

1 **Running Head:** Value of opportunistic individual data

2 **Investigating the potential of opportunistic sighting data to** 3 **inform wildlife conservation strategies**

4 SIMONE TENAN^{1*†}, PAOLO PEDRINI¹, NATALIA BRAGALANTI^{1,2}, CLAUDIO GROFF², CHRIS SUTHERLAND^{3†}

5 ¹ *Vertebrate Zoology Section, MUSE - Museo delle Scienze, Corso del Lavoro e della Scienza 3, 38122 Trento, Italy;*

6 ² *Provincia Autonoma di Trento, Servizio Foreste e Fauna, Via Trener 3, 38100 Trento, Italy;*

7 ³ *Department of Environmental Conservation, University of Massachusetts, Amherst, MA 01003, USA*

8 **ABSTRACT:**

9 Abundance and space use are key population-level parameters used to inform management and conservation decisions
10 of rare and elusive species, for which monitoring resources can be limited, potentially affecting quality of
11 model-based inference. Recently-developed methods that integrate multiple data sources arising from the same
12 ecological process have typically been focused on data from well-defined sampling protocols, i.e. structured data sets.
13 Despite a rapid increase in availability of large datasets, the value of unstructured or opportunistic data to improve
14 inference about spatial ecological processes is, however, unclear. Using spatial capture-recapture (SCR) methods, we
15 jointly analyze opportunistic recovery of biological samples, traditional SCR data resulting from systematic sampling
16 of hair traps and rub trees, and satellite telemetry data, collected on a reintroduced brown bear population in the
17 central Alps. We compared the precision of sex-specific estimates of density and space use derived from models
18 using combinations of data sources ranging from traditional SCR to a fully integrated SCR model that includes both
19 telemetry and opportunistic data. Estimates of density and space use were more precise when unstructured data were
20 added compared to estimates from a classical SCR model. Our results demonstrate that citizen science data lend itself
21 naturally to integration with in the SCR framework and highlight the value of opportunistic data for improving
22 inference about space use, and in turn, of abundance and density. When individual identity and location can be
23 obtained from opportunistic observations, such data are informative about space use and thus have the potential to
24 improve estimates of movement and density using SCR methods. This is particularly relevant in studies of rare or
25 elusive species, where the amount of SCR encounters is usually small, but also budget restrictions and the difficulty of
26 collaring animals limit the number of individuals for which telemetry information is available. Spatially-referenced
27 opportunistic data thus potentially increase both the geographic extent of a study and the number of individuals with

28 available spatial information, providing an improved understanding of how individuals are distributed and how they
29 use space – fundamental components for calibrating conservation management actions.

30 **Key-words:** *Abundance, carnivore monitoring, density, opportunistic data, population estimation, space usage,*
31 *spatial capture-recapture, telemetry.*

32

33 INTRODUCTION

34 Obtaining precise estimates of population density and space use can lead to an increased understanding of the
35 processes governing spatio-temporal ecological dynamics and, in turn, improve wildlife management and
36 conservation practices. The task of estimating ecological state variables is, however, challenging, especially for rare
37 and elusive species such as large carnivores, and requires analytical approaches that account for the fact not all
38 individuals in a population can be observed (Williams et al. 2002). Regardless of methodology, the quality of
39 model-based inference is directly related to data quality, which can be an issue for elusive species, especially when
40 resources for monitoring are limited. This has led to an emphasis on developing methods that integrate multiple data
41 sources (Schaub and Abadi 2010; Gimenez et al. 2014) and, importantly, to a realization that the vast amounts of data
42 regularly collected outside of formal scientific studies, *unstructured* or *opportunistic* data, offer a potentially valuable
43 data source (Dickinson et al. 2012; Newman et al. 2012). Although the majority of data integration methods have
44 focussed on improving estimates of species distribution and temporal population trends, opportunistic data has great
45 potential to improve inferences about spatial ecological processes.

46 Integrated population models (IPMs: Besbeas et al. 2002; Schaub and Abadi 2010; Tenan et al. 2012) provide a
47 statistical framework for jointly modeling count data and demographic data, typically resulting in improved inferences
48 about the mechanisms regulating population dynamics. As a result, there has been continued development of more
49 general ‘integrated data models’ that seek to combine any independent data sources that arise from the same
50 ecological process (Gimenez et al. 2014). For example, occupancy and abundance are two directly related ecological
51 state variables and joint analysis of capture-recapture and occupancy data has been shown to improve estimates of
52 abundance (Conroy et al. 2008; Blanc et al. 2014), density (Chandler and Clark 2014), and even
53 colonization-extinction dynamics and dispersal (Sutherland et al. 2014). A common feature of the majority of studies

* E-mail: simone.tenan@muse.it

† These authors equally contributed to this work.

54 that use multiple data sources, aside from improving parameter precision, is that each independent data set is
55 collected according to a well-defined sampling protocol, i.e., it is an integration of structured data sets. The value of
56 unstructured or opportunistic data, such as that collected by many citizen scientists, is yet unclear. For example, [van
57 Strien et al. \(2013\)](#) argue that opportunistic data represents an important data source that, if analyzed appropriately,
58 can yield improved inferences about temporal trends in occurrence, while [Kamp et al. \(2016\)](#) caution against its use,
59 demonstrating that citizen science data were unable to detect significant species declines. Regardless, with the rapid
60 increase in citizen science initiatives, finding innovative ways to utilize opportunistic data will broaden the scope of
61 ecological enquiry that can be addressed within a single analytical framework ([Gimenez et al. 2014](#)).

62 Spatial capture-recapture methods (SCR: [Efford 2004](#); [Royle et al. 2014](#)) are now well-established in applied
63 ecology and produce estimates of population density using spatial encounter history data, i.e., observation data on
64 who was detected when, and importantly, where. Using spatial patterns of observations to account for heterogeneity
65 in detection probabilities caused by individual differences in trap exposure, and treating space as an explicit model
66 component, SCR produces unbiased estimates of density and space use across a range of conditions (e.g., [Borchers
67 and Efford 2008](#); [Royle et al. 2014](#); [Sollmann et al. 2012](#); [Sutherland et al. 2015](#)). Moreover, SCR has been used to
68 estimate density for elusive species from data collected using a variety of field methodologies including camera traps
69 ([Royle et al. 2009](#)), hair snares ([Gardner et al. 2010](#)), scat surveys ([Fuller et al. 2016](#)), and unstructured survey data
70 ([Kéry et al. 2011](#)). A core component of SCR is an explicit model for space use that relates encounter probability to
71 the distance between any location and an individual's activity center via the estimation of a spatial scale parameter σ
72 (Ch. 7 [Royle et al. 2014](#)). Estimating σ accounts for individual encounter heterogeneity so the effective sampling area
73 is explicitly described and as a result, absolute density can be directly estimated. It follows that to estimate density
74 well, σ must also be well estimated. As with other statistical methods, the precision of SCR-derived estimates of
75 space use and density depend on sample sizes, specifically, but not solely, the number of unique spatial locations
76 individuals are observed at (spatial recaptures). Thus, adding additional spatial information should, in theory, lead to
77 improved inference about space use, and in turn, density. For example, [Gopaldaswamy et al. \(2012\)](#) increased the
78 number of spatial recaptures by integrating camera trap and scat collection data which resulted in more precise
79 estimates of density, while [Royle et al. \(2013\)](#) and [Sollmann et al. \(2013c\)](#) demonstrated that space use (σ) and
80 density are estimated with higher precision when telemetry data are used in addition to traditional capture-recapture
81 data (See also Table 1).

82 Interestingly, in an SCR model, telemetry data require no information about sampling effort because observed

83 locations provide representative information only about the spatial scale parameter (σ), and thus any amount of
84 telemetry data are likely to be informative about space use. So, while the inability to quantify observer effort and bias
85 is often cited as a major limitation of data collected by citizen scientists (Dickinson et al. 2010), it appears that such
86 opportunistic data lends itself naturally to integration within the SCR modeling framework. Specifically, when
87 opportunistic observations can be made of individuals, i.e., via direct or indirect recognition of naturally marked,
88 collared or tagged individuals or the collection of DNA yielding biological samples such as hair or faeces, the
89 locations of those observations are informative about space use and therefore have the potential to improve estimates
90 of spatial parameters in SCR, and have the added benefit of potentially increase the geographic extent of monitoring
91 studies significantly.

92 Here we demonstrate how opportunistic data can be jointly analyzed using spatial capture-recapture methods to
93 improve estimates of density and space use for a reintroduced brown bear *Ursus arctos* population in the central Alps.
94 The area is one of the most populated regions to be occupied by brown bears (Chapron et al. 2014; De Barba et al.
95 2010b) meaning bear-human interactions are highly probable and any perceived threat is considered a key factor in
96 determining the success or failure of the reintroduction (Mustoni et al. 2003). It is important that any conservation
97 management decisions are based on the best available information, and every effort must be made to improve
98 estimates of bear density and space use.

99 MATERIALS AND METHODS

100 **Study area and population**

101 This study was conducted in 2013 in the Italian Alps, an area characterized by a mosaic of natural and
102 human-modified habitats, with a landscape fragmented by urban areas and roads. Elevation ranges from 65 m to more
103 than 3900 m a.s.l., with submontane, montane and subalpine vegetation covering areas below 2000 m, and human
104 population density concentrated below 1000 m (Mustoni et al. 2003). Between 1999 and 2002, nine bears (three
105 males and six females, 3-6 years old) were released in the area as part of a reintroduction project to establish a
106 self-sustained population (Dupré et al. 2000; Mustoni et al. 2003). At the time, the original brown bear population
107 consisted of at least three animals, which were assumed to have died without any genetic exchange with the
108 translocated bears and their progeny (De Barba et al. 2010a).

109 **Brown bear data**

110 *Non-invasive genetic sampling*

111 Bear hair samples were collected from 99 hair traps and 89 rub trees. Hair traps consisted of a strand of barbed wire
112 wound around trees at *c.* 50 cm above ground level enclosing an area of *c.* 25 m² and scent lure was placed in the
113 center (Woods et al. 1999). They were set from 15 May to 31 July, checked on five occasions and the number of days
114 between occasions ranged from 3 to 10 days (Fig. 1). Rub trees, barbed wire wrapped around trees, were monitored
115 during the same time period and were checked twice, first after six days and then after four days (Fig. 1). All hairs
116 were collected during each visit to hair traps and rub trees so that only newly deposited hairs were collected in
117 subsequent visits. Because the hair trap and rub tree data were collected according to a specific protocol, we refer to
118 this structured data as traditional SCR data, or simply ‘SCR data’. In addition to the structured data collection, we
119 also collected opportunistic hair and feces data (De Barba et al. 2010a,b; Groff et al. 2014; Tenan et al. 2016).
120 Following notification by third parties (typically members of the public), opportunistic sampling of hair and feces was
121 carried out by agency personnel at sites where bear damage occurred, e.g. depredation on livestock, beehives and/or
122 crops (Tenan et al. 2016). We refer to this data as ‘opportunistic data’.

123 Biological samples were genetically analyzed for individual identification using ten loci. For a detailed
124 description of DNA extraction methods, PCR protocols, protocols for individual identification, and molecular sexing,
125 see De Barba et al. (2010a,b). We considered only data belonging to the non-cub part of the population and
126 successfully identified a total of $n = 22$ individuals (12 females and 10 males). Of the 22 individuals, 19 were
127 detected using hair traps; two males and one female that were sampled only on rub trees. During the period of trap
128 deployment, 11 of the 22 individuals (four females and seven males) were detected opportunistically resulting in an
129 additional 30 unique spatial locations (Fig. 1, Fig. 2a, Fig. S1 in Appendix S1).

130 *Telemetry*

131 Two bears, a 5-year old male and a 15-year-old female, were monitored during the hair trap and rub tree sampling
132 period using Global Positioning System (GPS) collars (Vectronic GPS-GSM collars, Vectronic Aerospace GmbH,
133 Berlin, Germany). Collared bears were captured using culvert traps (female) and Aldrich snares (male) upon
134 approved capture protocols (2003-DPR 357/97, Groff et al. 2014). GPS collars collected positions at different
135 intervals ranging from 10 min to 1 h. For the analysis we selected one random record per day per individual, giving a

136 total of 143 unique telemetry locations (74 and 69 for the male and female respectively). The collared female was
137 detected at a hair trap but never detected at rub trees or opportunistically (Fig. 2b). The collared male was never
138 detected with hair traps, but was detected once at a rub tree and opportunistically observed in five occasions (Fig. 2c).

139 Data analysis

140 Spatial capture-recapture data

141 Spatial capture recapture models are hierarchical models (Royle and Dorazio 2008) that describe distance-dependent
142 encounter probabilities (the observation process), and the spatial distribution of individuals across the landscape
143 (density, the ecological state process). We adopt a Bayesian analysis of the model (Royle and Young 2008; Gardner
144 et al. 2010) and assume that individual encounter data, y_{ijk} , representing whether or not individual i was detected in
145 trap j in occasion k , are Bernoulli random variables with success probability p_{ijk} , i.e., the encounter probability:

$$y_{ijk} \sim \text{Bernoulli}(p_{ijk}). \quad (1)$$

146 Encounter probabilities in SCR are assumed to decline with distance between a trap (x) and an individuals activity
147 center (s) according to some decreasing function; here we use the commonly applied half-normal encounter model
148 and allow for sex-specific variation in the parameters:

$$p_{ijk} = p_{0,ijk} \exp\left(-\frac{1}{2\sigma_{\text{sex}}^2} d(x_j, s_i)^2\right), \quad (2)$$

149 where σ_{sex} is the sex-specific spatial scale parameter that determines the decrease in encounter probability as the
150 distance between trap j and individual i 's activity center ($d(x_j, s_i)$), increases. The parameter $p_{0,ijk}$ is the baseline
151 encounter probability and can itself be modelled as a function of individual- (i), trap- (j) and occasion- (k) specific
152 covariates. Specifically, we modelled the baseline encounter probability as a function of *sex*, trap type (*trap*: hair trap
153 or rub tree), and, to account for the different time elapsed between consecutive sample occasions in each trap, time
154 since last check (*time*) using standard logistic regression:

$$\text{logit}(p_{0,ijk}) = \gamma_0 + \gamma_1 \text{sex}_i + \gamma_2 \text{trap}_j + \gamma_3 \text{time}_{jk}. \quad (3)$$

155 The second component of the SCR model is a point process model that describes the distribution of individual
156 activity centers, s_i , within a defined state-space \mathcal{S} which should be large enough to contain all plausible activity

157 centers of all observed individuals (Royle et al. 2014). We were particularly interested in modelling density as a
158 function of spatially varying covariates, i.e., an inhomogeneous point process model, and so we used a discrete
159 representation of S defined as the center points of each pixel. In our case, where we are considering a population that
160 was established from a single release point, we modelled variation in density as a function of the distance from the
161 point that the founding population was released between 1999 and 2002 (*d.release*; see ‘Study area and population’).
162 Using a binomial point process model, the per-pixel intensity, $\mu(s)$, is modelled as a log-linear function of ‘d.release’:

$$\log(\mu(s_g)) = \beta_0 + \beta_1 \text{d.release}_g, \quad (4)$$

163 and the probability that an individual activity center is located in a pixel, $\pi(s)$ is given by:

$$\pi(s_g) = \frac{\mu(s_g)}{\sum \mu(s_g)}. \quad (5)$$

164 This is the standard formulation of a Bayesian SCR model with a sex-specific half-normal encounter probability
165 model, an inhomogeneous point process density model, and the estimation of sex-specific total population size N
166 using data augmentation (see Chapters 7 and 10 in Royle et al. 2014). Sex was known for all observed individuals but
167 not for unobserved (augmented) individuals so was modelled as an individual random effect to be estimated:
168 $\text{sex}_i \sim \text{Bern}(\omega_{sex})$, where ω_{sex} is the population-level sex ratio. We expected detectability to vary significantly
169 between sexes and trap type, and to be positively related to the time since last check, expected space use to vary by
170 sex, and thus that sex-specific σ values would differ, and finally we expected density to decline with distance from the
171 release point.

172 The spatial encounter histories for the standard SCR analysis were generated for detections from a $J = 188$ -trap
173 array consisting of hair traps and rub trees across $K = 5$ sampling occasions (Fig. 1 and Fig. 3). Data were formatted
174 in a 3-dimensional $M \times J \times K$ array, \mathbf{Y}_{SCR} , where M is the number of augmented ‘all-zero’ encounter histories, a
175 proportion of which are the estimated unobserved individuals. The additional data required to fit the model are: the
176 coordinates of each hair trap and tree rub, a vector of sex determination of each individual, a $J \times K$ trap operation
177 matrix which is a binary indicator denoting whether each trap was operational during sampling occasion, and the
178 $J \times K$ matrix of ‘time since last check’ covariates, which were scaled to have zero mean and unit variance (Fig. 3).

179 *Telemetry and opportunistic data*

180 Unlike the traditional SCR data described above, both telemetry locations, \mathbf{I}_{tel} and opportunistic data \mathbf{I}_{opp} are not
181 restricted to trap locations and therefore provide important additional information about individual movement, i.e.,
182 they are both direct observations of space use (Royle et al. 2013). We combine the individual telemetry and
183 opportunistic locations and refer to them collectively as \mathbf{I}_i for individual $i = 1, \dots, n$. These additional locations can
184 be modelled using a bivariate normal movement model (Sollmann et al. 2013b,a):

$$\mathbf{I}_i \sim BVN(s_i, \Sigma) \quad (6)$$

where,

$$\Sigma = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}. \quad (7)$$

185 The parameters of this bivariate normal model can be related directly to the SCR half-normal encounter probability
186 model (Eq. 2) through the shared parameters s and σ , which means that telemetry data, opportunistic data and
187 traditional SCR data can be jointly modelled, each contributing to the estimation of the latent activity center, s_i
188 location and the spatial scale parameter σ .

189 The telemetry and opportunistic data, for R_{tel} and R_{opp} locations, were formatted in two $R \times n \times K$ arrays each,
190 one containing x-coordinates and the other containing corresponding y-coordinates, for the n observed individuals
191 and K occasions. This array structure allows the unstructured data (telemetry and opportunistic) to be related to the
192 SCR data in the integrated model (Fig. 3). In addition, an $M \times K$ matrix denotes the number of unstructured locations
193 for each individual in each occasion (Fig. 3).

194 *Model comparisons*

195 The main objective of this work was to investigate the value of integrating telemetry data and opportunistic sightings
196 data into a traditional SCR model for estimating density and space use for the reintroduced brown bear population to
197 inform future management. To do so we fitted the SCR models described above using four SCR data sets:

198 1. Traditional SCR data only (data from hair traps and tree rubs)

199 2. SCR data + telemetry data

200 3. SCR data + opportunistic sightings

201 4. SCR data + telemetry + opportunistic sightings

202 For each model, we estimated sex-specific total population size, N_{sex} and sex-specific σ_{sex} , and compared point
203 estimates (posterior median) and precision (95% Bayesian Credible Interval width, BCI from here) of the estimated
204 parameters.

205 We adopted a Bayesian analysis of the SCR models using Markov chain Monte Carlo implemented using the
206 program JAGS (Plummer 2003) executed from R (R Core Team 2012). For the parameters of the linear predictors (γ
207 and β), we used an uninformative Normal(0, 100) prior. We used an equally uninformative Normal(0, 10) prior for
208 $\alpha_{1,\text{sex}}$, where $\alpha_{1,\text{sex}} = 1/(2\sigma_{\text{sex}}^2)$. After testing a range of resolution values for the state space, we used a resolution of
209 $4 \times 4 \text{ km}$, a value that was small enough to yield stable parameter estimates, and large enough to ensure the model was
210 computationally tractable. To ensure the state space was large enough to contain all plausible activity centers, we used
211 a 21 km buffer around the most extreme coordinates of all the data (telemetry, opportunistic and trapping data, Fig.
212 2a). Data were augmented with $M - n$ ‘all-zero’ encounter histories, where $M = 300$. Summaries of the posterior
213 distribution were calculated from 30,000 post-burn-in posterior samples (burn-in = 3,000 iterations). The code for
214 the fully-integrated model is available as supporting information.

215 RESULTS

216 Overall, integrating all available sources of information (traditional SCR, telemetry and opportunistic data) produced
217 more precise estimates of population size and spatial scale when compared to the use of either SCR data alone or
218 integrated with only one additional source of data (Table 2-3, Fig. 4). In particular, the gain in precision achieved by
219 jointly modeling all three data types was particularly relevant for sex-specific population size estimates (N_m and N_f).
220 In addition to the increase in precision, integrating additional sources of information resulted in a shift in the median
221 abundance point estimates: from 15.679 (BCI: 7.994 – 30.349) under the SCR-only model to 12.629 (BCI: 6.840 –
222 21.953) under the fully integrated model for the number of males, and from 58.008 (BCI: 24.285 – 147.788) to
223 24.578 (BCI: 13.275 – 51.722) for females. Precision gains in estimates of σ were minimal when adding
224 opportunistic data and highest when integrating telemetry information only, markedly so for males (Table 2-3, Fig. 4).
225 As with the estimates of bear population size, the integration of additional information led to a change in the point

226 estimates of σ ; compared to the SCR-only model, there was a noticeable increase in the scale of space use when any
227 of the additional data was used. The female 95% home range size estimated from the half normal encounter model
228 under the SCR-only model was 212 km^2 whereas for the fully integrated model, the estimate was 1375 km^2 .
229 Conversely, estimates of male space use, σ_m , were consistent across models (Fig. 4, Table 2), as were the
230 corresponding 95% home range size estimates: 1800 km^2 (SCR data only) and 2024 km^2 (fully integrated model).

231 Estimates of the parameter relating density to distance to the reintroduction point (β_1) were negative under all
232 models, and although there was some variation in the strength of the effect, this result supports the hypothesis that
233 density decreased with distance from the point where founders were released (Table 2). The estimated sex ratio in the
234 population, ω_{sex} , did not vary significantly between the four models based on 95% Bayesian Credible Intervals and
235 ranged from 0.22 (BCI: 0.09 – 0.41) in the fully integrated model, to 0.36 (BCI: 0.19 – 0.57) in the SCR + telemetry
236 model (Table 2). Across all models, detectability was higher for males, higher at hair traps, and increased with
237 increasing time between checks (Table 2).

238 DISCUSSION

239 We developed a formulation of a spatial capture-recapture model that integrates multiple data sources, and as a result,
240 improves inferences about key ecological parameters, namely density and space use, which we demonstrate using data
241 from a reintroduced population of brown bears in the Italian Alps. Specifically, we were able to jointly analyze
242 traditional SCR data resulting from systematic sampling of hair traps and rub trees, satellite telemetry data, and
243 opportunistic recovery of biological samples. Comparing estimates from models ranging from traditional SCR, to a
244 fully integrated SCR model that includes both telemetry data and opportunistic sightings data, we demonstrated that
245 the addition of unstructured data results in increased precision in estimates of population size and space usage for this
246 species of conservation interest.

247 Estimates of male population size were stable across all models and precision was highest for the fully integrated
248 model, i.e., the model with most spatial data, and lowest for the SCR-only model, which had the fewest spatial
249 locations (Fig. 4). For females, estimates of population size from models with additional data were comparable but
250 were different from the SCR-only model estimates, both in terms of precision and point estimates. The addition of
251 telemetry and opportunistic data reduced parameter uncertainty when compared to the SCR-only model but overall
252 precision was lower for the females. Although the number of individuals observed for the two sexes was similar (12

253 females and 10 males), the number of SCR encounters at different sites (hair traps and rub trees) was smaller for
254 females (mean 2.1, min 1, max 4) than males (mean 8.6, min 1, max 21) (Fig. S1 in Appendix S1). This suggests that
255 smaller SCR data sets might benefit most from an integrated data approach, or conversely may be less stable and more
256 sensitive to combining data. It is encouraging however, that population size estimates based on the two independent
257 unstructured data sources produce similar estimates of abundance, suggesting the addition is beneficial rather than
258 due to parameter sensitivity. The ability to model sex-specific effects on detectability, which in turn affect estimates
259 of sex-specific abundance, can often be limited by insufficient observations of one sex or the other (Sollmann et al.
260 2011; Tobler and Powell 2013). In cases like these, including in our study, even small amounts of unstructured data,
261 like opportunistic sightings data, can resolve such limitations and increase the value of small SCR data sets.

262 As with abundance, estimates of sex-specific σ were more precise when more data were used, such that σ under
263 the fully integrated model had higher precision compared to the SCR-only model (Fig. 4). The most notable difference
264 was between point estimates of female space use from the SCR-only model and the three comparable data integration
265 models. The change in the point estimates is due to the marked difference in the spatial distribution of telemetry
266 locations compared to the SCR data (hair traps and tree rubs) which is likely related to the link between detector type
267 and behavior. SCR data was represented by a few spatially clustered encounters which may mostly reflect female hair
268 deposition patterns related to territoriality potentially at the core of a home range, telemetry and opportunistic data
269 reflect overall space use (Fig. 2b). Estimates of abundance are explicitly linked to estimates of space use because σ
270 controls overall expected encounter probabilities. Here, with the integration of added spatial information, estimates of
271 σ are higher compared to SCR-only data, and as a result, estimates of abundance are reduced (Fig. 4). Again, this
272 points to the value of added spatial information to refine estimates of space use, and in turn, of abundance and density.

273 The integration of additional spatially-referenced information with traditional SCR data ameliorate inference
274 accuracy, and even a small amount of opportunistic data yields improvements, meaning that citizen science type data
275 are potentially high value data sets in SCR regardless of whether survey effort is known. In some cases, like for the
276 male movement parameter σ_m , incorporation of opportunistic information may counterbalance the possible
277 inconsistency between spatial information provided by telemetry and SCR data (Fig. 4). This is particularly relevant
278 in studies of rare or elusive species, where the amount of SCR encounters is usually small but also budget restrictions
279 and the difficulty in collaring animals limit the number of individuals for which telemetry information is available.
280 Opportunistic data provide additional information not limited to the extent of a trap array, like telemetry, but which
281 can also be available for larger number of individuals than those equipped with devices. Unsurprisingly, the addition

282 of telemetry data alone generally results in more precision gains than adding opportunistic data alone because the
283 former is more information rich (Fig. 4). However, the addition relatively few opportunistic locations acted to
284 increase precision relative to inference from the SCR-only model (Fig. 4), suggesting that even in the absence of
285 telemetry data, opportunistic data are important sources of spatial information.

286 When the number of collared individuals is very low (e.g. 3 individuals in [Royle et al. 2013](#) and [Sollmann et al.](#)
287 [2013a](#), or two in our study) telemetry information may be more or less representative of population space use, with a
288 variable degree of concordance between spatial information provided by telemetry and SCR data. On the other hand,
289 information mismatch can also arise in the presence of sparse SCR data as a consequence of inappropriate trap
290 spacing, variation in group-specific (e.g. age or sex) exposure to traps, or when the trap array is small relative to
291 individual movement, inducing a geographic bias for the most mobile component of the population. Opportunistic
292 information may also not be representative of the entire population, as in the case of records collected at sites where a
293 damage occurred (e.g. depredation on livestock, beehives and crops) where individuals more prone to commit damage
294 can be sampled more times than others. Finally, as suggested above, the encounter data may reflect altogether different
295 behavioral states depending on the particular method used to detect individuals. This suggests that the process of data
296 integration requires more than simply the development of integrated data models, but rather that those models take
297 into consideration the variation in behavior that might be reflected in the independent data sets being used.

298 We provide evidence that incorporating unstructured opportunistic data to SCR and telemetry information, by
299 conceptually treating opportunistic records as thinned telemetry data, improves inference on abundance and space
300 usage, which are key population-level parameters to inform conservation decisions of elusive and difficult-to-study
301 species. However, care must be taken to assess the potential mismatch in spatial information provided by the different
302 data sets, where telemetry is both the most informative source of space use but also often available only for a few
303 individuals, whose movement may not be representative of population space use. Understanding how animal density
304 changes in space and how the latter is used is crucial when addressing practical issues in population management and
305 conservation ([Bischof et al. 2009](#)). For this aim, the use of opportunistic information increases availability of
306 spatially-referenced individual information, that can be suitably modelled along with other data within a unified
307 framework, thus reducing the need for additional invasive methods.

308 ACKNOWLEDGEMENTS

309 This research was partially funded by the Autonomous Province of Trento and the MUSE - Museo delle Scienze. We
310 would like to thank the Autonomous Province of Bolzano, the ISPRA, the personnel of the Servizio Foreste e Fauna
311 of the Autonomous Province of Trento, of the Adamello Brenta Natural Park, and of the Stelvio National Park. We
312 also thank the many forestry wardens and volunteers for field support, and Aaron Iemma for IT assistance.

313 REFERENCES

- 314 Besbeas, P., S. N. Freeman, B. J. T. Morgan, and E. A. Catchpole, 2002. Integrating mark-recapture-recovery and
315 census data to estimate animal abundance and demographic parameters. *Biometrics* **58**:540–547.
- 316 Bischof, R., J. E. Swenson, N. G. Yoccoz, A. Mysterud, and O. Gimenez, 2009. The magnitude and selectivity of
317 natural and multiple anthropogenic mortality causes in hunted brown bears. *Journal of Animal Ecology*
318 **78**:656–665.
- 319 Blanc, L., E. Marboutin, S. Gatti, F. Zimmermann, and O. Gimenez, 2014. Improving abundance estimation by
320 combining capture–recapture and occupancy data: example with a large carnivore. *Journal of Applied Ecology*
321 **51**:1733–1739.
- 322 Borchers, D. L. and M. Efford, 2008. Spatially explicit maximum likelihood methods for capture–recapture studies.
323 *Biometrics* **64**:377–385.
- 324 Chandler, R. B. and J. D. Clark, 2014. Spatially explicit integrated population models. *Methods in Ecology and*
325 *Evolution* **5**:1351–1360.
- 326 Chapron, G., P. Kaczensky, J. D. Linnell, M. Von Arx, D. Huber, H. Andrén, J. V. López-Bao, M. Adamec,
327 F. Álvares, O. Anders, et al., 2014. Recovery of large carnivores in Europe’s modern human-dominated
328 landscapes. *Science* **346**:1517–1519.
- 329 Conroy, M. J., J. P. Runge, R. J. Barker, M. R. Schofield, and C. J. Fonnesebeck, 2008. Efficient estimation of
330 abundance for patchily distributed populations via two-phase, adaptive sampling. *Ecology* **89**:3362–3370.

- 331 De Barba, M., L. Waits, E. Garton, P. Genovesi, E. Randi, A. Mustoni, and C. Groff, 2010a. The power of genetic
332 monitoring for studying demography, ecology and genetics of a reintroduced brown bear population. *Molecular*
333 *Ecology* **19**:3938–3951.
- 334 De Barba, M., L. P. Waits, P. Genovesi, E. Randi, R. Chirichella, and E. Cetto, 2010b. Comparing opportunistic and
335 systematic sampling methods for non-invasive genetic monitoring of a small translocated brown bear population.
336 *Journal of Applied Ecology* **47**:172–181.
- 337 Dickinson, J. L., J. Shirk, D. Bonter, R. Bonney, R. L. Crain, J. Martin, T. Phillips, and K. Purcell, 2012. The current
338 state of citizen science as a tool for ecological research and public engagement. *Frontiers in Ecology and the*
339 *Environment* **10**:291–297.
- 340 Dickinson, J. L., B. Zuckerberg, and D. N. Bonter, 2010. Citizen science as an ecological research tool: challenges
341 and benefits. *Annual review of ecology, evolution and systematics* **41**:149–72.
- 342 Dupré, E., P. Genovesi, and L. Pedrotti, 2000. Studio di fattibilità per la reintroduzione dell’Orso bruno (*Ursus*
343 *arctos*) sulle Alpi centrali. Istituto nazionale per la fauna selvatica ‘Alessandro Ghigi’.
- 344 Efford, M., 2004. Density estimation in live-trapping studies. *Oikos* **106**:598–610.
- 345 Fuller, A. K., C. S. Sutherland, J. A. Royle, and M. P. Hare, 2016. Estimating population density and connectivity of
346 american mink using spatial capture–recapture. *Ecological Applications* **26**:1125–1135.
- 347 Gardner, B., J. A. Royle, M. T. Wegan, R. E. Rainbolt, and P. D. Curtis, 2010. Estimating black bear density using
348 DNA data from hair snares. *The Journal of Wildlife Management* **74**:318–325.
- 349 Gimenez, O., S. T. Buckland, B. J. Morgan, N. Bez, S. Bertrand, R. Choquet, S. Dray, M.-P. Etienne, R. Fewster,
350 F. Gosselin, et al., 2014. Statistical ecology comes of age. *Biology letters* **10**:20140698.
- 351 Gopaldaswamy, A. M., J. A. Royle, M. Delampady, J. D. Nichols, K. U. Karanth, and D. W. Macdonald, 2012. Density
352 estimation in tiger populations: combining information for strong inference. *Ecology* **93**:1741–1751.
- 353 Groff, C., N. Bragalanti, R. Rizzoli, and P. Zanghellini, 2014. 2013 Bear report. Technical report, Forestry and
354 Wildlife Department of the Autonomous Province of Trento.

- 355 Kamp, J., S. Oppel, H. Heldbjerg, T. Nyegaard, and P. F. Donald, 2016. Unstructured citizen science data fail to
356 detect long-term population declines of common birds in denmark. *Diversity and Distributions* **Early View**, DOI:
357 **10.1111/ddi.12463**.
- 358 Kéry, M., B. Gardner, T. Stoeckle, D. Weber, and J. A. Royle, 2011. Use of spatial capture-recapture modeling and
359 DNA data to estimate densities of elusive animals. *Conservation Biology* **25**:356–364.
- 360 Mustoni, A., E. Carlini, B. Chiarenzi, S. Chiozzini, E. Lattuada, E. Dupré, P. Genovesi, L. Pedrotti, A. Martinoli,
361 D. Preatoni, et al., 2003. Planning the Brown Bear *Ursus arctos* reintroduction in the Adamello Brenta Natural
362 Park. a tool to establish a metapopulation in the Central-Eastern Alps. *Hystrix* **14**:3–27.
- 363 Newman, G., A. Wiggins, A. Crall, E. Graham, S. Newman, and K. Crowston, 2012. The future of citizen science:
364 emerging technologies and shifting paradigms. *Frontiers in Ecology and the Environment* **10**:298–304.
- 365 Plummer, M., 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In
366 *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*.
- 367 R Core Team, 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical
368 Computing, Vienna, Austria. ISBN 3-900051-07-0.
- 369 Royle, J. A., R. B. Chandler, R. Sollmann, and B. Gardner, 2014. Spatial Capture-Recapture. Academic Press,
370 Waltham, MA.
- 371 Royle, J. A., R. B. Chandler, C. C. Sun, and A. K. Fuller, 2013. Integrating resource selection information with spatial
372 capture–recapture. *Methods in Ecology and Evolution* **4**:520–530.
- 373 Royle, J. A. and R. Dorazio, 2008. Hierarchical modeling and inference in ecology: the analysis of data from
374 populations, metapopulations and communities. Academic Press, San Diego.
- 375 Royle, J. A., K. U. Karanth, A. M. Gopalaswamy, and N. S. Kumar, 2009. Bayesian inference in camera trapping
376 studies for a class of spatial capture-recapture models. *Ecology* **90**:3233–3244.
- 377 Royle, J. A. and K. V. Young, 2008. A hierarchical model for spatial capture-recapture data. *Ecology* **89**:2281–2289.

- 378 Schaub, M. and F. Abadi, 2010. Integrated population models: a novel analysis framework for deeper insights into
379 population dynamics. *Journal of Ornithology* **152**:227–237.
- 380 Sollmann, R., M. M. Furtado, B. Gardner, H. Hofer, A. T. Jácomo, N. M. Tôrres, and L. Silveira, 2011. Improving
381 density estimates for elusive carnivores: accounting for sex-specific detection and movements using spatial
382 capture–recapture models for jaguars in central brazil. *Biological Conservation* **144**:1017–1024.
- 383 Sollmann, R., B. Gardner, and J. L. Belant, 2012. How does spatial study design influence density estimates from
384 spatial capture-recapture models? *PloS one* **7**:e34575.
- 385 Sollmann, R., B. Gardner, R. B. Chandler, D. B. Shindle, D. P. Onorato, J. A. Royle, and A. F. O’Connell, 2013a.
386 Using multiple data sources provides density estimates for endangered florida panther. *Journal of Applied Ecology*
387 **50**:961–968.
- 388 Sollmann, R., B. Gardner, A. W. Parsons, J. J. Stocking, B. T. McClintock, T. R. Simons, K. H. Pollock, and A. F.
389 O’Connell, 2013b. A spatial mark–resight model augmented with telemetry data. *Ecology* **94**:553–559.
- 390 Sollmann, R., N. M. Tôrres, M. M. Furtado, A. T. de Almeida Jácomo, F. Palomares, S. Roques, and L. Silveira,
391 2013c. Combining camera-trapping and noninvasive genetic data in a spatial capture–recapture framework
392 improves density estimates for the jaguar. *Biological Conservation* **167**:242–247.
- 393 Sutherland, C., D. Elston, and X. Lambin, 2014. A demographic, spatially explicit patch occupancy model of
394 metapopulation dynamics and persistence. *Ecology* **95**:3149–3160.
- 395 Sutherland, C., A. K. Fuller, and J. A. Royle, 2015. Modelling non-euclidean movement and landscape connectivity
396 in highly structured ecological networks. *Methods in Ecology and Evolution* **6**:169–177.
- 397 Tenan, S., J. Adrover, A. M. Navarro, F. Sergio, and G. Tavecchia, 2012. Demographic consequences of
398 poison-related mortality in a threatened bird of prey. *PloS ONE* **7**:e49187.
- 399 Tenan, S., A. Iemma, N. Bragalanti, P. Pedrini, M. Barba, E. Randi, C. Groff, and M. Genovart, 2016. Evaluating
400 mortality rates with a novel integrated framework for non-monogamous species. *Conservation Biology* **Early**
401 **View, DOI: 10.1111/cobi.12736.**

- 402 Tobler, M. W. and G. V. Powell, 2013. Estimating jaguar densities with camera traps: problems with current designs
403 and recommendations for future studies. *Biological Conservation* **159**:109–118.
- 404 van Strien, A. J., C. A. Swaay, and T. Termaat, 2013. Opportunistic citizen science data of animal species produce
405 reliable estimates of distribution trends if analysed with occupancy models. *Journal of Applied Ecology*
406 **50**:1450–1458.
- 407 Williams, B. K., J. D. Nichols, and M. J. Conroy, 2002. Analysis and management of animal populations. Academic
408 Press, San Diego.
- 409 Woods, J. G., D. Paetkau, D. Lewis, B. N. McLellan, M. Proctor, and C. Strobeck, 1999. Genetic tagging of
410 free-ranging black and brown bears. *Wildlife Society Bulletin* **27**:616–627.

411 TABLES

Table 1: Summary of contributions that provide an integrated framework for spatially-referenced individual data. Systematic data are collected under specific study designs: spatial capture-recapture (SCR), telemetry, and counts or binary detections (survey). Parameter shared: ψ , Data Augmentation parameter; σ , scale parameter of the observation model ; ϕ , survival probability; δ , individual-level recruitment probability.

Paper	Systematic			Opportunistic	Parameter	Study species
	SCR	Telemetry	Survey			
Sollmann et al. (2013c)	• ¹				σ	jaguar
Gopaldaswamy et al. (2012)	• ¹				ψ, σ	tiger
Sollmann et al. (2013b)	• ²	•			σ	raccoon
Sollmann et al. (2013a)	• ²	•			σ	Florida panther
Royle et al. (2013)	•	• ³			σ	black bear
Chandler and Clark (2014)	•		•		ϕ, δ	black bear
Present study	•	•		•	σ	brown bear

¹ camera trapping and scat collection;

² extended to mark-resight;

³ resource selection function data

Table 2: Posterior parameter estimates achieved using SCR data alone or integrated with other systematic (telemetry) and opportunistic (opp) information available for the brown bear population in the Italian Alps. Parameters are denoted as follows: number of males and females, N_m and N_f , respectively; scale parameter of the Gaussian kernel, which relates to space usage of females and males, σ_f and σ_m , respectively; baseline encounter probability (p_0) intercept, γ_0 ; effect of being male on p_0 , γ_1 ; effect of trap type on p_0 , γ_2 ; effect of time since last check on p_0 , γ_3 ; density intercept, β_0 ; effect of distance from the release point on density, β_1 ; data augmentation parameter, ψ ; probability of being a male, ω_{sex} . Estimates are reported for a 4-km resolution of the state space grid.

Parameter	mean	SD	2.5%	50%	97.5%
SCR + tel + opp					
N_m	13.088	3.893	6.840	12.629	21.953
N_f	26.587	10.249	13.275	24.578	51.722
σ_m	10.372	0.474	9.496	10.351	11.348
σ_f	8.548	0.481	7.669	8.523	9.558
γ_0	-3.714	0.308	-4.335	-3.705	-3.125
γ_1	1.623	0.312	1.024	1.618	2.246
γ_2	-1.153	0.387	-1.899	-1.155	-0.384
γ_3	0.619	0.185	0.248	0.621	0.976
β_0	-3.245	0.670	-4.719	-3.192	-2.081
β_1	-1.128	0.467	-2.125	-1.107	-0.276
ψ	0.198	0.064	0.109	0.186	0.351
ω_{sex}	0.341	0.096	0.171	0.335	0.542
SCR + opp					
N_m	13.687	4.297	6.965	13.166	23.675
N_f	36.890	18.900	15.703	32.157	89.949
σ_m	11.800	0.905	10.207	11.741	13.743
σ_f	7.092	1.103	5.380	6.936	9.669
γ_0	-3.414	0.345	-4.116	-3.406	-2.758
γ_1	1.417	0.360	0.734	1.409	2.143
γ_2	-1.160	0.392	-1.923	-1.164	-0.390
γ_3	0.621	0.189	0.248	0.623	0.985
β_0	-2.713	0.600	-3.960	-2.687	-1.562
β_1	-0.739	0.435	-1.620	-0.731	0.113
ψ	0.252	0.106	0.124	0.227	0.542
ω_{sex}	0.291	0.096	0.125	0.285	0.498
SCR + tel					
N_m	15.514	5.297	7.747	14.684	28.297
N_f	28.206	11.284	13.501	25.819	57.590
σ_m	8.556	0.425	7.773	8.538	9.441
σ_f	8.339	0.482	7.461	8.316	9.345
γ_0	-3.479	0.332	-4.139	-3.478	-2.833
γ_1	1.791	0.343	1.119	1.787	2.472
γ_2	-1.207	0.402	-1.991	-1.207	-0.413
γ_3	0.670	0.193	0.288	0.671	1.046
β_0	-2.956	0.627	-4.313	-2.899	-1.869
β_1	-0.890	0.460	-1.842	-0.866	-0.043
ψ	0.218	0.074	0.115	0.204	0.406
ω_{sex}	0.364	0.098	0.187	0.359	0.569
SCR					
N_m	16.574	5.766	7.994	15.679	30.349
N_f	65.057	31.086	24.285	58.008	147.788
σ_f	3.355	0.582	2.460	3.274	4.729
σ_m	9.796	1.277	7.739	9.644	12.756
γ_0	-2.220	0.407	-3.040	-2.213	-1.451
γ_1	0.804	0.441	-0.046	0.799	1.683
γ_2	-1.207	0.421	-2.037	-1.204	-0.389
γ_3	0.658	0.203	0.263	0.656	1.060
β_0	-2.076	0.549	-3.259	-2.032	-1.119
β_1	-0.459	0.387	-1.267	-0.437	0.244
ψ	0.406	0.168	0.177	0.371	0.850
ω_{sex}	0.223	0.084	0.090	0.212	0.414

Table 3: Gain in precision of parameter estimates expressed as per cent reduction in the SD of the posterior estimates achieved by integrating different data types, compared to those obtained using classical SCR models. Population size (N_m and N_f) and the scale parameter of the observation model (σ_m and σ_f) are reported for males and females respectively.

	N_m	N_f	σ_m	σ_f
SCR + telemetry	8%	63%	67%	17%
SCR + opportunistic	26%	39%	29%	-90%
SCR + telemetry + opportunistic	33%	67%	63%	18%

412 FIGURE LEGENDS

413 **Figure 1:** Timeline of data collection during the period when SCR data were systematically collected using an array
414 of hair traps checked on five occasions of variable length (black blocks), and from rub trees checked for hairs in two
415 period (grey blocks). Telemetry data were thinned by randomly selecting on record per day, and opportunistic
416 recovery of biological samples was performed in 23 days.

417 **Figure 2:** (a) Distance from the point were founders were released (in km) and location of bear captures from
418 systematic sampling with hair traps and rub trees (SCR), telemetry and opportunistic records. (b-c) Location of the
419 records for the two collared individuals from which telemetry information was derived. Grey dots indicate the
420 location of all observed individuals.

421 **Figure 3:** Graphical representation of the data involved in the integrated analysis. Circles represent estimated
422 parameters. Observed data for n individuals, detected during K visits and members of the population of size N , were
423 augmented with $M - n$ all-zero detections (Y_{SCR} matrix). SCR data were collected at J sites, consisting of J_h hair
424 traps and J_r rub trees. Data set names in Courier font correspond to the names used in the model code. Coordinates,
425 trap deployment, and (standardized) time since last check data sets for the $J = J_h + J_r$ traps are denoted by
426 `SCR.traps`, `active`, and `time.elapsed.sc` labels, respectively. Raster data contain information on the distance
427 from the point were founders were released for each of the nG pixels (`d2core`) and was used to model density.
428 Telemetry and opportunistic data were formatted in the same way, with an augmented matrix for number of records
429 available for individual i in occasion k (`n.obs.TEL` and `n.obs.OPP`, respectively) and the x (`TEL.y_x`, `OPP.y_x`)
430 and y (`TEL.y_y`, `OPP.y_y`) coordinates of those records for each individual in each of the R_{tel} or R_{opp} locations and
431 occasion k .

432 **Figure 4:** Posterior parameter estimates (mean and 95% Bayesian Credible Interval) achieved using SCR data only,
433 or integrating them with telemetry ('tel') and opportunistic ('opp') data available for the brown bear population in the
434 Italian Alps. Reported are the sex-specific population sizes (N) and scale parameters of the Gaussian kernel (σ).

435 FIGURES

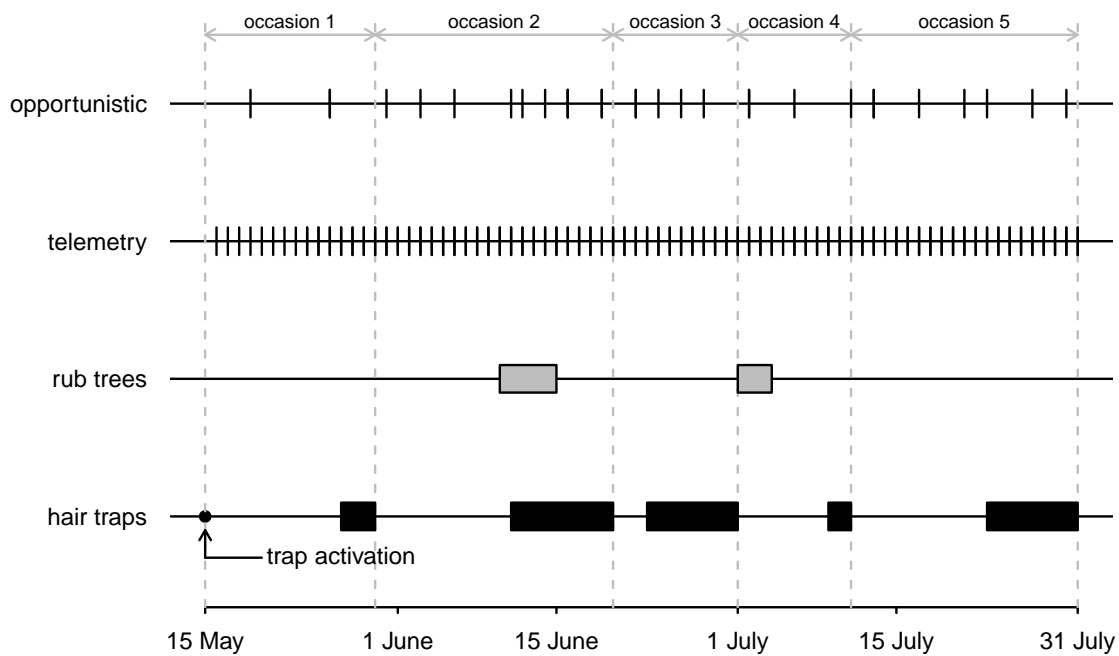


Figure 1

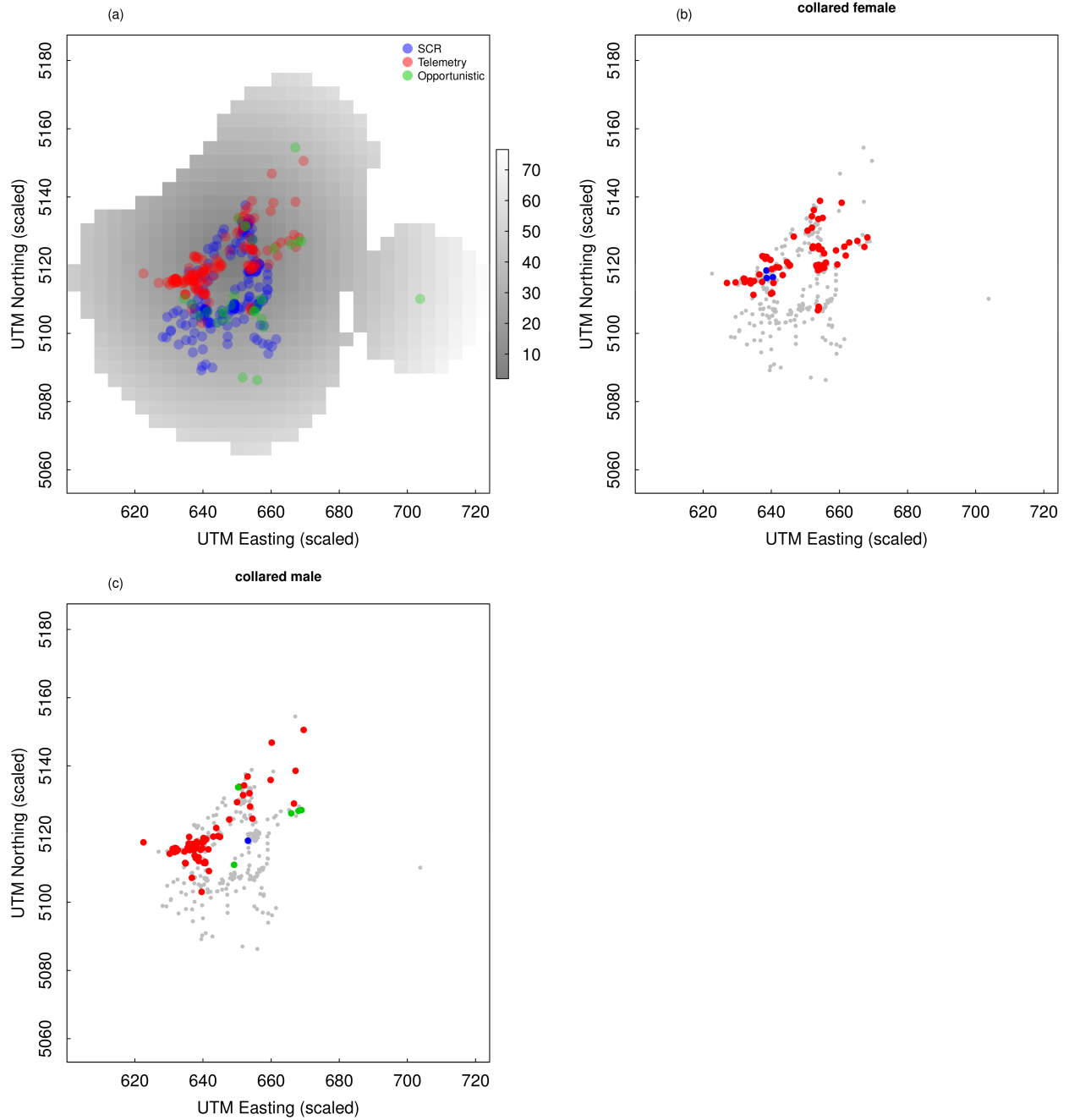


Figure 2

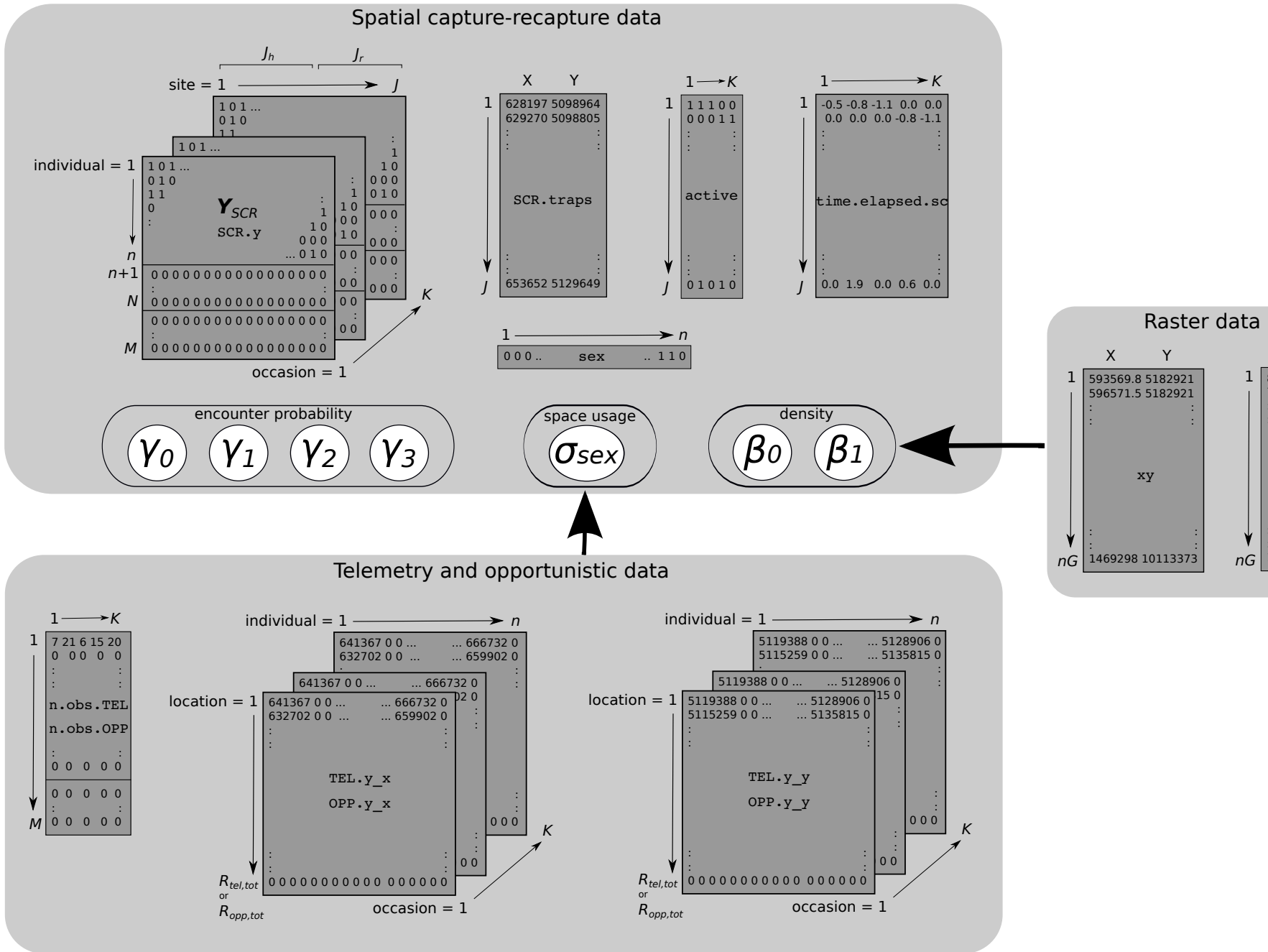


Figure 3

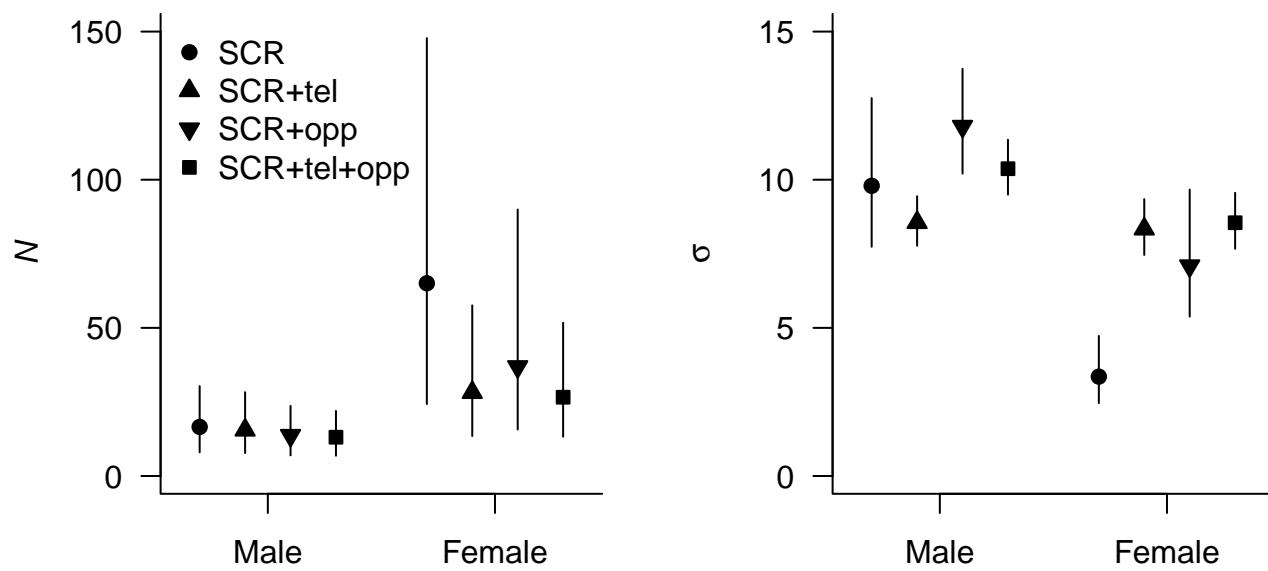


Figure 4