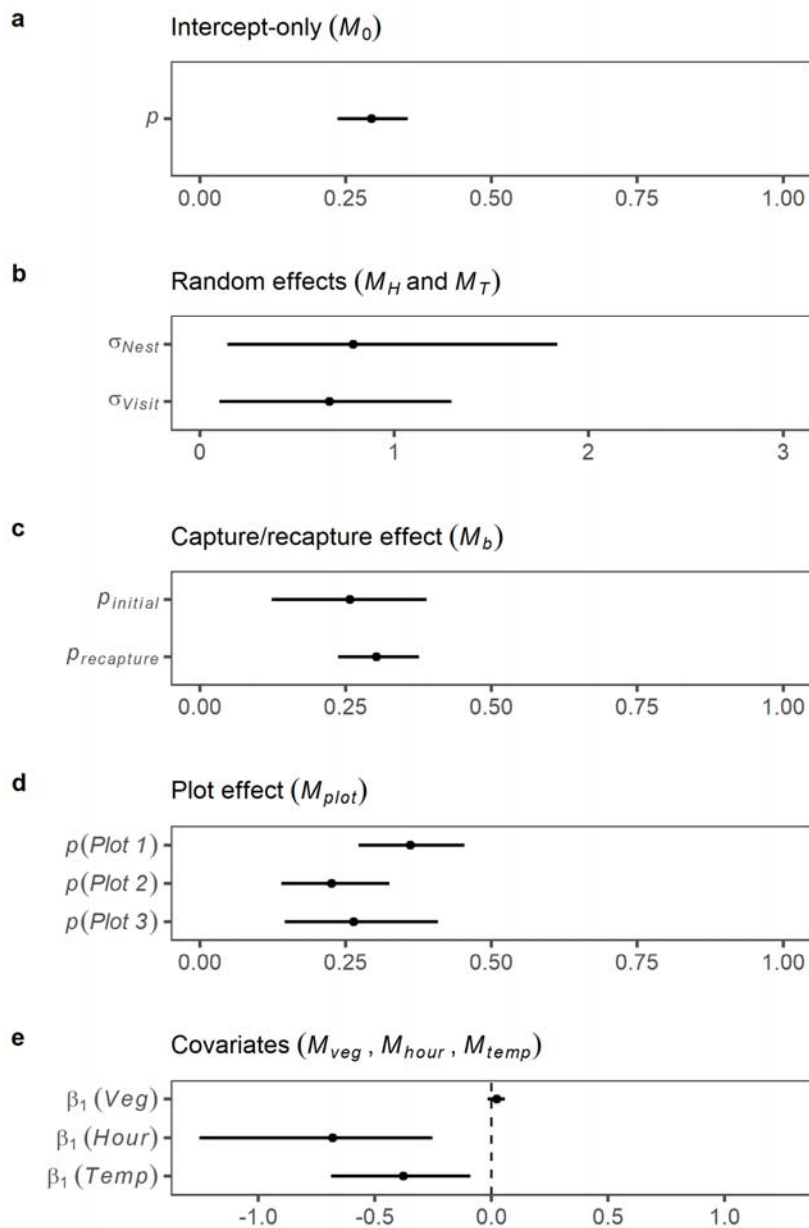
- 425 Brown LM, Breed GA, Severns PM, Crone EE (2017) Losing a battle but winning the war:  
426 moving past preference–performance to understand native herbivore–novel host plant  
427 interactions. *Oecologia* 183:441–453. <https://doi.org/10.1007/s00442-016-3787-y>
- 428 Chen G, Kéry M, Plattner M, Ma K, Gardner B (2013) Imperfect detection is the rule rather than  
429 the exception in plant distribution studies. *J Ecol* 101:183–191.  
430 <https://doi.org/10.1111/1365-2745.12021>
- 431 Couvillon MJ, Fitzpatrick G, Dornhaus A (2010) Ambient air temperature does not predict  
432 whether small or large workers forage in bumble bees (*Bombus impatiens*). *Psyche: A*  
433 *Journal of Entomology* 2010:536430. <http://dx.doi.org/10.1155/2010/536430>
- 434 Crone EE, Williams NM (2016) Bumble bee colony dynamics: quantifying the importance of  
435 land use and floral resources for colony growth and queen production. *Ecol Lett* 19:460–  
436 468. <http://dx.doi.org/10.1111/ele.12581>
- 437 Cumber R (1953) Some aspects of the biology and ecology of humble-bees bearing upon the  
438 yields of red-clover seed in New Zealand. *New Zealand Journal of Science and*  
439 *Technology* 34:227–240
- 440 Darvill B, Knight ME, Goulson D (2004) Use of genetic markers to quantify bumblebee foraging  
441 range and nest density. *Oikos* 107:471–478. [https://doi.org/10.1111/j.0030-  
442 1299.2004.13510.x](https://doi.org/10.1111/j.0030-1299.2004.13510.x)
- 443 Goulson D (2010) *Bumblebees: behaviour, ecology, and conservation*, 2nd edn. Oxford  
444 University Press, New York, pp 317
- 445 Goulson D, Lepais O, O'Connor S, Osborne JL, Sanderson RA, Cussans J, Goffe L, Darvill B  
446 (2010) Effects of land use at a landscape scale on bumblebee nest density and survival. *J*  
447 *Appl Ecol* 47:1207–1215. <https://doi.org/10.1111/j.1365-2664.2010.01872.x>

- 448 Gu W, Swihart RK (2004) Absent or undetected? Effects of non-detection of species occurrence  
449 on wildlife–habitat models. *Biol Conserv* 116: 195–203. [https://doi.org/10.1016/S0006-](https://doi.org/10.1016/S0006-3207(03)00190-3)  
450 [3207\(03\)00190-3](https://doi.org/10.1016/S0006-3207(03)00190-3)
- 451 Harder LD (1986) Influences on the density and dispersion of bumble bee nests (Hymenoptera:  
452 Apidae). *Ecography* 9:99–103. <https://doi.org/10.1111/j.1600-0587.1986.tb01196.x>
- 453 Heard M, Carvell C, Carreck N, Rothery P, Osborne J, Bourke A (2007) Landscape context not  
454 patch size determines bumble-bee density on flower mixtures sown for agri-environment  
455 schemes. *Biol Lett* 3:638–641. <https://doi.org/10.1098/rsbl.2007.0425>
- 456 Herrmann F, Westphal C, Moritz RF, Steffan Dewenter I (2007) Genetic diversity and mass  
457 resources promote colony size and forager densities of a social bee (*Bombus pascuorum*)  
458 in agricultural landscapes. *Mol Ecol* 16:1167–1178. [https://doi.org/10.1111/j.1365-](https://doi.org/10.1111/j.1365-294X.2007.03226.x)  
459 [294X.2007.03226.x](https://doi.org/10.1111/j.1365-294X.2007.03226.x)
- 460 Kellner KF, Swihart RK (2014) Accounting for imperfect detection in ecology: a quantitative  
461 review. *PLoS ONE* 9:e111436. <https://doi.org/10.1371/journal.pone.0111436>
- 462 Kellner KF (2018) jagsUI: A Wrapper Around 'rjags' to Streamline 'JAGS' Analyses. R package  
463 version 1.5.0. <https://CRAN.R-project.org/package=jagsUI>
- 464 Kells AR, Goulson D (2003) Preferred nesting sites of bumblebee queens (Hymenoptera:  
465 Apidae) in agroecosystems in the UK. *Biol Conserv* 109:165–174.  
466 [https://doi.org/10.1016/S0006-3207\(02\)00131-3](https://doi.org/10.1016/S0006-3207(02)00131-3)
- 467 Kéry M, Schaub M (2012) Bayesian population analysis using WinBUGS: a hierarchical  
468 perspective. Academic Press, Cambridge
- 469 Kéry M, Schmidt B (2008) Imperfect detection and its consequences for monitoring for  
470 conservation. *Community Ecol* 9:207–216. <https://doi.org/10.1556/ComEc.9.2008.2.10>

- 471 Kwon YJ, Saeed S (2003) Effect of temperature on the foraging activity of *Bombus terrestris*  
472 L.(Hymenoptera: Apidae) on greenhouse hot pepper (*Capsicum annuum* L.). Appl  
473 Entomol Zool 38:275–280. <https://doi.org/10.1303/aez.2003.275>
- 474 Lye GC, Osborne JL, Park KJ, Goulson D (2012) Using citizen science to monitor *Bombus*  
475 populations in the UK: nesting ecology and relative abundance in the urban environment.  
476 J Insect Conserv 16:697–707. <https://doi.org/10.1007/s10841-011-9450-3>
- 477 O'Connor S, Park KJ, Goulson D (2012) Humans versus dogs; a comparison of methods for the  
478 detection of bumble bee nests. J Apicult Res 51:204–211.  
479 <https://doi.org/10.3896/IBRA.1.51.2.09>
- 480 O'Connor S, Park KJ, Goulson D (2017) Location of bumblebee nests is predicted by counts of  
481 nest-searching queens. Ecol Entomol 42:731–736. <https://doi.org/10.1111/een.12440>
- 482 Osborne JL, Martin AP, Shortall CR, Todd AD, Goulson D, Knight ME, Hale RJ, Sanderson RA  
483 (2008) Quantifying and comparing bumblebee nest densities in gardens and countryside  
484 habitats. J Appl Ecol 45:784–792. <https://doi.org/10.1111/j.1365-2664.2007.01359.x>
- 485 Otis DL, Burnham KP, White GC, Anderson DR (1978) Statistical inference from capture data  
486 on closed animal populations. Wildl Monogr 62: 3–135
- 487 Rao S, Skyrn KM (2013) Nest Density of the Native Bumble Bee, *Bombus nevadensis* Cresson  
488 (Hymenoptera: Apoidea), in an Agricultural Landscape. J Kans Entomol Soc 86:93–97.  
489 <https://doi.org/10.2317/JKES120708.1>
- 490 Reid S (2016) Search effort and imperfect detection: Influence on timed-search mussel (Bivalvia:  
491 Unionidae) surveys in Canadian rivers. Knowl Manag Aquat Ecosyst 1–8.  
492 <https://doi.org/10.1051/kmae/2016004>

- 493 Slade NA, Alexander HM, Dean Kettle W (2003) Estimation of population size and probabilities  
494 of survival and detection in Mead's milkweed. *Ecology* 84:791–797.  
495 [https://doi.org/10.1890/0012-9658\(2003\)084\[0791:EOPSAP\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2003)084[0791:EOPSAP]2.0.CO;2)
- 496 Svensson B, Lagerlöf J, Svensson BG (2000) Habitat preferences of nest-seeking bumble bees  
497 (Hymenoptera: Apidae) in an agricultural landscape. *Agr Ecosyst Environ* 77:247–255.  
498 [https://doi.org/10.1016/S0167-8809\(99\)00106-1](https://doi.org/10.1016/S0167-8809(99)00106-1)
- 499 Waters J, O'Connor S, Park KJ, Goulson D (2010) Testing a detection dog to locate bumblebee  
500 colonies and estimate nest density. *Apidologie* 42:200–205.  
501 <https://doi.org/10.1051/apido/2010056>
- 502 White GC, Burnham KP (1999) Program MARK: survival estimation from populations of  
503 marked animals. *Bird Study* 46:S120–S139. <https://doi.org/10.1080/00063659909477239>
- 504 Williams BK, Nichols JD, Conroy MJ (2002) Analysis and management of animal populations.  
505 Academic Press
- 506 Williams NM, Regetz J, Kremen C (2012) Landscape-scale resources promote colony growth  
507 but not reproductive performance of bumble bees. *Ecology* 93:1049–1058.  
508 <https://doi.org/10.1890/11-1006.1>
- 509  
510  
511

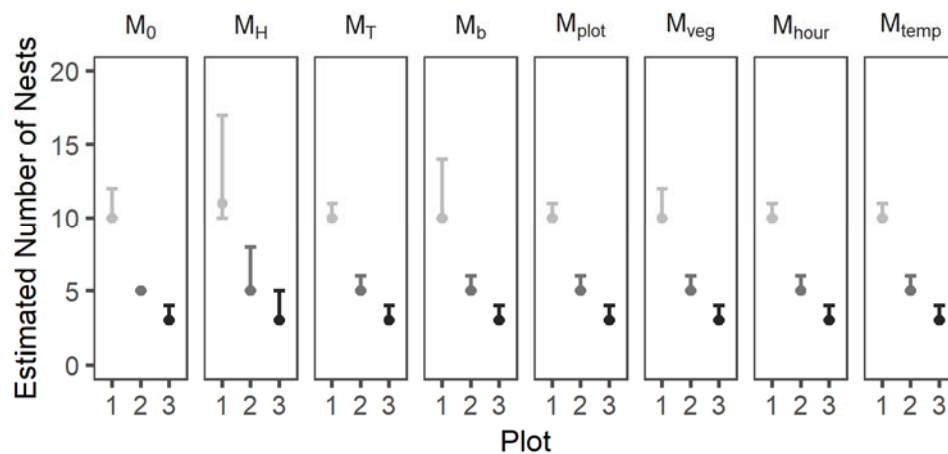
512 **Figures**



513

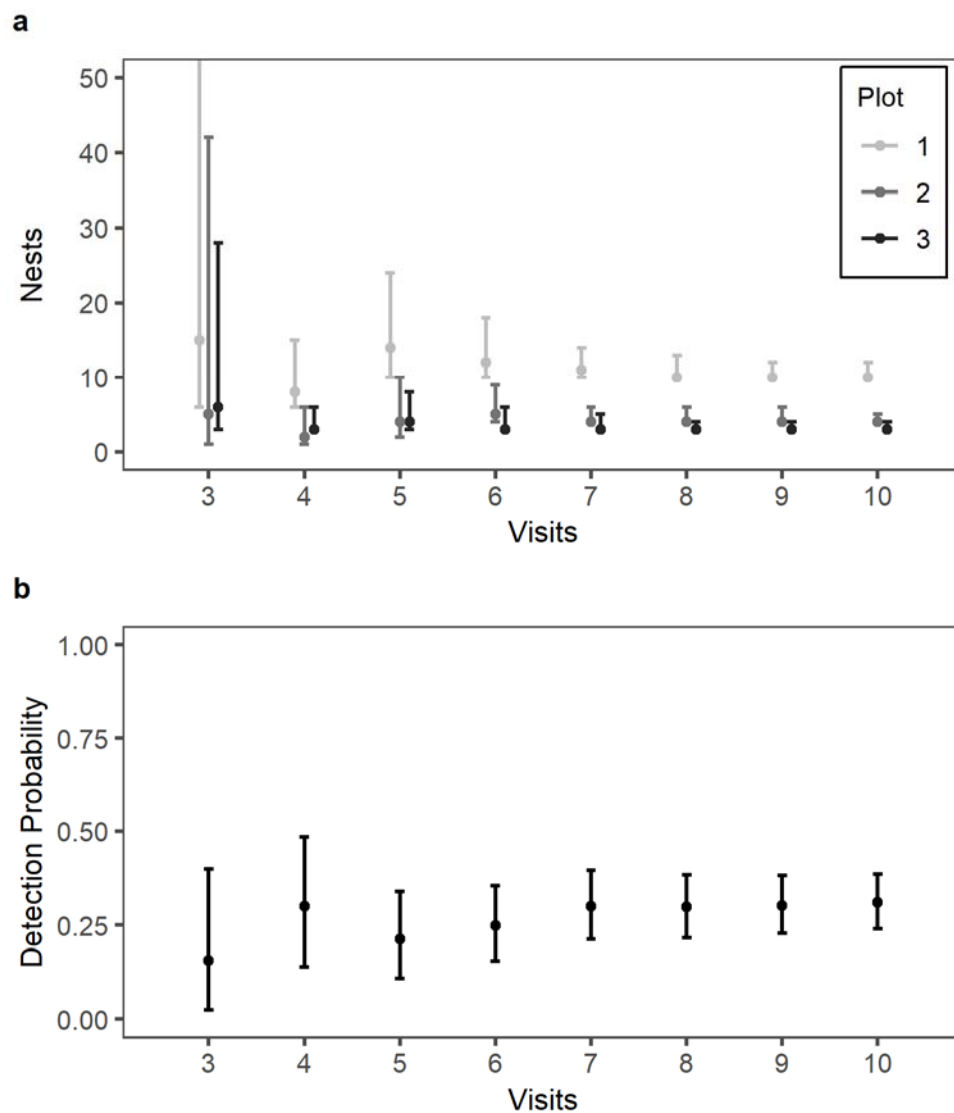
514 **Fig. 1** Effects on bumble bee nest detection probability from closed population models (n = 18  
515 nests). Points represent median estimate of effect from Bayesian posterior distributions; lines  
516 represent 95% credible intervals





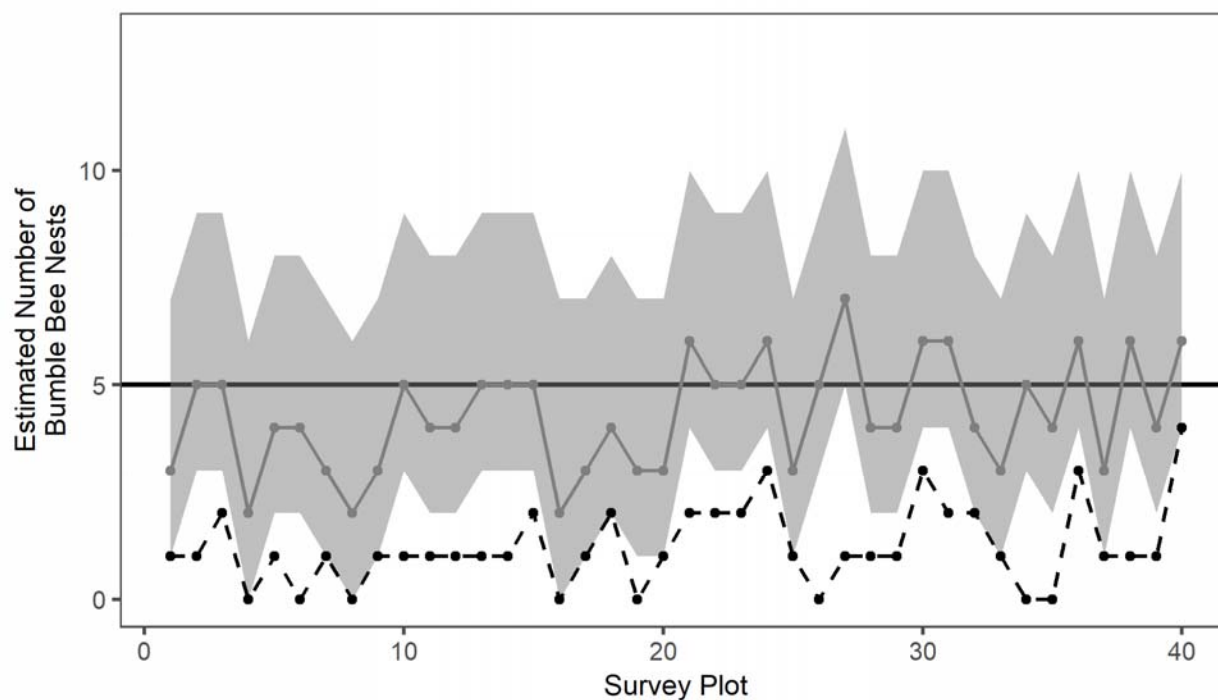
517

518 **Fig. 2** Estimated abundance of nests in each of the three survey plots (with associated 95%  
519 Bayesian credible intervals) based on each of the eight closed population models and all survey  
520 data



521

522 **Fig. 3** Estimated nests in each plot and associated mean detection probability from closed  
523 population models, based on model  $M_0$  after different numbers of visits to each plot. Points  
524 represent median estimates from Bayesian posterior distribution; lines denote associated 95%  
525 credible intervals



526

527 **Fig. 4** Results of a simulation with 40 survey plots, each with exactly 5 nests, and each nest with  
528 a 0.3 probability of detection on a single visit. Black dashed lines and black dots represent counts  
529 on a single visit to each plot. Gray lines and dots depict estimates based on a closed population  
530 model with 2 visits to each plot. Gray ribbon indicates 95% credible intervals for estimates in  
531 each plot. Thick solid line is true number of nests in each plot ( $n = 5$ )

532 **Tables**

533 **Table 1** Classic closed population model structures used in our study to examine variation in the  
 534 probability of detecting bumble bee nests.  $p_{i,t}$  refers to the detection probability of nest  $i$  on survey  $t$ ,  
 535 and  $\beta_i$  refer to various fitted model coefficients

536  
 537

#	Model Abbrev.	Model Name	Formula	Description
1	$M_0$	Intercept-only	$\text{logit}(p_{i,t}) = \beta_0$	Constant value of $p$ across all $i$ nests and $t$ visits
<u>Models with Random Effects</u>				
2	$M_H$	Individual heterogeneity	$\text{logit}(p_{i,t}) = \beta_0 + \alpha_i$ with $\alpha_i \sim \text{Normal}(0, \sigma_{nest}^2)$	Random effect for $p$ among $i$ nests
3	$M_T$	Temporal heterogeneity	$\text{logit}(p_{i,t}) = \beta_0 + \alpha_t$ with $\alpha_t \sim \text{Normal}(0, \sigma_{time}^2)$	Random effect for $p$ among $t$ visits
<u>Models with Categorical Covariates</u>				
5	$M_b$	Capture/recapture effect	$\text{logit}(p_{i,t}) = \beta_0 + \beta_1 * \text{recap}_{i,t}$	$p$ differs between initial capture and subsequent recaptures (where $\text{recap}_{i,t}$ is an indicator variable set to 0 until initial capture, 1 after initial capture)
4	$M_{plot}$	Plot effect	$\text{logit}(p_{i,t}) = \beta_0 + \beta_1 * \text{Plot}_i$	$p$ differs among survey plots (i.e., group effect)
<u>Models with Continuous Covariates</u>				
6	$M_{veg}$	Veg. height	$\text{logit}(p_{i,t}) = \beta_0 + \beta_1 * \text{veg}_i$	$p$ depends on vegetation height above the nest entrance
7	$M_{hour}$	Hour of day	$\text{logit}(p_{i,t}) = \beta_0 + \beta_1 * \text{hour}_t$	$p$ depends on the time of day the survey was conducted
8	$M_{temp}$	Air temperature	$\text{logit}(p_{i,t}) = \beta_0 + \beta_1 * \text{temp}_t$	$p$ is affected by air temperature during the survey

538

539

540 **Appendix 1**

541 **Table 2** Summary of previous studies that estimated bumble bee nest density, with description of  
 542 methods used, study region, species detected, habitat types associated with density estimates,  
 543 area searched, and estimated density. Note that search area does not apply to molecular methods  
 544 as nest density is inferred from genetic relatedness between workers and estimates of worker  
 545 foraging range

Study	Method	Study Region	Species Detected	Habitat Type	Search area (ha)	Density (nests·ha <sup>-1</sup> )
Osborne et al. 2008	Fixed Search	UK	Multiple	Grassland <10 cm	0.44	11.4
Osborne et al. 2008	Fixed Search	UK	Multiple	Grassland >10 cm	0.75	14.6
Osborne et al. 2008	Fixed Search	UK	Multiple	Woodland	0.19	10.8
Osborne et al. 2008	Fixed Search	UK	Multiple	Fence line	0.16	37.2
Osborne et al. 2008	Fixed Search	UK	Multiple	Hedgerow	0.41	29.5
Osborne et al. 2008	Fixed Search	UK	Multiple	Woodland edge	0.25	19.9
Osborne et al. 2008	Fixed Search	UK	Multiple	Lrg Garden	0.89	34.9
Osborne et al. 2008	Fixed Search	UK	Multiple	Med Garden	1.13	31.9
Osborne et al. 2008	Fixed Search	UK	Multiple	Sm Garden	0.40	50.4
Osborne et al. 2008	Fixed Search	UK	Multiple	Garden (all)	2.43	35.8
O'Connor et al. 2012	Fixed Search	UK	Multiple	Woodland	0.14	27.8
O'Connor et al. 2012	Detection Dog	UK	Multiple	Woodland	6.94	1.41
O'Connor et al. 2012	Free Search	UK	Multiple	Woodland	6.94	1.44
O'Connor et al. 2017	Free Search	UK	Multiple	Grassland	5.00	3.6
O'Connor et al. 2017	Free Search	UK	Multiple	Woodland	5.00	3
Cumber 1953	Free Search	UK	Multiple	Refuse dump	0.80	48.6
Darvill et al. 2004	Molecular	UK	<i>B. pascuorum</i>	Forest/farmland	NA	1.9
Knight et al. 2005	Molecular	UK	<i>B. pascuorum</i>	Farmland	NA	0.3
Knight et al. 2009	Molecular	UK	<i>B. pascuorum</i>	Farmland	NA	1.7
Darvill et al. 2004	Molecular	UK	<i>B. terrestris</i>	Forest/farmland	NA	0.13
Knight et al. 2005	Molecular	UK	<i>B. terrestris</i>	Farmland	NA	0.3
Knight et al. 2005	Molecular	UK	<i>B. lapidarius</i>	Farmland	NA	1.2
Knight et al. 2005	Molecular	UK	<i>B. pratorum</i>	Farmland	NA	0.3
Waters et al. 2010	Sniffer Dog	UK	<i>B. muscorum</i>	Upland Heath	NA	0.5
Waters et al. 2010	Sniffer Dog	UK	Multiple	Lowland Heath	NA	0.27

Waters et al. 2010	Sniffer Dog	UK	Multiple	Machair	NA	2.13
Waters et al. 2010	Sniffer Dog	UK	Multiple	Dune	NA	1.47
Rao and Skyrme 2013	Free Search	USA	<i>B. nevadensis</i>	Crop field	NA	18.8
Rao and Strange 2012	Molecular	USA	<i>B. vosnesenskii</i>	Crop field	NA	0.76
Harder 1986	Free Search	CAN	Multiple	Old field	3.20	10.93

---

546

## 547 **Description of supplementary information**

548 Supplementary material 1: R code used to examine variation in the probability of detecting  
549 bumble bee nests using bayesian closed population models and generate Fig. 1 and Fig. 2

550

551 Supplementary material 2: R code used to demonstrate that imperfect detection can obscure  
552 differences in colony density among plots and generate Fig. 3

553

554 Supplementary material 3: R code used to estimate bumblebee populations for simulated mark-  
555 recapture data and generate Fig. 4

556

557 Supplementary material 4: Capture histories for 10 bumble bee nests for survey plot 1

558

559 Supplementary material 5: Capture histories for 5 bumble bee nests for survey plot 2

560

561 Supplementary material 6: Capture histories for 3 bumble bee nests for survey plot 3

562

563 Supplementary material 7: Covariate data for each nest, including the height of the tallest  
564 vegetation immediately above the nest entrance

565

566    Supplementary material 8: Hourly ambient air temperature measurements during each survey  
567    from the nearest weather station  
568