Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
`http://siba-ese.unisalento.it/index.php/ejasa/index`
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v12n3p604

Counting the unmarked: Estimating animal population using count data

By Gubbi et al.

Electronic Journal of Applied Statistical Analysis
Vol. 12, Issue 03, November 2019, 604-35
DOI: 10.1285/i20705948v12n3p604

Counting the unmarked: Estimating animal population using count data

Sanjay Gubbi*[a,b], Srikanth Seshadri[c], Vijaya Kumara[b], and Suresh Chandra K[d]

[a]*Nature Conservation Foundation, 3076/5, IV Cross, Gokulam Park, Mysore - 570 002, India*
[b]*Department of Wildlife and Management, Kuvempu University, Jnanasahyadri, Shankarghatta - 577 451, India*
[c]*Sahaj Software Solutions, 365, 3rd Floor, First Cross, Road, 3rd Block, Koramanagala, Bangalore - 560 034, India*
[d]*Meghana and Shalini Apartments, 100 Feet Ring Road, Bangalore - 560 070, India.*

Published: 20 November 2019

Understanding population parameters are important tools for wildlife management, and one of the key objectives of ecological research. Motion sensor cameras are a widely used tool to estimate abundance and densities of species that are identifiable based on the natural markings on their bodies. Though camera trapping provides information such as count data, on species that are not individually identifiable, estimating population size using conventional capture-recapture methodologies is not possible hindering estimating population information of several wildlife species. However, recent methodologies help use camera trapping data to bridge this gap. Here we extend the model of Chandler and Royle (2013), with suitable modifications, and used camera trap detection data to estimate abundance and density of eight wild prey, and five domestic prey species of leopards (*Panthera pardus fusca*). In this context, a new procedure has been proposed, based on grouping of the count data, which is useful in cases of large encounters.

The current model should apply widely to a range of other unmarked wildlife species such as dholes, lions, golden jackal, Indian fox, ratel, to name a few, that could help understand prey-predator relationships, competition, trophic interactions, species interactions and other similar ecological questions. The methodology could also reduce costs, and maximise the utilisation of existing camera trapping data. The methodology helps understanding

---

[*] Corresponding author: gubbi@ncf-india.org

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

population parameters of several endangered, unmarked species to draw up conservation strategies whose estimates are currently largely based on educational guess.

keywords: Abundance estimation, camera-trapping, count data, density, ungulates, unmarked species.

1 Introduction

Understanding density and abundance are important tools for wildlife management, and one of the key objectives of ecological research. It also acts as an important indicator for evaluating wildlife management.

In the recent years motion sensor cameras are a widely used tool to collect information on populations sizes, distribution, species richness and other useful information especially on cryptic species (Champion, 1992; Griffiths and Van Schaik, 1993; Garshelis et al., 1999). Importantly, it''s commonly used for estimating abundance and densities of naturally marked animals. However, many other wildlife species are also captured on the camera traps providing valuable data but are discarded, in most circumstances, as estimating density or abundance of species that cannot be individually identifiable (unmarked henceforth) is not possible under the current conventional capture-recapture methodologies. This has hindered density estimation of species of several globally significant wildlife species, or species that provide useful information on interrelatedness amongst wildlife species such as prey-predator relationships (Karanth et al., 2004; Carbone et al., 2010), trophic interactions, linkages (Owen-Smith and Mills, 2008), harvest quotas, diesease (Ramsey et al., 2015), or species interactions. Such monitoring is especially important for conservation and monitoring of endangered species (Campbell et al., 2002).

Since abundance and density estimation of unmarked species is not possible through the traditional capture-recapture, or the recently developed spatially-explicit capture-recapture models that use photographic information of individually identifiable animals through the natural markings on their bodies, trapping rate (photographs/trapping effort) is widely used as an alternative measure of relative abundance (Carbone et al., 2001). However trapping rates does not account for imperfect detection hence limiting their utility.

However, challenging conventional analytical sampling methods, Chandler and Royle (2013) demonstrate that individual recognition of species is not obligatory for population density estimation. The study uses spatially correlated count data of a species at multiple sample locations in close proximity to derive density, and abundance estimates for species that do not possess any natural markings. Hence this model used on unmarked species is considered as an extension of the existing spatial capture-recapture models (Ramsey et al., 2015). In addition, the traditional capture-recapture, or spatial capture-recapture methods use Binomial encounter data which is replaced by count data based on Poisson encounters (aggregated over the unknown number of individuals). This has a significant utility for researchers who encounter unmarked species in camera traps to make inferences about animal population sizes.

1.1 *Distance sampling versus spatial count models*

The distance sampling methodology is widely used for the estimation of abundance of unmarked animals. In fact the spatial capture-recapture models can be viewed as extension of distance sampling models (Royle et al., 2013). Further, the fact that the distance sampling methodology does not need the explicit identity of the animal makes it a competitor for the spatial count model discussed in this paper. Both methodologies have their relative merits and demerits. But the distance sampling methods involve relatively more human efforts, and measurements at various stages of data collection that can contribute to non-sampling errors in the final inference. It is highly resource consuming if data has to be collected over a vast area.

We also note that the distance sampling cannot be applied to collect data on some species due to their inherent behaviour such as nocturnal activity, low visual detection rates during data collection, and other similar characteristics. This has hindered population estimates of key wildlife species such as sloth bears (*Melursus ursinus*), dholes (*Cuon alpinus*), and others. These limitations coupled with the requirement of large data for efficient analysis in distance sampling methodology, do enhance the practical utility of using spatial count data.

2 Materials and methods

2.1 *Study area*

The study was carried out in the contiguous forests of Bukkapatna (142.8 km^2), Suvarnamukhi (22.5 km^2), Muthagadahalli (4.4 km^2), Mathikere (2.4 km^2) reserved forests, and other adjoining unprotected forests in southern India Figure 1. The habitat is largely woodland savannah, scrub forests, and supports a wide variety of dry habitat wildlife species including leopard (*Panthera pardus fusca*), striped hyaena (*Hyaena hyaena*), sloth bear (*Melursus ursinus*), golden jackal (*Canis aureus*), and other carnivore species. Herbivore species includes Indian gazelle (*Gazella bennetti*), four-horned antelope (*Tetracerus quadricornis*), blackbuck (*Antilope cervicapra*), wild pig (*Sus scrofa*), and others.

2.2 *Camera trapping*

In this study, we use existing camera trap data of 12 unmarked species to estimate density and abundance of leopard prey species, and for sloth bear by adopting the recently

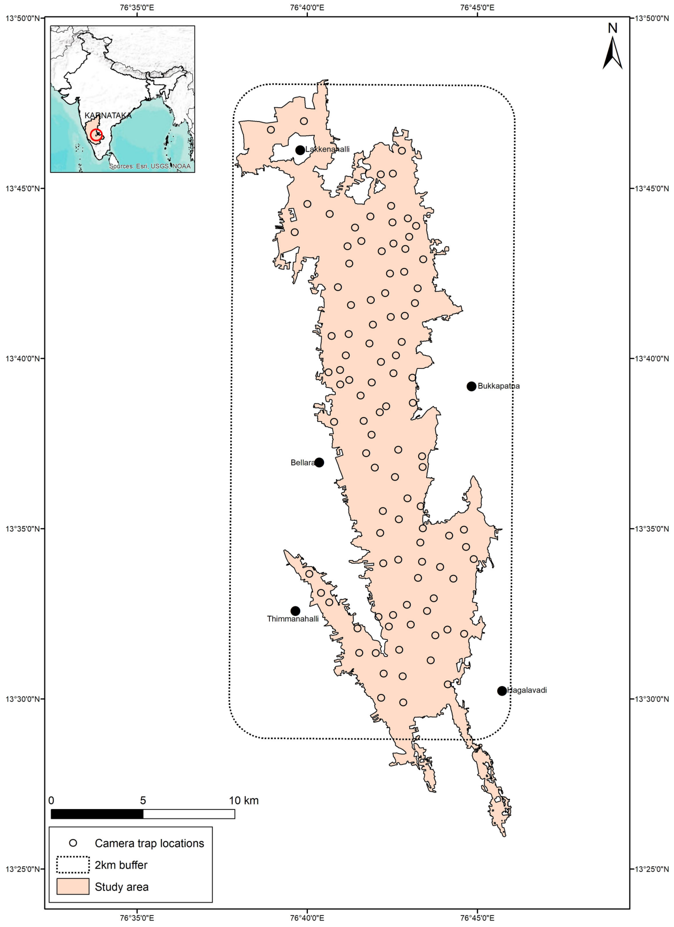

Figure 1: Study area, and locations of camera traps (indicated by open circles) at Bukkapatna and adjoining areas in southern India. The polygon surrounding the camera trap locations defined the area of the state-space used for inference.

developed Chandler and Royle (2013) method with suitable modifications. Data was collected as a by-product while carrying out population estimation studies on leopards using Panthera V4 passive infrared motion detection digital cameras (Panthera Corp. NY, USA).

Camera traps were placed at 99 locations for 16 trap days (24 hour period). All cameras were deployed along forest roads or trails and secured to trees approximately 45 cm above ground. Cameras operated 24 h per day and were checked every day to retrieve data and to ensure functionality of the units. The actual study area was 221.3 km^2, and a two km buffer around the convex hull over trap locations yielded a super area of 525.4 km^2.

2.3 *Sampling design and data*

In this study we assume that there are J traps (camera traps) monitored on K occasions involving N (unknown) individuals. Denoted by z_{ijk} the encounter data on the i^{th} individual by the j^{th} trap on the k^{th} occasion (i=1,2,...,N; j=1,2,...,J; k=1,2,...,K), we note that, especially in the case of unmarked individuals, the individual-specific encounter data is not available. What is observed of, is the count data

$$n_{jk} = \sum_{i=1}^{N} z_{ijk}$$

and that N (which is our primary interest) is unknown. If one assumes that each z_{ijk} is a realization of a Poisson random variable with (rate) parameter λ_{ij} , and the underlying distributions are independent, we have

$$n_{jk} \sim Poisson(\Lambda)$$

where $\Lambda_j = \sum_{i=1}^{N} \lambda_{ij}$. Therefore the analysis based on the count data { n_{jk} } may not directly depend on the non-observable latent data on { z_{ijk} }. But in practical situations it may open out some statistical issues which will be elaborated later in this paper.

Towards making the underlying model behind the count encounter data spatially explicit, we assume that each individual has an activity center s_i (not known apriori) identified uniquely by its (x,y) co-ordinates, in relation to the given study area. Similarly if x_j indicates the location of the j^{th} trap (which is known uniquely through its co-ordinates) then the Euclidean distance between s_i and x_j given by d_{ij} plays a vital role in the most commonly used modeling of the encounter rate given by

$$\lambda_{ij} = \lambda_0 \left\{ \sum_{i=1}^{N} e^{-\frac{d_{ij}^2}{2\sigma^2}} \right\}$$

Since N is unknown, the data augmentation approach Royle et al. (2013) introduces two more quantities: M a conveniently chosen known upper limit to N, and a parameter, $(0 < \psi < 1)$, the data augmentation parameter. In fact, if w_i is the binomial random variable indicating whether the i^{th} individual is the member of the population with parameter , it is easy to see that $N = \sum_{i=1}^{M} w_i$ and that $\psi = E(N)/M$. Thus the final shape of the model is

$$w_i \sim Bernoulli(\psi)$$

$z_{ijk} \sim Poisson(\lambda_{ij} w_i)$; k=1,2,...,K

$$\lambda_{ij} = \lambda_0 \left\{ \sum_{i=1}^{M} e^{-\frac{d_{ij}^2}{2\sigma^2}} \right\}$$

$$n_{jk} \sim \sum_{i=1}^{M} z_{ijk}$$

There are three parameters σ, λ_0 and ψ and two sets of latent (unobservable) random variables z, and w in the model leaving the Bayesian methodology as the only option for the estimation of the parameters.

It may be noted that, unlike in marked data sets, the interpretation of σ and λ_0 is not straightforward. However, Chandler and Royle (2013) highlight that σ is an indicator of the degree of spatial correlation among counts and the information in this spatial correlation facilitates, indirectly, to obtain encounter rate parameters and the density.

2.4 *Data application*

Analyses was carried out on nine of the commonly camera trapped wild prey species that had no natural markings that would help in individual identification. Species included four-horned antelope (*Tetracerus quadricornis*), blackbuck (*Antilope cervicapra*), chinkara (*Gazella bennetti*), wild pig (*Sus scrofa*), Indian crested porcupine (*Hystrix indica*), black-naped hare (*Lepus nigricollis*), bonnet macaque (*Macaca radiata*), Indian peafowl (*Pavo cristatus*), all important wild prey species for leopards. In addition, we considered cow (*Bos taurus*) and buffalo (*Bubalus bubalis*) both combined and termed as large livestock, sheep (*Ovis aries*) and goat (*Capra aegagrus hircus*) combined and called as small livestock, and domestic dog (*Canis lupus familiaris*). Since the area has a population of sloth bears, and sufficient data was available during this study it was included into the analyses.

Considering the mobility of the specific species the buffer area was identified for data analysis. For all species (except for porcupine and blackbuck) a two km buffer was used. For porcupine a buffer of 500 meters was used as we assumed the species has smaller home ranges, and would not use agricultural fields in the study area. Since blackbuck is found within the woodland savannah habitat and extensively uses agricultural fields surrounding the study area, the entire super area was considered as habitat.

Potential activity centers, being centroids of suitably chosen pixels in the super area were identified after dichotomising the entire area into habitat and non-habitat regions. For all species other than domestic dogs, small and large livestock, and blackbuck, the habitat region accounted for 221.3 km^2 mostly consisting of open woodland savannah, and dry deciduous habitats. In case of domestic dogs, small and large livestock which normally spill over to forest area from non-forest areas, necessary changes in the interpretation of their activity centers and densities have been suitably adopted. Since blackbuck are found both within the open woodland savannah and agricultural fields the entire super area was assumed as their habitat.

2.5 *The R code used for the analysis*

We have used the Metropolis-within-Gibbs algorithm developed by Chandler and Royle (2013) in which the algorithm is unconditional on latent encounter frequencies which is the most suitable algorithm for our study data. The original algorithm was modified to facilitate the input data in the csv format, and to include the goodness of fit (Freeman-Tukey type) statistics based both on the "individual x trap frequencies" and on the "trap frequencies" (Royle et al., 2013) for computing the Bayesian p-values. The Bayesian p-values arising out of these statistics are denoted by p_1 and p_2 respectively. Although we have presented both the values, with at least one of them within the interval (0.05, 0.95), the second statistic mentioned above leading to the p_2 value has been claimed to be more appropriate fit statistic in the analysis of spatial count models (Royle et al., 2013). We had to modify the procedure to update the activity centers (S), since there was a necessity to use "habitat mask" for our analyses. The MCMC diagnostics were also incorporated into the code using the package CODA, a library within R. Since the code required a larger computational time, necessary changes had to be incorporated to hasten the computations, wherever possible.

2.6 *Analysis and issues*

The modified algorithm used for the analysis requires two data input files: Trap-wise encounter counts across the sampling occasions with trap locations (co-ordinates) and potential activity centers as centroids of suitably constructed pixels covering the super area (sampled area and the buffer), with species specific habitat masks (coded as '1' for habitat, and as '0' for non-habitat).

An inevitable issue that prevails is that of the 'large encounters' that can either be due to 'influential observations' and/or 'outliers' in the count data. Outliers do occur when animals, especially as noted for some of species in this study, tend to move in groups of varying sizes (ex: livestock, wild pigs) for obvious reasons. Although the Poisson law theoretically allows relatively large values as "possible values", the fit may not be statistically acceptable, both in terms of Log Likelihood (LL) as well as the Bayesian p-values. We tried several strategies for dealing with such situations as listed below, and one or more of the following could be incorporated in the analyses.

(i) Ignore the entry or the entire data for the specific trap. This is not desirable since the outliers do contribute to the abundance.

(ii) It appears that one can try large data augmentation parameter (M) to deal with large encounters. If such encounters are not unduly large (say less than 6) this may work but will end up with very large or very small p-values. As noted in many other studies increasing the value of M poses many computational difficulties.

(iii) Square root transformation of count data with subsequent normality assumptions.

(iv) Regrouping the encounter data. This procedure will be elaborated in the sequel.

(v) Treat outlier counts as "marked animals" and carry out the analysis as discussed by Chandler and Royle (2013).

(vi) Set a uniform prior for ψ with a higher lower limit.

(vii) One can reduce the time duration of all sampling occasions so that the postulates of the Poisson process get reasonably validated. This will result in increased number of sample occasions. We could not try this for practical reasons. But such exercises are known to be helpful in Queuing theories.

Since our main interest is the estimation of abundance/density, we were tracking the LL and p-values for each option. Whenever the analysis based on the count data did not provide statistically satisfactory p-values, we tried one or more of the above mentioned alternatives, and chose the one with relatively large LL, and statistically acceptable p-values. Regrouping the encounter count data was found to be a useful option in our investigation. The features of this new procedure are highlighted in the following section.

2.7 *Salient features of the proposed grouping procedure adopted in the presence of large encounters*

Since both influential observations and outliers, per se, have a direct impact on the estimate of abundance (and hence on the estimate of density), it is necessary to involve them implicitly in the working model. The proposed procedure uses the count data on the groups (or herds) of individuals encountered at each trap location/ on each occasion, instead of the observed count data on the individuals, To elaborate the procedure, we first identify the group size of the animal (based on data), 'm' (say) which in a way represents the optimal size of the group mobility. To be specific, we transform the variable X representing the individual encounter data to another variable Y (encountered group size)such that

$$Y = 0 \text{ if } X = 0$$
$$Y = i \text{ whenever } X = (i-1)m+1 \text{ or } (i-1)m+2, ..., \text{ or im; for } i = 1, 2, ...$$

For example, if m=3, the count data such as 1, 3, 0, 5, 2 will be transformed as 1, 1, 0, 2, 1 which represents the group sizes of animals encountered. In short, the individual count data is recorded as count of groups of 'm' individuals, treating the residuals, if any, with less than 'm' individuals also as a separate group.

We assume that the transformed variable Y has a Poisson distribution. For this, Poisson assumption can be justified from two angles: (a) From a practical point of view, data relating to the species with large encounters have supported the Poisson assumptions on the transformed count data in terms of the resulting LL and Bayesian p-values. An illustrative example is discussed in the sequel. (b) Efforts to use other distributional assumptions like the normal and gamma did not yield satisfactory fit.

The transformed count data is analysed using the spatial count model to get the estimate for the group abundance (N^g say) which can be any measure of location of the posterior distribution of group abundance. The estimate of the abundance (*'N'*) of the species is (on assuming that the group sizes less than 'm' are equally likely) then given by

$$N = N^g \mathrm{m}$$

being the average number of individuals for the group size equal to N^g. The standard error of the proposed estimate can be obtained from the standard deviation (SD) of the posterior distribution of *N*. That is,

$$\text{SD of } N = \mathrm{m}(\text{ SD of } N^g)$$

In order to decide on the application of the grouping procedure one should know if the data has outliers. There are indicators discussed in outlier theories applicable to specific situation. A simple method applicable to Poisson encounters is based on the median of square root transformed data. For the given data, say Δ, on encounters, one can conclude the presence of large encounters if some of the observations exceed the value [median(2*sqrt(Δ))+3] where sqrt(Δ) is the data obtained by taking the square root of the observations in the data. This suggests that data sets with encounter rates around 2, we can expect Poisson assumptions to hold if the maximum encounter recorded is less than 6.

Two important issues arise in the proposed grouping method: (a) the validity of the crucial Poisson assumptions and (b) the determination of the group size in the analysis. Towards this end we present the findings relating to an investigation based on a real encounter data affected by some large encounters. The real data set used is the encounter data on wild pig (*Sus scrofa*) in Bukkapatna forest area in Karnataka. We used the encounter count data at the 99 camera trap locations recorded on 18 occasions. Poisson distribution (not the spatial count model) was fitted for both the given and the transformed data sets for different values of group size (m), on remembering the heterogeneous encounter rates across the trap locations. The values of Log Likelihood (LL) and the p-values based on Freeman-Tukey goodness of fit statistic over 100000 iterations are presented in Table 1.

It may be noted that the grouping of counts reduces the variability among the encounters. While small grouping value (m) retains the effect of outliers, the larger values

Table 1: Log likelihood and Bayesian p values

Group size (m)	Log Likelihood (LL)	p value
1 (given data)	-900.5408	1
2	-649.5435	1
4	-535.6989	0.9959
6	-501.1819	0.7549
8	-485.515	0.4071
10	-489.1831	0.2959
11	-476.1362	0.2243
12	-476.1362	0.2243
40 (Bernoulli data)	-461.7548	0.0877

can eventually transform the data to a Bernoulli (0,1) data. Initially the Log likelihood values show a decrease but they tend to stabilize for larger values of m. The Bayesian p-value shows that the fit may not be good for small and large values of m. Hence a good strategy seems to be to select the grouping size which has the p value around 0.5, with a stabilising value for the Log likelihood. But what ultimately decides is the performance of the estimate of abundance/density it provides when such group sizes are selected.

Finally, this grouping procedure was adopted only when the Poisson model did not fit the given count data in terms of the Bayesian p-value.

3 Results

Application of the model, perhaps for the first time in India, resulted in densities of 0.19 four-horned antelope, 0.07 blackbuck, 0.11 chinkara, 2.9 wild pigs, 0.44 Indian porcupine, 2.42 black-naped hare, 0.04 bonnet macaque, 0.14 peafowl, 0.39 large livestock, 15.45 small livestock, and 0.23 domestic dogs/km^2. We also extended the methodology for sloth bear (*Melursus ursinus*), which yielded a density estimate of 0.26 animals/km^2.

Tables 2 to 5 summarize the results for the 12 unmarked species taken up for analysis under this study. The posterior distribution of N was seen to be mostly skewed to the right, and hence one could use either the median, or the mean as a good measure of location. Except for black-naped hare, large livestock, small livestock and wild pig, the observed count data was used without any modification. Small log likelihood values and large p_1 and p_2 values (away from 0.5) were mainly due to large counts in the corresponding encountered count data.

Table 2: Density and abundance of four-horned antelope (*Tetracerus quadricornis*), blackbuck (*Antilope cervicapra*) and chinkara (*Gazella bennetti*) obtained from spatially correlated count data.

	Four-horned antelope		Blackbuck		Chinkara	
	Mean (SD)	95% HPDR	Mean (SD)	95% HPDR	Mean (SD)	95% HPDR
N	43.65 (16.08)	18 - 77	40.86 (15.55)	14 - 72	26.26 (9.03)	13 - 45
sigma	364 (74.56)	230.7 - 505.77	312.57 (45.11)	226.66 - 401.54	438.06 (72.83)	293.87 - 565.63
lam0	0.241 (0.076)	0.112 - 0.398	0.599 (0.336)	0.148 - 1.302	0.472 (0.174)	0.198 - 0.809
psi	0.438 (0.164)	0.168 - 0.789	0.410 (0.160)	0.144 - 0.745	0.267 (0.1)	0.107 - 0.469
LL (SD)	-194.151 (5.42)		-101.015 (3.5)		-275.276 (5.48)	
p_1 , p_2 values	0.858 , 0.566		0.890 , 0.510		0.820 , 0.569	
Density (SD)	0.19 (0.07)		0.07 (0.03)		0.12 (0.04)	

SD is standard deviation, 95% HPDR is highest posterior density region, N is the estimated number of individuals in the super area and 'Density' is per kilometer square

Table 3: Density and abundance of wild pig (*Sus scrofa*), crested porcupine (*Hystrix indica*) and black-naped hare (*Lepus nigricollis*) obtained from spatially correlated count data.

	Wild pig[1]		Crested porcupine[2]		Black-naped hare[3]	
	Mean (SD)	95% HPDR	Mean (SD)	95% HPDR	Mean (SD)	95% HPDR
N	642.7 (359.3)	185.5 - 1425.5	90.85 (47.91)	21 - 182	536.51 (57.51)	416 - 599
sigma	213.05 (69.98)	126.19 - 340.38	250.31 (79.41)	128.29 - 402.47	342.77 (35.01)	281.91 - 413.95
lam0	1.327 (1.563)	0.013 - 3.942	0.307 (0.095)	0.151 - 0.488	0.390 (0.062)	0.275 - 0.514
psi	0.325 (0.178)	0.086 - 0.721	0.455 (0.239)	0.115 - 0.927	0.892 (0.097)	0.691 - 1
LL (SD)	-74.859 (4.64)		-200.324 (5.12)		-763.943 (7.02)	
p_1 , p_2 values	0.476 , 0.326		0.808 , 0.519		0.003 , 0.617	
Density (SD)	2.9 (1.62)		0.45 (0.24)		2.42 (0.26)	

SD is standard deviation, 95% HPDR is highest posterior density region, N is the estimated number of individuals in the super area and 'Density' is per kilometer square

[1] Groups of 10 individuals was used as one count data.

[2] A 500 meter buffer with pixel size $316.22 * 316.228$ square meter was used.

[3] Groups of 3 individuals was used as one count data

Table 4: Density and abundance of sloth bear (*Melursus ursinus*), bonnet macaque (*Macaca radiata*), peafowl (*Pavo cristatus*) and obtained from spatially correlated count data.

	Sloth bear		Bonnet macaque		Peafowl	
	Mean (SD)	95% HPDR	Mean (SD)	95% HPDR	Mean (SD)	95% HPDR
N	58.06 (15.94)	28 - 89	9.834 (13.394)	1 - 39	33.68 (12.04)	12 - 56
sigma	256.46 (39.78)	190.46 - 337.87	222.31 (111.66)	73.23 - 348.32	321.77 (52.01)	237.68 - 434.44
lam0	0.410 (0.173)	0.138 - 0.791	3.837 (1.558)	0.585 - 6.035	0.677 (0.295)	0.288 - 1.432
psi	0.579 (0.163)	0.289 - 0.918	0.106 (0.134)	0 - 0.398	0.172 (0.065)	0.060 - 0.302
LL (SD)	-147.353 (5.05)		-15.383 (1.13)		-234.151 (4.27)	
p_1 , p_2 values	0.945 , 0.763		0.553 , 0.536		0.941 , 0.606	
Density (SD)	0.26 (0.07)		0.04 (0.26)		0.15 (0.05)	

SD is standard deviation, 95% HPDR is highest posterior density region, N is the estimated number of individuals in the super area and 'Density' is per kilometer square

Table 5: Density and abundance of large livestock (*Bos taurus and Bubalus bubalis*), small livestock (*Ovis aries* and *Capra aegagrus hircus*) and domestic dog (*Canis lupus familiaris*) obtained from spatially correlated count data.

	Large livestock[1]		Small livestock[2]		Domestic dog	
	Mean (SD)	95% HPDR	Mean (SD)	95% HPDR	Mean (SD)	95% HPDR
N	87.74 (32.51)	29 - 142	3421.5 (1824)	1055.5 - 7555.5	52.48 (23.33)	24 - 100
sigma	161.34 (33.95)	116.93 - 235.48	291.81 (73.69)	176.04 - 446.15	401.37 (62.65)	266.49 - 518.36
lam0	1.271 (0.395)	0.491 - 1.994	0.475 (0.499)	0.052 - 1.885	0.337 (0.087)	0.185 - 0.513
psi	0.089 (0.034)	0.029 - 0.148	0.350 (0.190)	0.088 - 0.770	0.265 (0.119)	0.102 - 0.508
LL (SD)	-228.575 (4.03)		-98.197 (4.6)		-331.942 (5.6)	
p_1 , p_2 values	0.868 , 0.476		0.637 , 0.678		0.880 , 0.617	
Density (SD)	0.39 (0.15)		15.45 (8.24)		0.23 (0.11)	

SD is standard deviation, 95% HPDR is highest posterior density region, N is the estimated number of individuals in the super area and 'Density' is per kilometer square

[1] Groups of 5 individuals was used as one count data

[2] Groups of 100 individuals was used as one count data

For each species, several iterations of the code was run across the varying values of M, and data transformation (only when the encountered count data on hand did not provide statistically acceptable p-values), and the final data presented is the one that was the best from a statistical point of view.

4 Discussion

Using the approach developed by Chandler and Royle (2013) we were able to use field data to derive density and abundance estimates of a wide range of species. To our knowledge, this is the first time in India such population density estimates for unmarked species, which is one of the most challenging issues in population ecology, have been obtained using detection/non-detection data .

In India most studies undertaken to estimate abundance and densities of wild ungulates are carried out mostly within protected areas (Varman and Sukumar, 1995; Karanth and Nichols, 2000; Biswas and Sankar, 2002; Jathanna et al., 2003; Bagchi et al., 2004; Gopal et al., 2010; Dave and Jhala, 2011; Kumara et al., 2012), and such results are unavailable for habitats outside protected areas despite several wildlife species are found beyond the boundaries of protected areas. Since carrying out distance sampling is both time and resource consuming, especially over wider areas, our paper highlights that by using existing camera trapping protocols, population density and abundance of unmarked species is a possibility for many ungulate species which was hitherto discarded or unutilised to it's true potential.

Our density results, except for wild pig, fall below the reported densities that used distance sampling for prey species monitoring (Karanth and Nichols, 2000; Bagchi et al., 2004). But the studies cited here were all carried out within protected areas where protection against threats such as poaching, that have an impact on ungulate densities

(Madhusudan and Karanth, 2002; Karanth et al., 2004), is perhaps better inside these protected zones leading to higher ungulate densities.

In addition, we have also been able to estimate density and abundance of sloth bear, a unmarked species of high conservation relevance due to its role as a key seed disperser (Sreekumar and Balakrishnan, 2002). Secondly, in India sloth bears are a highly conflict-prone species, and data on its abundance can play an important role in its management and conflict mitigation. The only study that estimated abundance of sloth bears was based on radio collar data (Garshelis et al., 1999) which is highly resource consuming. Previous attempts to estimate sloth bear abundance using camera trap data have not provided fruitful results (Gopalaswamy, 2006). In addition, distance sampling is also not a useful methodology for estimating population parameters of sloth bears due to extremely low visual detection rates. Hence this methodology could be replicated to understand population of sloth bears, and several other such species.

As noted by Chandler and Royle (2013), and as expected due to Poisson encounters, the results exhibited relatively large variances, and highly positively skewed posterior distributions when compared to analyses for marked individuals using the conventional abundance estimation software such as SPACECAP and SECR that uses capture-recapture framework. We first tried to assess the validity of the algorithm by comparing the results based on SPACECAP, with results obtained by using the same data using the code developed by us for unmarked animals. Although the 95% HPD (Highest Posterior Density) regions for our estimate covered the SPACECAP estimate of abundance, the posterior distribution was highly dispersed, and positively skewed.

The anticipated large variances and high positive skewness in the posterior distribution of N prompt large number of iterations during MCMC operations. The MCMC diagnostics did offer some help in this context. However the larger spread of the posterior distribution suggests larger values for M. But in that case, in addition to the increased computing time, one should have larger number of potential activity centers in the habitat area. To meet this requirement we tried to reduce the pixel sizes in order to increase the number of activity centers in the habitat to accommodate large M.

As widely commented, it is a challenge to venture into the inference for unmarked animals using spatial count data. From the data analysis point of view, it is more so when we have to move from the data based on Binomial encounters to Poisson encounter count data. Perhaps several approaches of analysis arising from different statistical/ecological perspectives, may be necessary to deal with count data before arriving at a plausible outcome.

The theoretically sound methodology requires planning for a realistic and executable sampling design which is both species specific as well as area specific. This will be our endeavour for the future.

The grouping method suggested and adopted in this exercise was the best option, among many other strategies, in dealing with large encounters. The classical root transformation of count data and the subsequent use of normality assumptions did not perform well when compared to the new grouping procedure mainly due to the fact that the inherent skewness in count data gets aggravated by the presence of large encounters. The poisson assumptions in the latter retain the skewness flavour in the data. It appears that there is scope for further theoretical probes in this direction.

Overall the methodology used in this paper can now provide highly valuable results using camera trap data that are currently used on a variety of wildlife species both within and outside protected areas. Now the by-catch data of these studies can prove invaluable.

Acknowledgement

This work was supported by the Kaplan Graduate Fellowship and Panthera which is greatly acknowledged. The authors thank M.D. Madhusudan for guidance, Harish N. S., Ashritha Anoop, Aparna Kolekar and H. C. Poornesha and all those who helped in data collection. The permission and support provided by the Karnataka Forest Department and all the staff is acknowledged and highly appreciated.

References

Bagchi, S., Goyal, S. P., and Sankar, K. (2004). Herbivore density and biomass in a semi-arid tropical dry deciduous forest of western india. *Journal of Tropical Ecology*, 20(4):475–478.

Biswas, S. and Sankar, K. (2002). Prey abundance and food habit of tigers (panthera tigris tigris) in pench national park, madhya pradesh, india. *Journal of Zoology*, 256(3):411–420.

Campbell, S. P., Clark, J. A., Crampton, L. H., Guerry, A. D., Hatch, L. T., Hosseini, P. R., Lawler, J. J., and O'Connor, R. J. (2002). An assessment of monitoring efforts in endangered species recovery plans. *Ecological Applications*, 12(3):674–681.

Carbone, C., Christie, S., Conforti, K., Coulson, T., Franklin, N., Ginsberg, J., Griffiths, M., Holden, J., Kawanishi, K., Kinnaird, M., et al. (2001). The use of photographic rates to estimate densities of tigers and other cryptic mammals. In *Animal Conservation forum*, volume 4, pages 75–79. Cambridge University Press.

Carbone, C., Pettorelli, N., and Stephens, P. A. (2010). The bigger they come, the harder they fall: body size and prey abundance influence predator–prey ratios. *Biology Letters*, page rsbl20100996.

Champion, F. W. (1992). *With a camera in tiger-land.* Chatto and Windus.

Chandler, R. B. and Royle, J. A. (2013). Spatially explicit models for inference about density in unmarked or partially marked populations. *The Annals of Applied Statistics*, 7(2):936–954.

Dave, C. and Jhala, Y. (2011). Is competition with livestock detrimental for native wild ungulates? a case study of chital (axis axis) in gir forest, india. *Journal of Tropical Ecology*, 27(3):239–247.

Garshelis, D. L., Joshi, A. R., and Smith, J. L. (1999). Estimating density and relative abundance of sloth bears. *Ursus*, pages 87–98.

Gopal, R., Qureshi, Q., Bhardwaj, M., Singh, R. J., and Jhala, Y. V. (2010). Evaluating the status of the endangered tiger panthera tigris and its prey in panna tiger reserve, madhya pradesh, india. *Oryx*, 44(3):383–389.

Gopalaswamy, A. M. (2006). *Estimating sloth bear abundance from repeated presence-absence data in Nagarahole-Bandipur National Parks, India.* PhD thesis, University of Florida.

Griffiths, M. and Van Schaik, C. (1993). Camera-trapping: a new tool for the study of elusive rain forest animals. *Tropical biodiversity*, 1(2):131–135.

Jathanna, D., Karanth, K. U., and Johnsingh, A. (2003). Estimation of large herbivore densities in the tropical forests of southern india using distance sampling. *Journal of Zoology*, 261(3):285–290.

Karanth, K. U. and Nichols, J. D. (2000). *Ecological status and conservation of tigers in India. Final Technical Report (February 1995 to January 2000).* Centre for Wildlife Studies.

Karanth, K. U., Nichols, J. D., Kumar, N. S., Link, W. A., and Hines, J. E. (2004). Tigers and their prey: predicting carnivore densities from prey abundance. *Proceedings of the National Academy of Sciences of the United States of America*, 101(14):4854–4858.

Kumara, H. N., Rathnakumar, S., Sasi, R., and Singh, M. (2012). Conservation status of wild mammals in biligiri rangaswamy temple wildlife sanctuary, the western ghats, india. *Current Science*, pages 933–940.

Madhusudan, M. and Karanth, K. U. (2002). Local hunting and the conservation of large mammals in india. *AMBIO: A Journal of the Human Environment*, 31(1):49–54.

Owen-Smith, N. and Mills, M. G. (2008). Predator–prey size relationships in an african large-mammal food web. *Journal of Animal Ecology*, 77(1):173–183.

Ramsey, D. S., Caley, P. A., and Robley, A. (2015). Estimating population density from presence–absence data using a spatially explicit model. *The Journal of Wildlife Management*, 79(3):491–499.

Royle, J. A., Chandler, R. B., Sollmann, R., and Gardner, B. (2013). *Spatial capture-recapture.* Academic Press.

Sreekumar, P. and Balakrishnan, M. (2002). Seed dispersal by the sloth bear (melursus ursinus) in south india. *Biotropica*, 34(3):474–477.

Varman, K. S. and Sukumar, R. (1995). The line transect method for estimating densities of large mammals in a tropical deciduous forest: An evaluation of models and field experiments. *Journal of Biosciences*, 20(2):273–287.