Methods in Ecology and Evolution

Wearn Oliver R (Orcid ID: 0000-0001-8258-3534)

Estimating animal density for a community of species using information obtained only from camera-traps

Oliver R. Wearn[1,2*], Thomas E. M. Bell[1], Adam Bolitho[3], James Durrant[4], Jessica K. Haysom[3,5], Sahil Nijhawan[1,4], Jack Thorley[3,6], J. Marcus Rowcliffe[1]

[1]Institute of Zoology, Zoological Society of London, Regent's Park, London, UK

[2]Fauna & Flora International, Vietnam Programme, 118 Tu Hoa, Tay Ho, Hanoi, Vietnam

[3]Imperial College London, Silwood Park, Buckhurst Road, Berkshire, UK

[4]University College London, Gower Street, London, UK

[5]Durrell Institute of Conservation and Ecology, University of Kent, Canterbury, UK

[6]Department of Earth, Oceans and Ecological Sciences, University of Liverpool, Liverpool, UK

*Corresponding author: Wearn, O. R. (oliver.wearn@gmail.com)

Running headline: Multi-species density estimation with camera-traps

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/2041-210X.13930

Abstract

1. Animal density is a fundamental parameter in ecology and conservation, and yet has remained difficult to measure. For terrestrial mammals and birds, camera-traps have dramatically improved our ability to collect systematic data across a large number of species, but density estimation (except for species with natural marks) is still faced with statistical and logistical hurdles, including the requirement for auxiliary data, large sample sizes, and an inability to incorporate covariates.

2. To fill this gap in the camera-trapper's statistical toolbox, we extended the existing Random Encounter Model (REM) to the multi-species case in a Bayesian framework. This multi-species REM can incorporate covariates and provides parameter estimates for even the rarest species. As input to the model, we used information directly available in the camera-trap data. The model outputs posterior distributions for the REM parameters – movement speed, activity level, the effective angle and radius of the camera-trap detection zone, and density – for each species. We applied this model to an existing dataset for 35 species in Borneo, collected across old-growth and logged forest. Here, we added animal position data derived from the image sequences in order to estimate the speed and detection zone parameters.

3. The model revealed a decrease in movement speeds, and therefore day-range, across the species community in logged compared to old-growth forest, whilst activity levels showed no consistent trend. Detection zones were shorter, but of similar width, in logged compared to old-growth forest. Overall, animal density was lower in logged forest, even though most species individually occurred at higher density in logged forest. However, the biomass per unit area was substantially higher in logged compared to old-growth forest, particularly among herbivores and omnivores, likely because of increased resource availability at ground level.

We also included body mass as a variable in the model, revealing that larger-bodied species were more active, had more variable speeds, and had larger detection zones.

4. Caution is warranted when estimating density for semi-arboreal and fossorial species using camera-traps, and more extensive testing of assumptions is recommended. Nonetheless, we anticipate that multi-species density estimation could have very broad application.

Keywords: density estimation, camera-trap, wildlife monitoring, Random Encounter Model, animal movement rate, animal activity, animal day-range, selective logging

1. Introduction

Density – the number of animals per unit area – is a fundamental parameter in ecology and conservation. For example, it underpins our ideas about the flow of energy and matter in ecosystems (Barnes et al., 2014), it underpins our understanding of global change (McRae et al., 2017), and it allows for robust assessment of conservation impact (Donald et al., 2007).

Despite its importance, density is a difficult parameter to measure. Many species are cryptic, actively avoid human observers, or occur at low density, making it difficult to gather sufficient data. For terrestrial mammals and birds, camera-traps have dramatically improved our ability to detect and monitor species, and even whole communities (Wearn & Glover-Kapfer, 2017, 2019). Camera-traps can be deployed in large numbers and left to autonomously collect data continuously for months, making even the rarest species a potential target for monitoring. However, logistical and statistical hurdles to density estimation with camera-traps mean that it has remained rare in the literature (Gilbert et al., 2021).

There are at least nine approaches currently available for density estimation with camera-traps, and all have difficulties associated with their implementation (Gilbert et al., 2021; Table S1.1, Appendix S1). For the Random Encounter Model (REM), the first approach that was developed specifically for camera-traps (Rowcliffe et al., 2008), uptake has remained low because of the widespread use of non-random camera placement (violating a key REM assumption), the inability to incorporate covariates during modelling, and the requirement for estimates of certain parameters that are difficult to measure in practice. REM consists of four component parts: 1) animal day-range, itself a product of movement speed (Rowcliffe et al., 2016) and activity levels (Rowcliffe et al., 2014); 2) the size of the detection zone, defined by the angle and radius (Rowcliffe et al., 2011); 3) the number of animal detections made at

each location, and 4) the sampling effort at each location. Whilst (3) and (4) are known with certainty, (1) and (2) are usually unknown. Day-range (the average daily distance travelled), in particular, is difficult to measure in wild populations. Previous REM studies have usually relied on auxiliary data to estimate this parameter, e.g. from collared animals (Caravaggi et al., 2016) or follows of habituated animals (Cusack et al., 2015).

Whilst the gap in the camera-trapper's toolbox for density estimation remained unfilled, an alternative – occupancy modelling – matured and rose to prominence (Delisle et al., 2021). Although occupancy is a metric of distribution, not density, best-practice methods have been well-described, and it only requires simple detection/non-detection data. However, it was really the development of community occupancy models, which simultaneously estimate occupancy (including covariate effects) for multiple species (Kéry & Royle, 2016), which has led to the primacy of occupancy over density estimation in recent camera-trap studies (Delisle et al., 2021).

Here, we extend the REM approach in several ways by implementing it in a Bayesian framework. This allows us to: 1) incorporate multiple species simultaneously, including rare species, 2) add covariates on any aspect of the model, and 3) propagate errors through the multiple components of the model. In addition, we rely only on information available in the camera-trap images to estimate parameters; no auxiliary data are used. We apply this approach to a community of Bornean mammals and birds, estimating the density of each species across land-use type.

2. Materials and methods

2. 1. Camera-trapping field methods

The study design and field protocols used to collect the data are described in full elsewhere (Wearn et al., 2017). For this work, we used data collected between May 2011 and November 2013 from 510 camera-trap locations spread across two land-use types in Malaysian Borneo: old-growth forest in the Maliau Basin Conservation Area and Brantian-Tatulit Virgin Jungle Reserve (n = 206), and selectively-logged forest in the Kalabakan Forest Reserve (n = 304). Cameras were left in place for a mean of 51 days (range: 1-109), for a total effort of 26,084 days (old-growth: 9,467; logged: 16,617). Crucial to the assumptions of REM, we placed cameras randomly with respect to animal movement (strictly within 5 m of pre-determined points). Camera-traps were set with the sensor approximately 30 cm off the ground and were programmed to take 10 images per trigger (at ~2 frames per second).

We also collected two types of calibration imagery – for camera locations and for camera-traps – which would later assist with the reconstruction of animal positions and speeds (described below and in Appendix S2). To calibrate camera locations, one person moved an object of known dimensions, in this case a 1 m pole, around the camera-trap's field of view (setup in its final deployed position), aiming for approximately 6 to 20 pole locations, whist a second person close to the camera triggered the passive infrared sensor with their hand. This process took 5-10 minutes per location. To calibrate the camera-trap model used (Reconyx HC500), we took images of an array of objects with known dimensions (again, 1 m poles) at known distances and angles from the camera. This was done at a field station, in a single location. We calibrated 16 camera-traps and here use the average calibration parameters across this sample.

2. 2. Animal position and speed data

We tracked the image pixel positions of animals in two ways: 1) the pixel positions of animals when the camera's passive infrared (PIR) sensor was definitely triggered, and 2) the pixel positions of animals as they moved across image sequences. The former would provide data to model the size of the camera-trap detection zone, whilst the latter would provide estimates of animal speed. The ground 'position' of animals was standardised as midway between the forelimbs. For definite trigger locations, we used animal positions only from the first image of each sequence of 10, given that PIR sensors can fail to detect animals even when apparently inside the notional detection zone. For the movement sequences, we tracked all straight-line movements of animals and paused the tracking if an animal appeared to show a behavioural response to the camera (e.g. moving or orientating towards the camera), if an animal was resting or sleeping, or if an animal's location could not accurately be determined (e.g. if the body was occluded by an object, or if they were very close to or very far from the camera).

Images were annotated with pixel positions using a modified version of the open-source graphics software, Blender (https://www.blender.org), with data stored in an underlying SQLite database. In total, 88,014 animal pixel positions were extracted (movement tracks and trigger locations combined), for 76 species across 307 locations. These data took approximately 180 person-days to generate. We here used only the data from old-growth and logged forest, and only for the focal species (defined below). This included 62,384 animal movement positions and 10,976 trigger locations, across 35 species and 249 locations (old-growth: 88; logged: 161).

Using the calibration imagery, we were able to convert each pixel position into radial and angular distances from the camera (see Appendix S2 for a detailed description). This involved using a pinhole camera model to describe two relationships: 1) between the size of the

calibration object in the image (in pixels) and its actual size and distance from the camera; and 2) between the X-Y pixel position in the image and the angular distance from the camera's principal axis. Calibration models for a deployment location used the first relationship above to calculate pole distances, and then used these distances, together with the second relationship, to describe the relationship between pixel positions in an image and radial and angular distances from the camera. Finally, the calibration models were used to predict radial and angular distances of animals from the camera based on their pixel positions within the images. These estimates were input directly to the detection zone modelling (see below), using only the definite trigger locations. To recover animal speed, trigonometry was used to calculate the total distance moved across a movement sequence, and then this distance was divided by the time elapsed (taken from the image date-time information).

2. 3. Bayesian multi-species REM

To date, no closed-form solution to the likelihood has been derived for the REM (but for a partial solution see: Jourdain et al., 2020). This has meant that it has not been possible to incorporate covariates into analyses. In addition, the sampling variances associated with each part of the REM have been propagated into the final density estimate using a combination of non-parametric bootstrapping and the delta approximation method (Rowcliffe et al., 2008).

Bayesian inference using Markov chain Monte Carlo (MCMC) methods offers a solution to these shortcomings, since the component parts of the REM model can be specified and then integrated with estimation of animal density in a single hierarchical model, all whilst propagating the errors associated with each component part. Moreover, specifying the REM as a hierarchical model allows complete flexibility over how and where covariates are introduced into the model. Here, we used land-use type as a covariate on density, as well as

in each of the REM sub-models for animal speed, activity level, detection angle, detection radius and the detection counts. We also included species body-mass as a hyper-covariate to help explain speed, activity level, detection angle and detection radius.

A Bayesian approach also makes REM possible for rare species. Likelihood-based approaches – as currently used for the activity level, speed and detection zone components of the REM – are subject to asymptotic assumptions that are unlikely to hold in the case of rare species with sparse data; Bayesian methods make no such assumptions. Perhaps most powerfully, a Bayesian approach makes a multi-species or "community-level" REM possible, by treating species as a random effect in the hierarchical model. This approach, which "borrows" information from common species in order to allow parameters for the rarer species to be estimated, has become widely used in community occupancy modelling (e.g. Tobler et al., 2015; Wearn et al., 2017).

Here, we defined our focal community as the 35 species of mammal and bird which were: large (> 500 g body mass), habitually active on the ground (i.e. can be reliably detected with camera-traps, if present), and non-migratory in the study areas. Small animals were excluded because they likely violate the assumptions of the distance sampling approach we used to model camera detection zones. Specifically: 1) we suspected that they might sometimes pass underneath camera-traps without being detected, and 2) detection probability for some species might always be < 1, even when directly in front of the camera.

2. 3. 1. Day-range

Day-range was estimated as the product of movement speed and activity level. Observed speeds *i* for species *s* were modelled using a size-biased (SB) log-normal distribution (Rowcliffe et al., 2016):

$$speed_{i,s} \sim \text{SB Lognormal}(\mu_{i,s}, \sigma_{i,s}) \quad (1)$$

where μ and σ are the mean and standard-deviation of the speeds, respectively. A size-biased distribution was appropriate because faster animals are more likely to contact camera-traps. A log-normal distribution was assumed, since it previously was found to fit real data better than a gamma or Weibull distribution (Rowcliffe et al., 2016). A log-normal distribution typically produces higher speed estimates than the alternatives, leading to lower estimates of animal density (Rowcliffe et al., 2016).

Covariates on movement speed and standard deviation were introduced via linear predictors, using an identity link and a log link, respectively. Here, we used land-use type, *LU,* as a sole covariate:

$$\mu_{i,s} = \alpha 0_s + \alpha 1_s LU_i \quad (2)$$

$$\log(\sigma_{i,s}) = \beta 0_s + \beta 1_s LU_i \quad (3)$$

The intercepts and slopes of these linear predictors for a given species were themselves random effects in hyper-distributions for the community-level mean speed and community-level variance. For example, for the intercept and slope for mean speed:

$$\alpha 0_s \sim \text{Normal}(mean_{\alpha 0_s}, sigma_{\alpha 0_s}) \quad (4)$$

$$\alpha 1_s \sim \text{Normal}(mean_{\alpha 1_s}, sigma_{\alpha 1_s}) \quad (5)$$

where the $mean$ and $sigma$ hyper-parameters are the mean and standard deviation, respectively, of a normal distribution.

In turn, the $mean$ hyper-parameters of the hyper-distributions for the intercepts $\alpha 0_s$ and $\beta 0_s$ were a function of species body mass (taken from Wearn et al., 2017). For example, for $mean_{\alpha 0_s}$:

$$mean_{\alpha 0_s} = \gamma 0 + \gamma 1.\log(body\ mass_s) \quad (6)$$

The use of body mass as a hyper-covariate means that the model can leverage information about the expected mean and standard deviation of movement speed for a given species not only from the observed speed data (which, for rare species, might be sparse), but also the broader scaling relationship between the body mass of a species and its mean speed and standard deviation. We expected that larger species would travel more quickly and have more variable speeds.

To model activity levels, we used a combination of non-parametric and Bayesian approaches. This was done because the key distribution – a circular kernel distribution (Ridout & Linkie, 2009) – was not available in the software used for Bayesian inference. First, we extracted independent times-of-detection for each species (i.e. detections separated by more than 1 hour) from the camera-trap data. We then used the 'activity' package in R (Rowcliffe et al., 2014) to fit circular kernel distributions to the data for each combination of species and land-use, and in each case estimate the mean and standard error of activity level. We then generated a sample of 1000 activity levels for each combination of species and land-use, using a beta distribution with parameters matching those that were estimated using the circular kernel distribution. These activity level samples were then taken forward into the Bayesian hierarchical model, propagating forwards the errors associated with the activity level estimation. For combinations of species and land-use which had fewer than 15 times-of-detection, activity level estimation in the 'activity' package was not attempted; instead

activity levels for these species were directly estimated in the Bayesian hierarchical model, by exploiting the modelled relationship between body mass and activity level (see below). For these species, we expect activity level estimates to be less precise and less accurate, since only their body mass information was included in the model.

The activity level samples *j* for species *s* were modelled using a beta distribution:

$$activity_{j,s} \sim \text{Beta}(m_{j,s}, v_{j,s}) \quad (7)$$

where m and v are the mean and sample-size parameters. This parameterisation of the beta distribution was chosen instead of the standard parameterisation in terms of two shape parameters so that covariates – here, only land-use type – on the mean activity level and sample-size parameters could be introduced. Linear predictors with a logit and log link were used for the mean and sample-size parameters, respectively:

$$\text{logit}(m_{j,s}) = \varepsilon 0_s + \varepsilon 1_s LU_j \quad (8)$$

$$\log(v_{j,s}) = \tau 0_s + \tau 1_s LU_j \quad (9)$$

Like the speed sub-model, the intercepts and slopes of these linear predictors were themselves random effects in hyper-distributions for the community-level mean and variance of activity level. In turn, the hyper-parameter for mean activity level at the community level was a function of the body mass of a species, with the expectation that larger species would be more active. No body mass covariate was used for the sample size parameter at the community level, since we had no specific hypothesis in this case.

2. 3. 2. Detection zone

Distances r (i.e. radii) and angles θ for the trigger locations, were modelled using a distance sampling approach (Rowcliffe et al., 2011). Whilst this approach corrects for imperfect

detection, an important assumption here is that detection is nonetheless certain at a single point directly in front of the camera-trap. If this is violated, density estimates will be negatively biased.

The approach for distances and angles was similar, so we here focus on the approach to the distance data and highlight where the approach for angles was different. Detection distances were modelled as an outcome of both a *state* process and an *observation* process. The state process was the latent number of animal 'passes', i.e. the number of times animals entered the notional detection zone of a camera-trap, c. These are *potential* triggers, of which not all actually cause cameras to be triggered. These animal passes were assumed to be Poisson-distributed:

$$passes_{c,s} \sim \text{Poisson}(\lambda_{c,s}) \quad (10)$$

The observation process consisted of the probability of an animal appearing at a given distance, multiplied by the probability that the animal would be detected by the passive infrared sensor at that given distance. Animal passes are more likely to occur further from a camera, because the width of the field-of-view is larger at greater distances. This is also the case in conventional distance sampling using point transects, wherein a triangular distribution is used (Buckland et al., 2001; Rowcliffe et al., 2011). The probability density of the distances of the latent passes was therefore:

$$f(r_{c,s}) = \frac{2.r_{c,s}}{{r_{max}}^2} \quad (11)$$

where r_{max} was the maximum distance at which camera-traps triggered in the study (here, 24.2 m). To model the triggering properties of a camera-trap sensor, specifically the declining sensitivity with distance, a hazard detection function was used:

$$p(r_{c,s}) = 1 - \exp(-\left(\frac{scale_{c,s}}{r_{c,s}}\right)^{shape_{c,s}}) \quad (12)$$

where the $scale$ and $shape$ parameters define the width and steepness of the hazard function, respectively. The hazard allows for a "shoulder" of certain detection close to the camera, and was previously found to fit real camera-trap data better than the half-normal (Rowcliffe et al., 2011).

Given probability density functions for the latent distances and their probabilities of observation, the overall detection probability of a given camera-trap, evaluated across the full length of the notional detection zone, was therefore:

$$p^*_{c,s} = \int_0^{r_{max}} f(r_{c,s}) g(r_{c,s})\, dr \quad (13)$$

Using the overall detection probability, the effective radius r^* was then calculated as a derived parameter. This was done by re-arranging the canonical distance sampling estimator (Buckland et al., 2001) like so:

$$r^*_{c,s} = \sqrt{p^*_{c,s} {r_{max}}^2} \quad (14)$$

Finally, the observed number of triggers was modelled with a binomial distribution, with the latent number of animal passes determining the number of 'trials' and the detection probability determining the probability of success:

$$triggers_{c,s} \sim \text{Binomial}(passes_{c,s}, p^*_{c,s}) \quad (15)$$

For modelling detection angles, the observation process was taken from the line transect case in distance sampling, and a uniform distribution for the latent angles was therefore used (Buckland et al., 2001; Rowcliffe et al., 2011). This uniform distribution was defined between zero and the maximum angle observed in the study (here, 58°). For the detection function, a

half-normal was used instead of the hazard, because it was previously found to fit real camera-trap data better (Rowcliffe et al., 2011).

Covariates were introduced into the detection zone sub-model using a similar approach to that already outlined for the speed and activity sub-models. Land-use covariates were introduced on the scale and shape parameters of the hazard detection function using linear predictors and log link functions. The same was done for the scale parameter of the half-normal detection function. Community-level hyper-distributions allowed the detection function parameters to vary by species. In addition, we expected that larger species would be detected at greater distances and greater angles, and so the community-level means for the two scale parameters (for the hazard and the half-normal) were also modelled as a function of species body mass. No species-level covariates were used for the shape parameter of the hazard, as we had no clear hypotheses in this case.

2. 3. 3. Observed counts and density

We modelled animal density using a gamma distribution, which is a natural choice for any continuous, positive variable, and can accommodate 'long tail' distributions (Bolker, 2008). The other choice was a log-normal distribution, which is often indistinguishable in practice.

$$density_{c,s} \sim \text{Gamma}\left(shape_{density\ c,s}, rate_{density\ c,s}\right) \quad (16)$$

The two parameters of the gamma – its shape and rate – were modelled as a function of land-use type, using log links. Hyper-distributions for the two parameters also allowed density to vary by species.

We added zero-inflation according to land-use type l, in order to allow for cases where a species simply did not occur in a given land-use. The occurrence of a species was modelled with a Bernoulli distribution:

$$occurrence_{l,s} = \text{Bernoulli}(\Omega_s) \tag{17}$$

For species which were detected in a given land-use during sampling, $occurrence_{l,s}$ is equal to 1 in every MCMC iteration. Otherwise, it varies across iterations, and the mean of these values provides the probability that the species occupies the given land-use, despite not being detected.

The final step of the Bayesian implementation of REM was to link the various component parts of the REM. We did this by re-arranging the REM to make the detection counts the subject, and then formulating it as a linear predictor inside a log link:

$$\log(\lambda_{counts_{c,s}}) = \log(density_{c,s}) + \log(effort_{c,s}) + \log(speed_{c,s}) + \log(activity_{c,s}) + \log(2 + \theta^*_{c,s}) + \log(r^*_{c,s}) - \log(\pi) \tag{18}$$

where $effort$ is the number of camera-trap days and θ^* is the effective detection angle. The speed and activity for a given species at a given location were calculated as derived variables, since they were not directly modelled. The detection counts were then modelled with a Poisson distribution multiplied by the zero-inflation by land-use type:

$$counts_{c,s} \sim \text{Poisson}(\lambda_{counts_{c,s}}).\,\text{occurrence}_{c,s} \tag{19}$$

For the detection counts, we tallied the number of times animals 'contacted' camera-traps, i.e. were observed to enter the field-of-view of a camera-trap. For group-living animals (in particular, bearded pig *Sus barbatus* and Sunda pig-tailed macaque *Macaca nemestrina*), it was difficult to verify for each individual when a contact had occurred (because other

individuals in the group may have already triggered the camera). We therefore counted for each group the total number of individuals seen across a sequence of images and assigned one contact to each individual. If individuals contacted cameras on average less than or more than once, then our density estimates for group-living species will be too high or too low, respectively.

2. 4. Inference via MCMC

We formulated this model in BUGS code and used JAGS (version 4.3.0; Plummer, 2017), called from R (version 3.6.1; R Development Core Team, 2019), to obtain samples of the joint posterior distribution. Modifications to the parameterisation of the model outlined above were necessary in order to formulate it in the BUGS language (Appendix S2). We used standard vague priors for all parameters (Kéry & Royle, 2016). We also assessed the goodness-of-fit of our detection count model with a posterior predictive check, based on the Pearson residuals of the counts (Kéry & Royle, 2016; Appendix S2).

3. Results

Across the species community, day-range (Fig. 1) was a median of 1.2 km per day shorter in logged forest compared to old-growth forest (95% credible interval (CI): -2.37 to -0.18), amounting to a median decrease of 20% (CI: -33.6 to -3.56 %). This change in day-range across land-use type was driven by changes in animal speed rather than activity levels (Fig. 2). Species in logged forest travelled a median of 0.12 km hr^{-1} (CI: -0.21 to -0.05 km hr^{-1}) or 18% (CI: -26.7 to -8.30 %) slower than in old-growth forest, whilst there was no clear difference in activity levels (median difference = -0.004; CI: -0.047 to 0.048). There was a positive relationship between (the log of) species body mass and activity level (slope parameter = 0.17;

CI: 0.06 to 0.28; Fig. 3A), but no clear relationship with (the log of) speed (slope = 0.04; CI: -0.12 to 0.19). Species with higher body masses did, however, have more variable speeds (slope for the precision of log-speed = -0.07; CI: -0.13 to -0.002; Fig. 3B).

Detection zones across species (Fig. 4) were smaller in logged forest compared to old-growth forest by a median of 0.61 m^2 (CI: -1.18 to -0.16 m^2), amounting to a median decrease of 16% (CI: -25.7 to -5.11 %). This difference was overwhelmingly due to the shorter effective detection distances in logged forest, rather than any change in effective detection angles (Fig. 5). Across species, the effective distance was a median of 0.3 m shorter in logged forest than in old-growth forest (CI: -0.53 to -0.09 m), amounting to a decrease of 8.3 % (CI: -13.4 to -2.72 %). There was a positive relationship between (log) body mass and both effective detection angles (slope = 0.38; CI: 0.31 to 0.48; Fig. 6A) and distances (slope = 0.27; CI: 0.11 to 0.42; Fig. 6B). Moreover, the slopes in both cases were different between land-use types (slope differences from old-growth to logged forest for angles and distances, respectively: -0.06 and 0.15; Fig. 6).

Density estimates ranged from as low as 0.01 animals km^{-2} (Malay weasel *Mustela nudipes*), up to 17 animals km^{-2} (greater mouse-deer *Tragulus napu*) (Table S3.1, Appendix S3). The uncertainty associated with the estimates, as measured using the coefficient of variation (CV), increased sharply when the sample size (number of 'contacts') for a species was less than 30 (Fig. S4.1, Appendix S4). However, in most cases (42 of 78 estimates), CVs were in an acceptable range (CV < 35%; Bessone et al., 2020).

It was not possible to independently validate the density estimates, but there was broad correspondence with previously-published density estimates from other sites, where data were available (Table S3.2, Appendix S3). For 11 of the 14 comparisons that were possible,

the REM estimates were within the range of previous estimates, and lower than previous estimates in the remaining cases (Fig. S4.2, Appendix S4). For the Bornean orangutan (*Pongo pygmaeus*) and Sunda clouded leopard (*Neofelis diardi*), REM densities were also compared *post-hoc* with alternative estimates from the same study landscape, derived from nest count surveys and spatially-explicit capture-recapture, respectively (Appendix S4). Although the comparisons were weakened by spatial and temporal mismatches in the datasets, the REM-derived orangutan densities appeared low when compared to alternative estimates, whilst the clouded leopard densities were comparable (Fig. S4.3, Appendix S4).

Most species increased in density from old-growth to logged forest (20 of 35 species; Fig. 7). Notably, sambar deer (*Rusa unicolor*), bearded pig (*Sus barbatus*) and Malay civet (*Viverra tangalunga*) increased by 674% (CI: 362.2 to 1190 %), 321% (CI: 196.2 to 489.3 %) and 255% (CI: 126.4 to 449.7 %), respectively. However, there was no clear evidence across species that the average change in density was positive, because any gains were matched by losses (median change = 12.4%; CI: -24.4 to 51.6 %). Indeed, there were 8.9 fewer animals km^{-2} overall in logged compared to old-growth forest (CI: -16.3 to -2.38 animals km^{-2}), which amounts to a 20% decline (CI: -32.4 to -5.91 %). This pattern was driven by large declines from old-growth to logged forest in some of the more common species. In particular, the greater and lesser mouse-deer species (*T. napu* and *T. kanchil*) declined by 98% (CI: -99.3 to -94.3 %) and 69% (CI: -83.8 to -42.0 %), respectively, and the Bornean yellow muntjac (*Muntiacus atherodes*) declined by 86% (CI: -93.3 to -67.2 %).

Finally, we calculated *post-hoc* the biomass of animals in major trophic groupings (groupings followed Wearn et al., 2017). The resulting 'trophic pyramids' (Fig. 8) illustrated the dramatic increase (by 754 kg km^{-2}; CI: 558 to 1021 kg km^{-2}) in herbivore biomass in logged versus old-

growth forest, as well as the reduction (-85.4 kg km^{-2}; CI: -118 to -61.1 kg km^{-2}) in frugivore biomass. Overall, animal biomass was 890 kg km^{-2} higher in logged forest compared to old-growth forest (CI: 682 to 1161 kg km^{-2}).

4. Discussion

We here extended the REM approach to a Bayesian mode of inference, allowing for multi-species density estimation from camera-trap data for the first time. We used as input to the model only information that is available in the camera-trap images, rather than relying on auxiliary data or intensive measurements in the field. The output of the model was a set of posterior distributions for the REM parameters for each species. Uncertainty about movement speed, activity level and the detection zone was fully propagated into the final density estimates. The Bayesian hierarchical modelling approach also allowed us to incorporate covariates into the REM. This provided us with REM parameter estimates by land-use type and revealed the importance of body-size in explaining differences across species in both their ecology and in the camera-trap detection process.

4. 1. New insights gained

The multi-species REM yielded density estimates for 35 species of Bornean mammal and bird, of which 17 are threatened with extinction (Vulnerable or higher on the IUCN Red List; IUCN, 2021). For many of these species, this is the first time their density has been estimated, improving our understanding of their conservation status. It is also now possible to make 'apples to apples' comparisons of densities, either across species or across sites. For example, a conservation manager can infer that sun bears (*Helarctos malayanus*) were rarer than banded civets (*Hemigalus derbyanus*), and might therefore be more sensitive to any offtake

(e.g. through snare trapping), but that banded civets became rarer in logged compared to old-growth forest (in contrast to sun bears), and therefore are more sensitive to disturbance from logging. These comparisons would not be as straightforward using occupancy as a metric, because occupancy is sensitive to differences in home-range size (Efford & Dawson, 2012) and the duration of sampling (Steenweg et al., 2018). This is also the case for relative abundance measures, such as local abundance (e.g. Granados et al., 2016; Wearn et al., 2017). A multi-species REM also opens up the possibility of monitoring population trends across a suite of species simultaneously, although detecting changes with high confidence will be remain difficult for the rarest species (with sample sizes < 30) due to the low precision of the estimates.

We can also combine density estimates across species and calculate emergent, ecosystem-level variables, such as the overall density of animals or the total biomass. This revealed that, even though most species occurred at higher density in logged forest (in line with previous studies), there were overall fewer animals per unit area in logged compared to old-growth forest. This has not previously been reported for Borneo, given the focus on occupancy and relative abundance (Granados et al., 2016; Wearn et al., 2017), and likely has implications for the ecosystem functioning of logged forests in the region.

REM also provides insights into the ecology of study species and is therefore much more than just a way of "counting animals". Similar benefits have been highlighted for spatially-explicit capture-recapture modelling of marked species (Royle et al., 2014), but the same cannot be said for any of the other density estimation methods for unmarked species. Here, the REM approach revealed that, compared to old-growth, logged forests were characterised by slower-moving, larger-bodied animals, with more variable speeds. The likely explanation for

this is that logged forests provide a higher density of resources, in particular resources at ground level for herbivores and omnivores (Meijaard & Sheil, 2008), leading to less time spent travelling and more time foraging. This is an example of how the deeper insights provided by the REM approach allow for more direct links to be made between pattern and process.

4. 2. Assumptions and caveats with unmarked density estimation

Whilst we have removed important barriers in this study towards broader use of REM, important assumptions in the approach remain (reviewed in: Wearn & Glover-Kapfer, 2017; Gilbert et al., 2021). These assumptions require further testing and validation before the REM-based density estimates we present here can be assumed to be accurate. We urge caution when interpreting the exact density estimates for each species, especially in an applied setting, where incorrect inferences might lead to perverse management and conservation outcomes. Ideally, REM-based estimates should be used in decision-making alongside other forms of evidence (e.g. expert knowledge) and density estimates should be compared to those derived from other approaches.

We consider the strongest assumption in the REM approach to be that all animals in the study population are active at the peak of activity (Rowcliffe et al., 2014). This is likely a reasonable assumption for terrestrial, non-fossorial species with clearly defined nocturnal or diurnal activity patterns. However, violations of the assumption might occur in cathemeral species and, most especially, in arboreal and fossorial species, which can carry out their activities fully outside the realm of ground-based camera-traps. If only a proportion of animals in the study area are active at the peak, then we may have overestimated activity levels, and underestimated density. We consider that this might have been the case for some of the most arboreal and fossorial species in our study, including the semi-arboreal yellow-throated

marten (*Martes flavigula*), binturong (*Arctictis binturong*), long-tailed macaque (*Macaca fascicularis*) and Bornean orangutan, as well as the fossorial Sunda pangolin (*Manis javanica*).

The REM approach is not alone in making strong assumptions about animal activity. All of the density estimation methods for unmarked species make assumptions about activity, and only the capture-recapture approaches are free of this. In common with REM, the Random Encounter and Staying Time (REST) and camera-trap Distance Sampling (CT-DS) approaches explicitly include activity level as a parameter (Howe et al., 2017; Nakashima et al., 2018; Bessone et al., 2020), whilst the Time-to-event (TTE), Space-to-event (STE) and Instantaneous Sampling (IS) approaches (Moeller et al., 2018) have an undeclared assumption that animals are always available for detection. Given the importance of understanding activity levels, further studies would be helpful to test assumptions, for example using arboreal camera-traps (e.g. Haysom et al., 2021).

Other aspects of the REM would also warrant further validation and-or direct comparison with other approaches, including the estimation of movement speed (e.g. by comparing with GPS-collaring) and detection zone dimensions. Nakashima et al. (2022) recently showed that a paired camera-trap approach can be used to validate the assumption that detection is certain directly in front of the camera-trap. This study (using a different camera-trap model to our study) showed that 9% of detections in a 1.6 m^2 notional detection zone were missed (for species > 500 g). This has implications for all the unmarked approaches to density estimation which exploit the camera-trap sensor (REM, REST, CT-DS and TTE).

4. 3. Applying the multi-species REM in practice

The two most important survey design requirements for REM (reviewed in: Wearn & Glover-Kapfer, 2017) are that cameras are placed randomly with respect to animal movement and that high-performance camera-traps are used, i.e. those with a fast trigger speed (< 0.3 s) and ability to record either video or sequences of images in quick succession. For modelling speed and the detection zone, animal position data are required, which can be generated either by: i) collecting intensive field measurements (Rowcliffe et al., 2016), ii) using reference points with a known distance that are placed, or occur naturally, in the field-of-view of camera-traps (Caravaggi et al., 2016; Palencia et al., 2021), or iii) calibrating camera-trap locations and the camera-trap models used by taking images of objects of known size (see Appendix S5 for guidance). The principal benefit of methods (ii) and (iii) are that the animal locations are generated in the lab rather than in the field. This provides opportunities to scale the data collection cost-effectively, e.g. using a citizen science approach or an automated approach to track animal locations (e.g. using a 'bounding box' classifier; Norouzzadeh et al., 2021). Collecting the position data *post-hoc* in the lab also means that the detection locations for rare species are already known, allowing those sequences to be prioritised.

Once generated, the position data can be analysed with a REM approach, as done here, but with some alterations in the processing steps can also be analysed using a REST, CT-DS or TTE approach. For REST, a notional zone of perfect detection needs to be defined, for which the effective detection zone could be used, and then the average 'staying time' within this zone can be calculated from the position data. For CT-DS, the distances from the camera at exact clock times (e.g. every 2 seconds) need to be extracted from the position data. TTE requires an estimate of animal speed and so the processing steps are similar to REM.

4. 4. A new method in the camera-trapper's toolbox

The multi-species REM represents a much sought-after new tool at the disposal of ecologists and conservationists, albeit with some strong accompanying assumptions. We expect that similar multi-species models could be developed for REST, CT-DS and the other unmarked methods. The analyses for unmarked density estimation are more laborious than for occupancy, but they provide the gold-standard of monitoring – density – rather than a measure of distribution. We also anticipate that the labour savings on the horizon with the rapid development of automation algorithms for camera-trap images will free up time that would have been spent on the detection and identification of species (Norouzzadeh et al., 2021). With investment in training, as well as further software development and testing of assumptions, we see no reason why multi-species monitoring of populations cannot expand greatly in use across a diversity of contexts, species and biomes.

Acknowledgments

This research drew upon camera-trap data collected in a previous research project, and we again thank Yayasan Sabah, Benta Wawasan, the Sabah Forestry Department, the Maliau Basin Management Committee (permit no. MBMC/2010/15) and the Sabah Biodiversity Council (permit no. JKM/MBS.1000-2/3-84) for providing access and research permissions. A grant from the Sime Darby Foundation to the Royal Society South East Asia Rainforest Research Programme, specifically the 'Stability of Altered Forest Ecosystems' (SAFE) Project, supported the collection of the original camera-trap dataset. Fieldwork would not have been possible without the collective effort of the SAFE Project's field staff. Models were run on the Google Cloud Platform, using credits provided under a 'Data Solutions for Change' grant from Google to the Zoological Society of London.

Data availability

All data analysed in this manuscript (including detection counts, movement speeds, activity times, detection angles and detection distances) are available in the Zenodo repository at: https://doi.org/10.5281/zenodo.6635649 (Wearn et al., 2022).

Conflict of interest

We declare that we have no conflict of interest.

Authors' contributions statement

Oliver R. Wearn and J. Marcus Rowcliffe conceived the ideas and designed the methods; Oliver R. Wearn, Jessica K. Haysom and Jack Thorley collected the field data; Oliver R. Wearn, Thomas Bell and Adam Bolitho digitised the animal locations, using software produced by James Durrant; J. Marcus Rowcliffe and Sahil Nijhawan collected the camera-trap calibration data; Oliver R. Wearn analysed the data and led the writing of the manuscript. All authors contributed to the drafts and gave final approval for publication.

References

Barnes, A. D., Jochum, M., Mumme, S., Haneda, N. F., Farajallah, A., Widarto, T. H., & Brose, U. (2014). Consequences of tropical land use for multitrophic biodiversity and ecosystem functioning. *Nature Communications*, *5*, 5351.

Bessone, M., Kühl, H. S., Hohmann, G., Herbinger, I., N'Goran, K. P., Asanzi, P., … Fruth, B. (2020). Drawn out of the shadows: Surveying secretive forest species with camera trap distance sampling. *Journal of Applied Ecology*, *57*, 963–974.

Bolker, B. (2008). *Ecological Models and Data in R*. Princeton University Press.

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., & Thomas, L.

(2001). *Introduction to Distance Sampling: Estimating Abundance of Biological Populations*. Oxford University Press, Oxford.

Caravaggi, A., Zaccaroni, M., Riga, F., Schai-Braun, S. C., Dick, J. T. A., Montgomery, W. I., & Reid, N. (2016). An invasive-native mammalian species replacement process captured by camera trap survey random encounter models. *Remote Sensing in Ecology and Conservation*, *2*, 45–58.

Cusack, J. J., Swanson, A., Coulson, T., Packer, C., Carbone, C., Dickman, A. J., … Rowcliffe, J. M. (2015). Applying a random encounter model to estimate lion density from camera traps in Serengeti National Park, Tanzania. *The Journal of Wildlife Management*, *79*, 1014–1021.

Delisle, Z. J., Flaherty, E. A., Nobbe, M. R., Wzientek, C. M., & Swihart, R. K. (2021). Next-generation camera trapping: systematic review of historic trends suggests keys to expanded research applications in ecology and conservation. *Frontiers in Ecology and Evolution*, *9*, 1–18.

Donald, P. F., Sanderson, F. J., Burfield, I. J., Bierman, S. M., Gregory, R. D., & Waliczky, Z. (2007). International conservation policy delivers benefits for birds in Europe. *Science*, *317*, 810–813.

Efford, M., & Dawson, D. (2012). Occupancy in continuous habitat. *Ecosphere*, *3*, 1–15.

Gilbert, N. A., Clare, J. D. J., Stenglein, J. L., & Zuckerberg, B. (2021). Abundance estimation of unmarked animals based on camera-trap data. *Conservation Biology*, *35*, 88–100.

Granados, A., Crowther, K., Brodie, J. F., & Bernard, H. (2016). Persistence of mammals in a

selectively logged forest in Malaysian Borneo. *Mammalian Biology*, *81*, 268–273.

Haysom, J. K., Deere, N. J., Wearn, O. R., Mahyudin, A., Jami, J. bin, Reynolds, G., & Struebig, M. J. (2021). Life in the Canopy: Using Camera-Traps to Inventory Arboreal Rainforest Mammals in Borneo. *Frontiers in Forests and Global Change*, *4*.

Howe, E. J., Buckland, S. T., Després-Einspenner, M.-L., & Kühl, H. S. (2017). Distance sampling with camera traps. *Methods in Ecology and Evolution*, 1558–1565.

IUCN. (2021). The IUCN Red List of Threatened Species. Version 2021.1.

Jourdain, N. O. A. S., Cole, D. J., Ridout, M. S., & Rowcliffe, J. M. (2020). Statistical development of animal density estimation using Random Encounter Modelling. *Journal of Agricultural, Biological, and Environmental Statistics*, *25*, 148–167.

Kéry, M., & Royle, J. A. (2016). *Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 1: Prelude and Static Models.* Academic Press.

McRae, L., Deinet, S., & Freeman, R. (2017). The diversity-weighted Living Planet Index: Controlling for taxonomic bias in a global biodiversity indicator. *PLOS ONE*, *12*, e0169156.

Meijaard, E., & Sheil, D. (2008). The persistence and conservation of Borneo's mammals in lowland rain forests managed for timber: observations, overviews and opportunities. *Ecological Research*, *23*, 21–34.

Moeller, A. K., Lukacs, P. M., & Horne, J. S. (2018). Three novel methods to estimate abundance of unmarked animals using remote cameras. *Ecosphere*, *9*, e02331.

Nakashima, Y., Fukasawa, K., & Samejima, H. (2018). Estimating animal density without individual recognition using information derivable exclusively from camera traps. *Journal of Applied Ecology*, *55*, 735–744.

Nakashima, Y., Hongo, S., Mizuno, K., Yajima, G., & Dzefck, Z. C. B. (2022). Double-observer approach with camera traps can correct imperfect detection and improve the accuracy of density estimation of unmarked animal populations. *Scientific Reports*, *12*, 1–10.

Norouzzadeh, M. S., Morris, D., Beery, S., Joshi, N., Jojic, N., & Clune, J. (2021). A deep active learning system for species identification and counting in camera trap images. *Methods in Ecology and Evolution*, *12*, 150–161.

Palencia, P., Rowcliffe, J. M., Vicente, J., & Acevedo, P. (2021). Assessing the camera trap methodologies used to estimate density of unmarked populations. *Journal of Applied Ecology*, *58*, 1583–1592.

Plummer, M. (2017). JAGS: Just Another Gibbs Sampler, version 4.3.0.

R Development Core Team. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ridout, M. S., & Linkie, M. (2009). Estimating overlap of daily activity patterns from camera trap data. *Journal of Agricultural, Biological, and Environmental Statistics*, *14*, 322–337.

Rowcliffe, J. M., Carbone, C., Jansen, P. A., Kays, R., & Kranstauber, B. (2011). Quantifying the sensitivity of camera traps: an adapted distance sampling approach. *Methods in Ecology and Evolution*, *2*, 464–476.

Rowcliffe, J. M., Field, J., Turvey, S. T., & Carbone, C. (2008). Estimating animal density using

camera traps without the need for individual recognition. *Journal of Applied Ecology*, *45*, 1228–1236.

Rowcliffe, J. M., Jansen, P. A., Kays, R., Kranstauber, B., & Carbone, C. (2016). Wildlife speed cameras: measuring animal travel speed and day range using camera traps. *Remote Sensing in Ecology and Conservation*, *2*, 84–94.

Rowcliffe, J. M., Kays, R., Kranstauber, B., Carbone, C., & Jansen, P. A. (2014). Quantifying levels of animal activity using camera trap data. *Methods in Ecology and Evolution*, *5*, 1170–1179.

Royle, J. A., Chandler, R. B., Sollman, R., & Gardner, B. (2014). *Spatial Capture-Recapture*. *Spatial Capture-recapture*. Elsevier.

Steenweg, R., Hebblewhite, M., Whittington, J., Lukacs, P., & McKelvey, K. (2018). Sampling scales define occupancy and underlying occupancy–abundance relationships in animals. *Ecology*, *99*, 172–183.

Tobler, M. W., Zúñiga Hartley, A., Carrillo-Percastegui, S. E., & Powell, G. V. N. (2015). Spatiotemporal hierarchical modelling of species richness and occupancy using camera trap data. *Journal of Applied Ecology*, *52*, 413–421.

Wearn, O. R., Bell, T. E., Bolitho, A., Durrant, J., Haysom, J. K., Nijhawan, S., … Rowcliffe, M. J. (2022). Supporting data for 'Estimating animal density for a community of species using information obtained only from camera-traps' [Data set]. *Zenodo*. https://doi.org/10.5281/zenodo.6635649

Wearn, O. R., & Glover-Kapfer, P. (2017). *Camera-trapping for conservation: a guide to best-*

practices. WWF Conservation Technology Series 1(1). WWF-UK, Woking, United Kingdom.

Wearn, O. R., & Glover-Kapfer, P. (2019). Snap happy: Camera traps are an effective sampling tool when compared with alternative methods. *Royal Society Open Science*, *6*, 181748.

Wearn, O. R., Rowcliffe, J. M., Carbone, C., Pfeifer, M., Bernard, H., & Ewers, R. M. (2017). Mammalian species abundance across a gradient of tropical land-use intensity: a hierarchical multi-species modelling approach. *Biological Conservation*, *212*, 162–171.

Figure captions

Figure 1. Day-range estimates from camera-trap data for 35 species of Bornean mammal and bird, in old-growth and logged forest. Species with clear differences in day-range between land-use types (i.e. credible intervals on the estimated differences did not cross zero) are shown with filled ridges, whilst those without clear differences are shown with transparent ridges.

Figure 2. Changes in speed and activity level across land-use type for a community of Bornean mammals and birds. Vectors show the direction of change in the two parameters from old-growth to logged forest. The product of activity level and speed is day-range (the distance travelled in a 24-hour period). Vectors for species showing clear changes in day-range across land-use type (i.e. credible intervals on the estimated change did not cross zero) are labelled with the species common name, whilst vectors for species showing no clear changes are unlabelled and are semi-transparent.

Figure 3. Relationships between (the log of) body mass versus A) activity level and B) the variance in (the log of) movement speed across species, including 95% credible intervals. Estimated activity levels and variances in (log) speed for each species are shown as points.

Figure 4. Estimated camera-trap detection zones in old-growth and logged forest for 35 species of Bornean mammal and bird. Detection zones are defined by two parameters – the effective detection angle and distance – which define the width and depth of detection zones, respectively. Points show the locations of animals when the camera-traps were triggered. Points further than 10 m from camera-traps (< 1% of the data) are here not shown to allow use of the same scale across species.

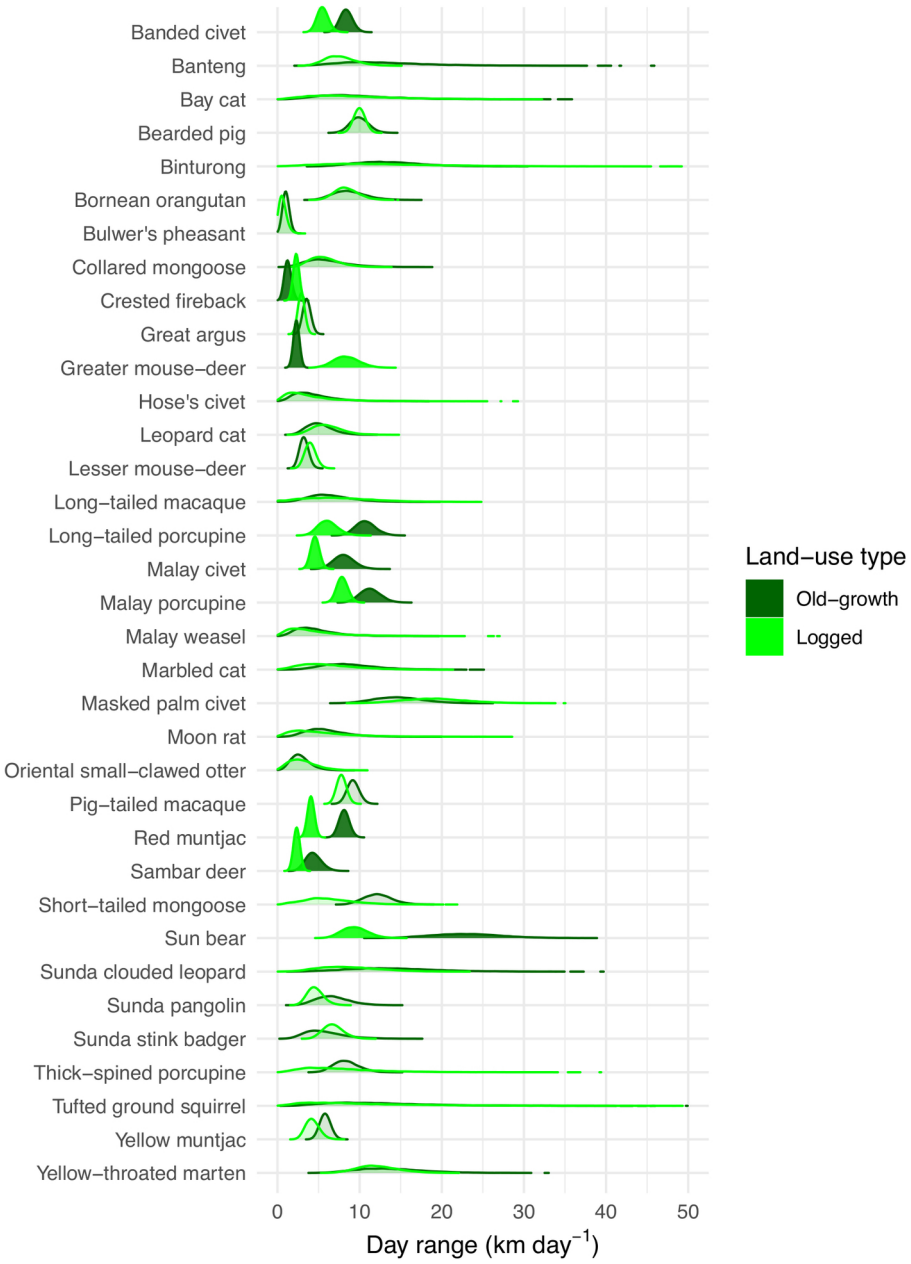
Figure 5. Changes in effective detection angle and radius across land-use type for a community of Bornean mammals and birds. Vectors show the direction of change in the two parameters from old-growth to logged forest. The combination of the radius and angle determine the size of the camera-trap detection zone. Vectors for species showing clear changes in detection zone across land-use type (i.e. credible intervals on the estimated change did not cross zero) are labelled with the species common name, whilst vectors for species showing no clear changes are unlabelled and are semi-transparent.

Figure 6. Relationships between (the log of) species body mass versus A) effective detection angle and B) effective detection distance across species, including 95% credible intervals. Estimated effective detection angles and distances for each species are shown as points.

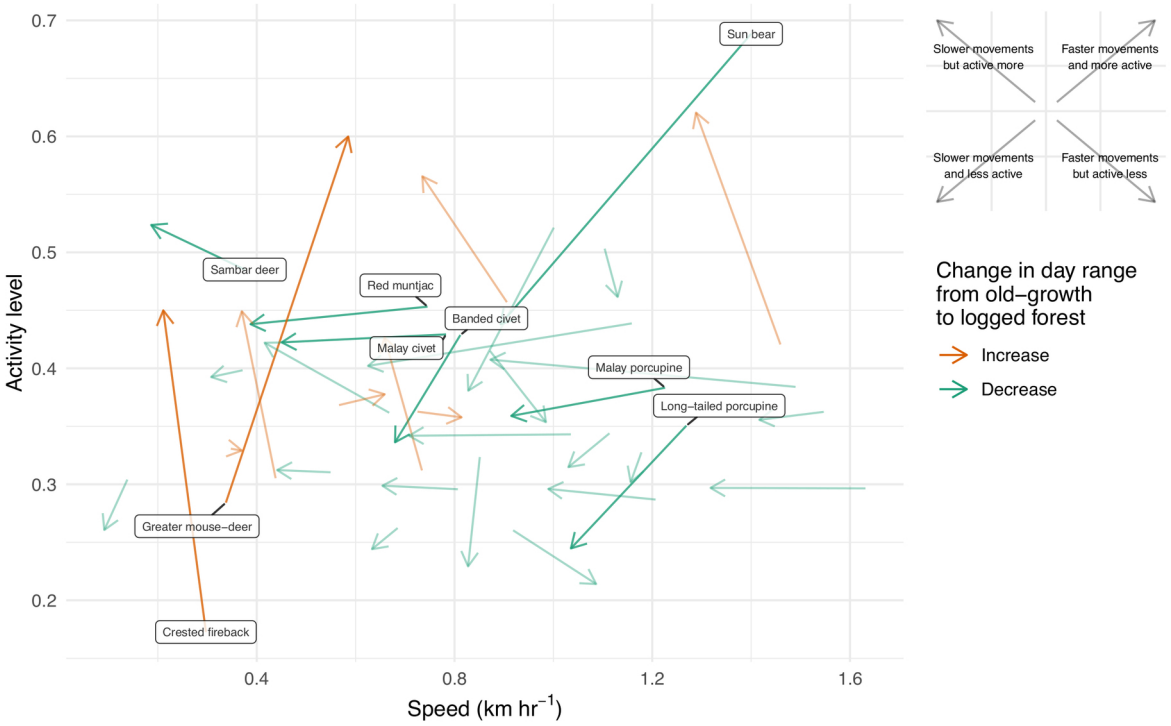
Figure 7. Density estimates from a hierarchical random encounter model for a community of Bornean mammals and birds, in old-growth and logged forest. Species are ordered by their density in old-growth forest, with 'common' species in the top panel and 'rare' species in the

bottom panel. Species with clear differences in density between land-use types (i.e. credible intervals on the estimated differences did not cross zero) are shown with filled ridges, whilst those without clear differences are shown with transparent ridges.

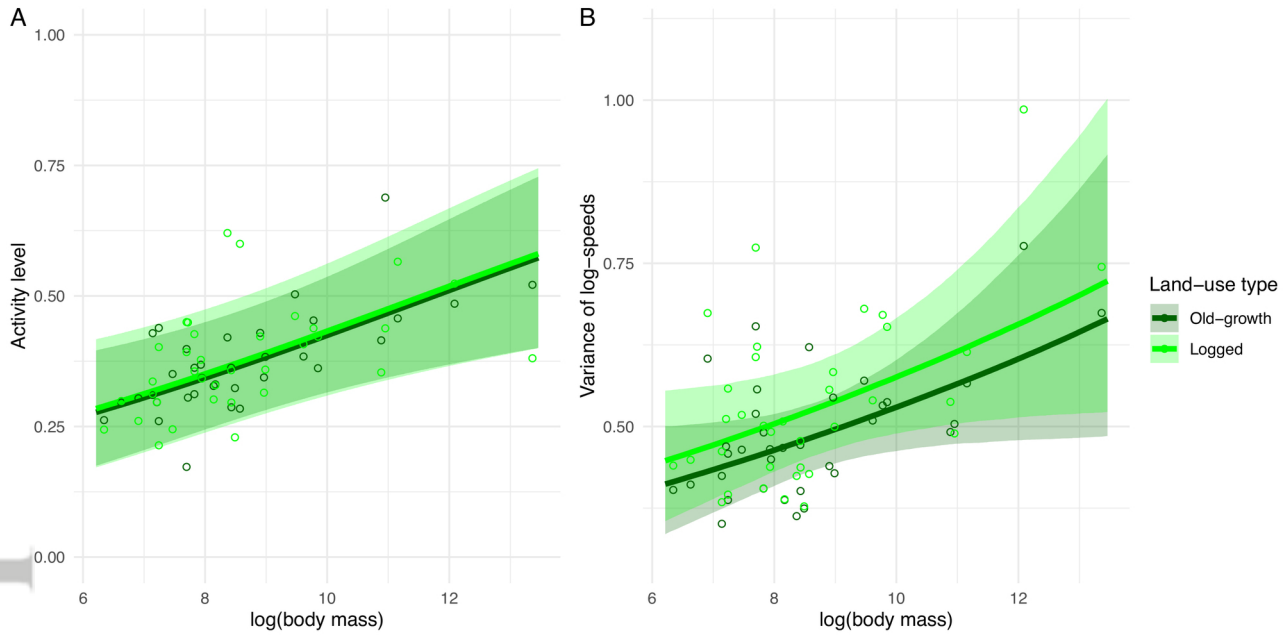
Figure 8. Estimated biomass of Bornean mammals and birds across broad trophic groupings, in old-growth and logged forest. The lower, middle and outer limits of each bar show the lower 95% credible intervals (CIs), medians and upper 95% CIs, respectively. Labels indicate the summed biomass for each trophic group.



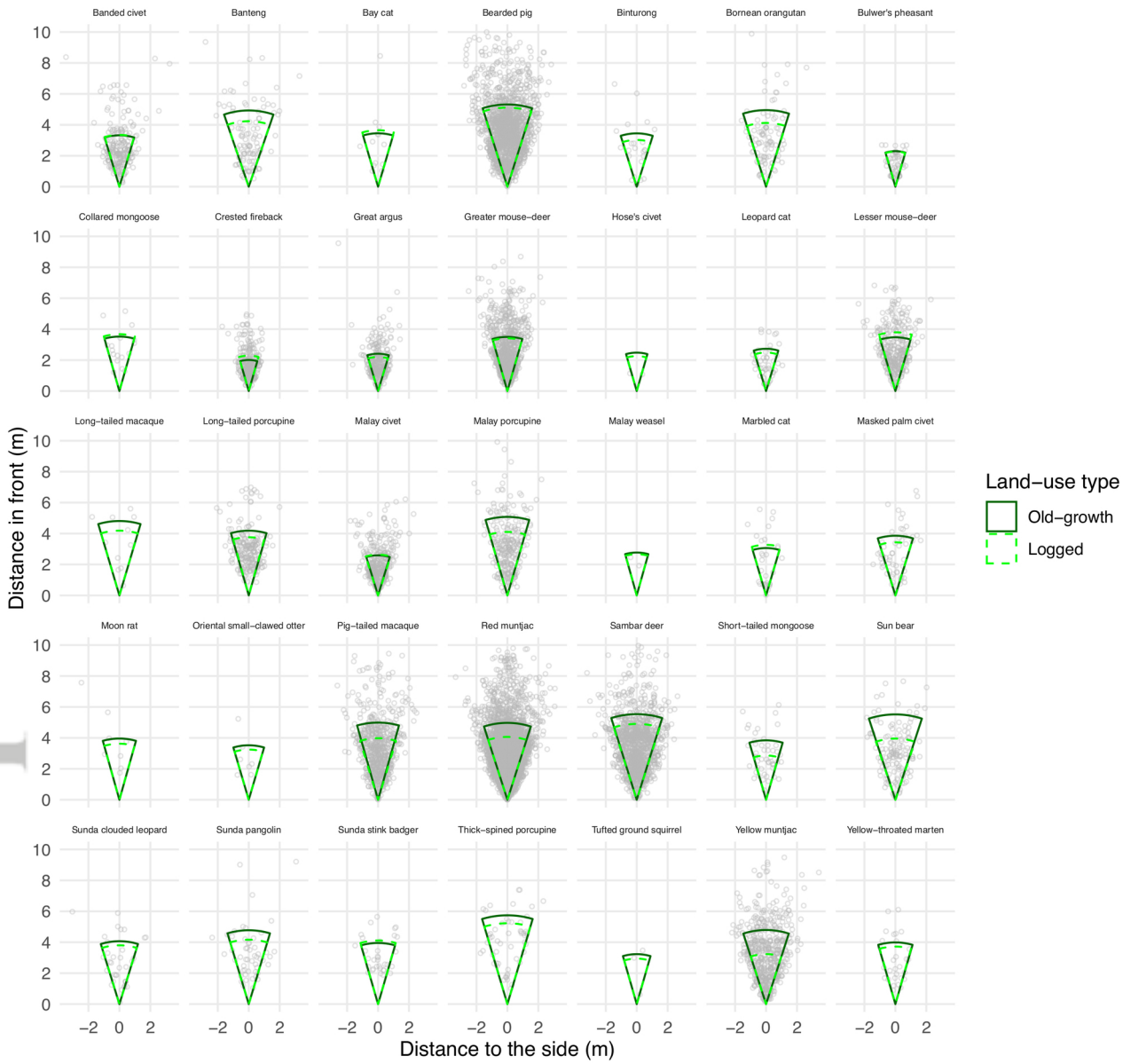
MEE3_13930_posteriors_ridgeplot_dayrange_byspecies(6x8).jpg



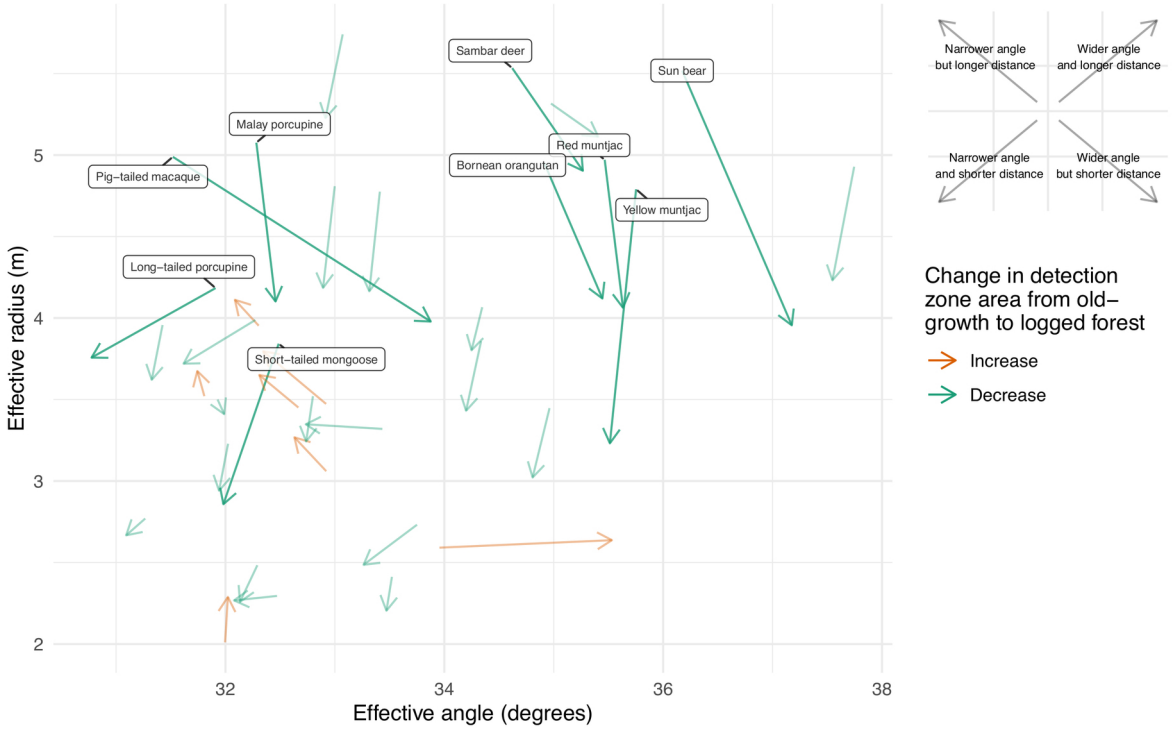
MEE3_13930_arrows_dayranges(9x6).jpg



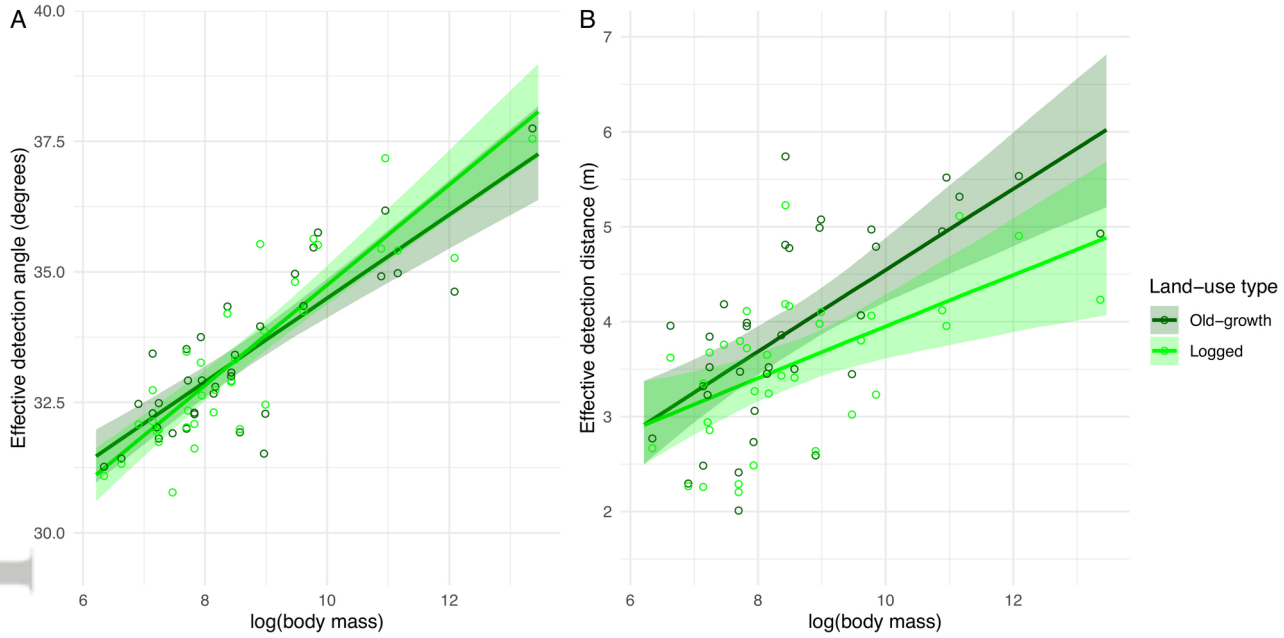
MEE3_13930_bodymasseffect_activity&logspeedvariance(10x5).jpg



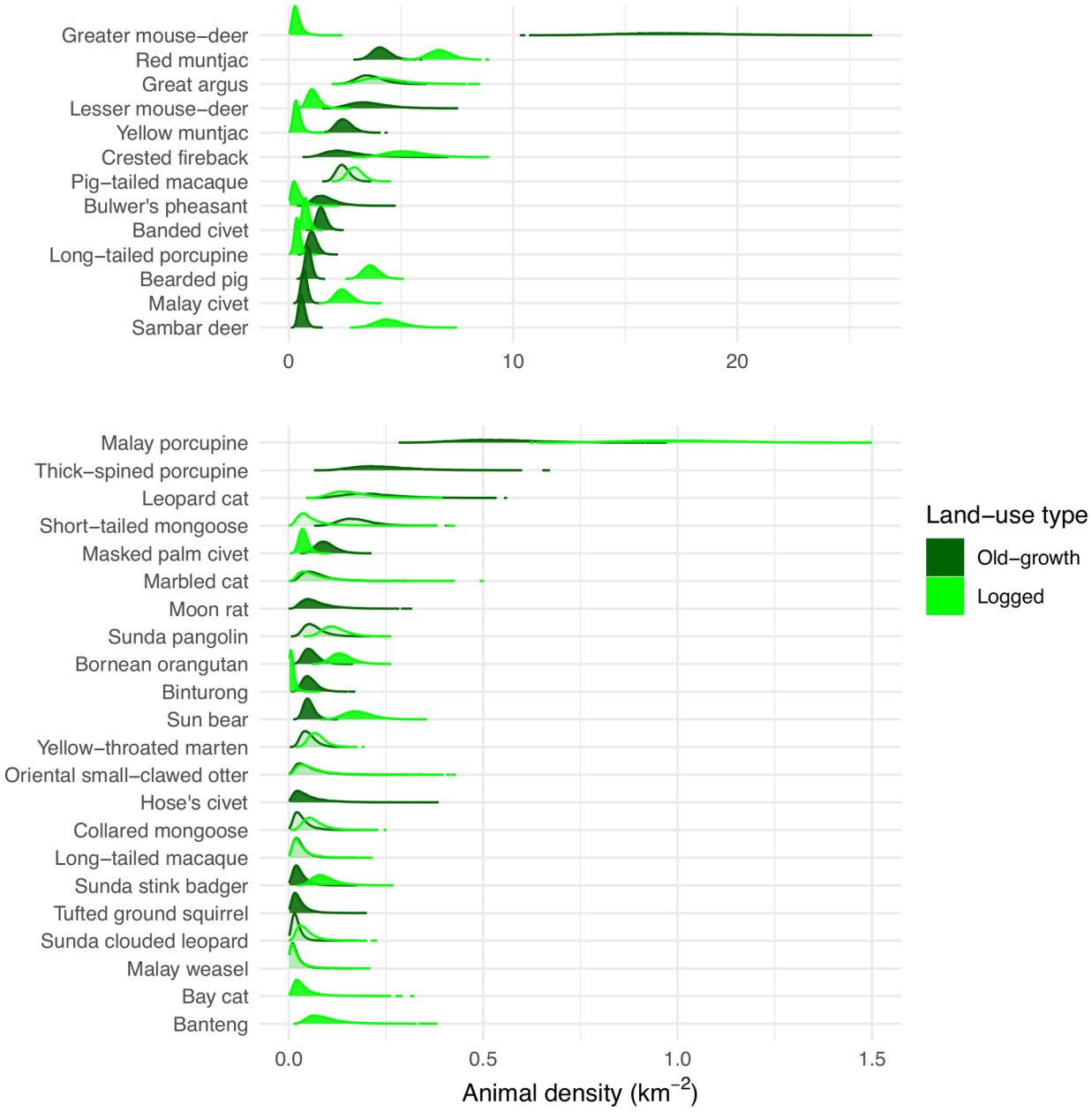
MEE3_13930_detectionzone_byspecies_withpoints(8-5x8-5)_v2.jpg



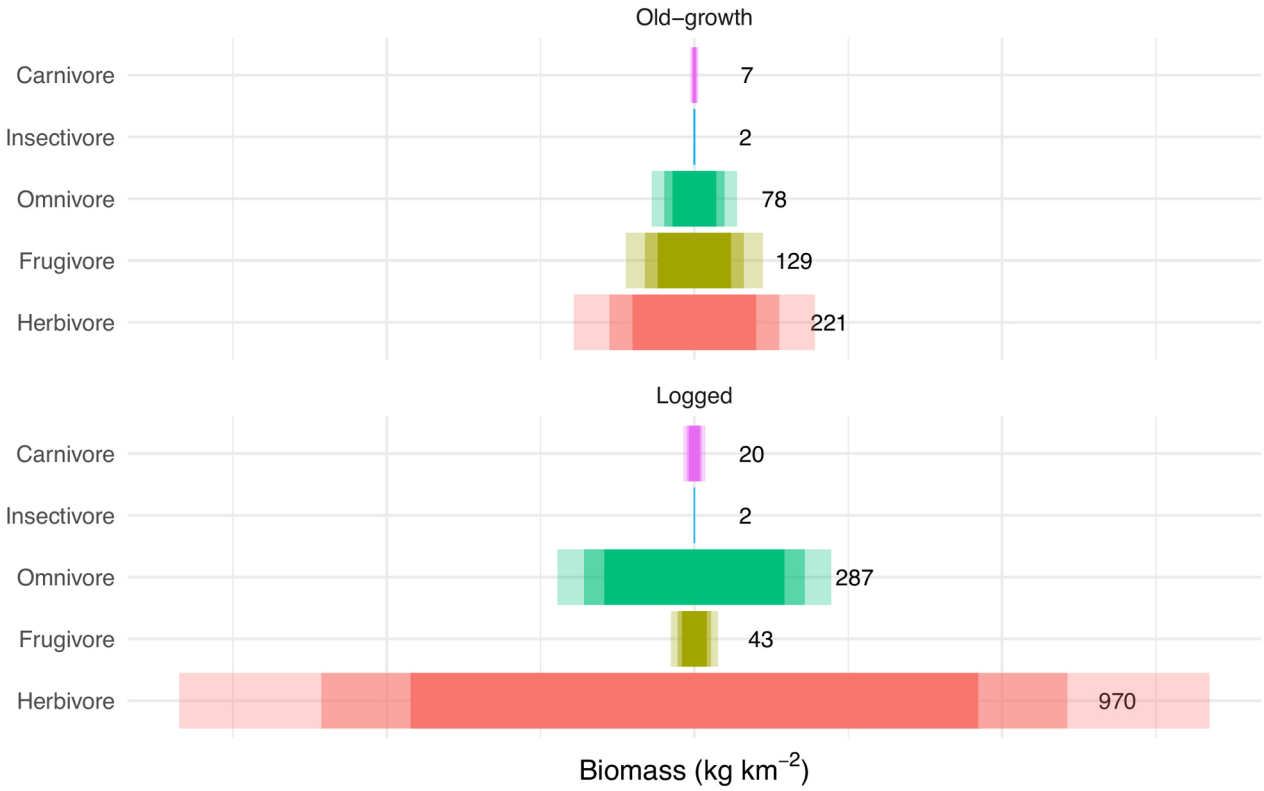
MEE3_13930_arrows_detectionzones(9x6).jpg



MEE3_13930_bodymasseffect_angle&radius(10x5).jpg



MEE3_13930_posteriors_ridgeplot_density_byspecies(7x7).jpg



MEE3_13930_biomasspyramid_withCIs(7x4-5).jpg