Human and environmental associates of local species-specific abundance in a multi-species deer assemblage

Authors: Valentina Zini [a,b]*, Kristin Wäber [a,c], Karen Hornigold [d], Ian Lake [a], Paul M. Dolman [a]

Author Affiliations:

[a] School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK; *E-mail addresses*: v.zini@uea.ac.uk; k.waeber@uea.ac.uk ; p.dolman@uea.ac.uk

[b] Natural Capital Solutions, 1 Lucas Bridge Business Park, 1 Old Greens Norton Road, Towcester, Northamptonshire, NN12 8AX UK

[c] National Trust, Westley Bottom, Westley, Bury Saint Edmunds IP33 3WD

[d] Woodland Trust, Kempton Way, Grantham, Lincolnshire, NG31 6JF

* *Corresponding author*: Valentina Zini; *Postal address*: School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK; *E-mail address*: v.zini@uea.ac.uk

Appendix A: Composition and configuration of forest blocks

Table 1S: Bio-geographic blocks of the Thetford Forest landscape mosaic, for each of 12 showing area (ha), composition (percentages) of: calcareous soil; grassland; mature tree crops (>45 years old); restocked tree crops (0-5 years from planting); and of the block perimeter that abuts arable (% arable perimeter).

Forest Block	Area (ha)	% calcareous	% grassland	% mature	% restock	% arable perimeter
1	1337	22	19	11	2	82
2	918	22	28	17	1	87
3	2401	51	16	3	3	68
4	3025	31	27	36	3	52
5	714	17	19	24	4	60
6	833	2	28	31	1	48
7	545	14	34	16	4	68
8	908	26	28	27	4	63

9	3464	32	13	36	3	39
10	1656	18	16	29	4	12
11	2333	27	17	27	5	60
12	592	18	32	34	2	51

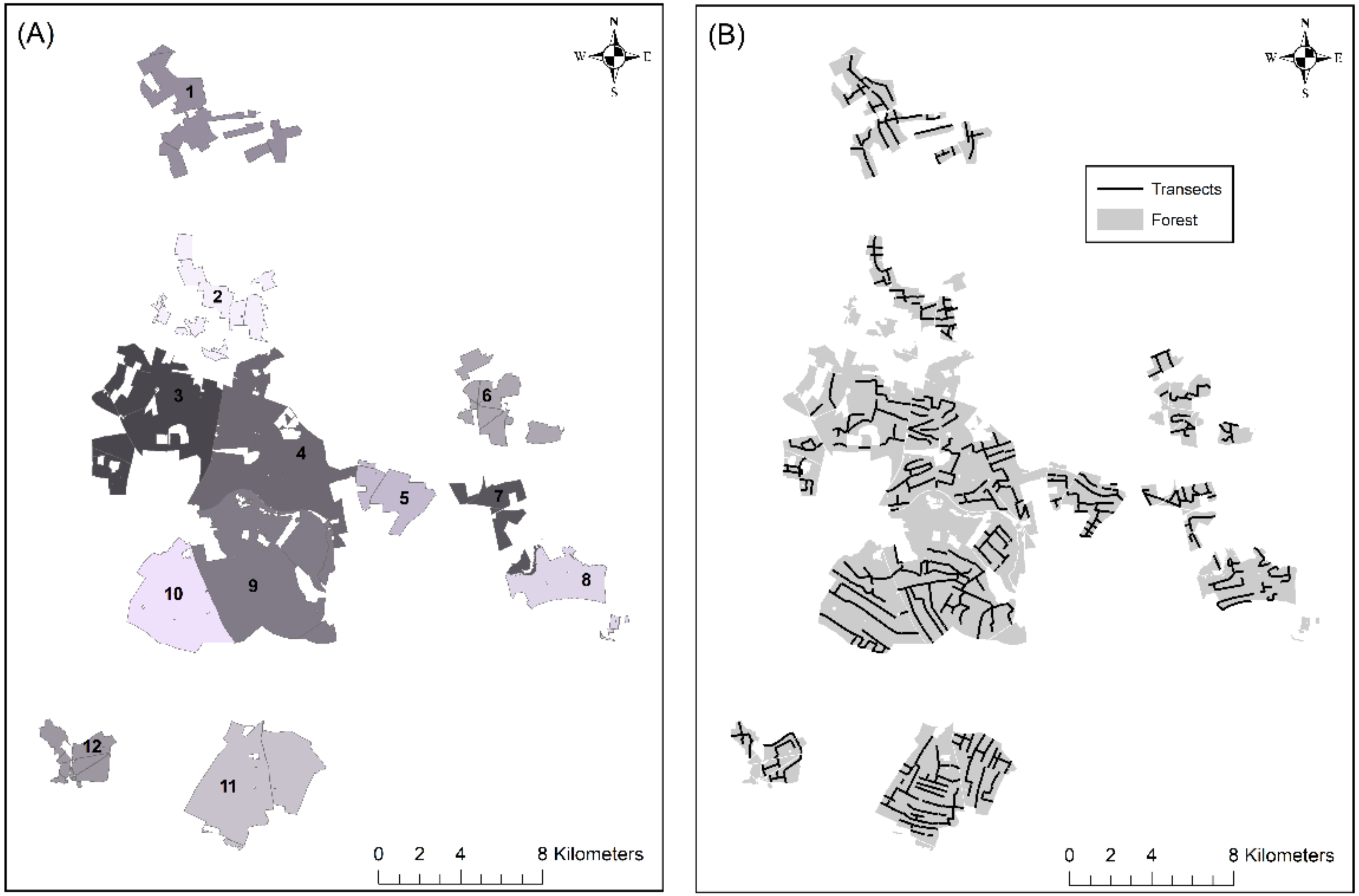

Figure 1S: Thetford Forest landscape mosaic showing: A n=12 bio-geographic blocks detailed in table 1S; B sampling network of nocturnal deer thermal imaging distance sampling transects, each one-sided driven in both directions sampling contiguous contrasting stands.

Appendix B: Detectability class covariate and detection function analysis

Criteria for combined sections

An observers ability to detect an individual deer (‘detectability’) during nocturnal thermal imaging distance sampling, reduces with perpendicular distance of the deer (or group) but also systematically differs between forest growth stages, according to structural closure (Hemami et al. 2007; Wäber and Dolman 2015). Detection functions in models of detectability, therefore, included an *a priori* 3-level categorical variable, at the scale of the stand in each transect section, as either:

‘open’ (comprising permanent open areas, clearfelled unplanted stands, and recently restocked stands 0-3 years since planting);

‘dense’ (comprising restocked stands 4-5 years old, pre-thicket stands 6-10 years old and thicket stands 11-20 years old, and those pole stage stands, between 21-40 years old, that had not yet received a first thinning cut, when rows are removed), or

‘mature’ (comprising thinned pole stage stands 20-40 years old and mature stands >40 years old).

Individual forest stands defined by growth stage and potentially by visibility class, surveyed as contiguous or sequential single-sided driven transects, generating a series of short (n=2414, mean length 225m, SD=128) transect ‘sections’ (with the start and end defined for each individual forest stand). Large numbers of sections with zero observations caused zero-inflation in the density surface model (DSM) and generated inflated high values of local density in those sections with one or more deer observed. To improve model precision and avoid zero-inflation, sections were combined into fewer, longer composite ‘segments’, each comprising a single visibility class, either contiguous or interrupted (interspersed with sections of another detectability class, lying on either side of the transect centre-line) but extending for no more than 600m between start and end point, by an algorithm. In a few instances (n=29) where longer individual sections (max length=944m, mean =700m, SD=102) ran through or alongside a single large stand (of uniform visibility class), these were retained in their entirety despite being >600m in length. Figure 2S shows three examples of this process, whereby the initial 2414 sections were merged and reduced to 1162 segments (mean length=462m, SD=297).

Candidate variables were extracted from a series of buffers (sequentially sampling radii from 100m to 650 for roe and 100m to 700m for muntjac) around each combined segment (segment buffered are referred to as 'localities'); therefore, individual localities sampled similar environmental attributes to neighbouring, overlapping or interspersed localities that differed in detectability class.

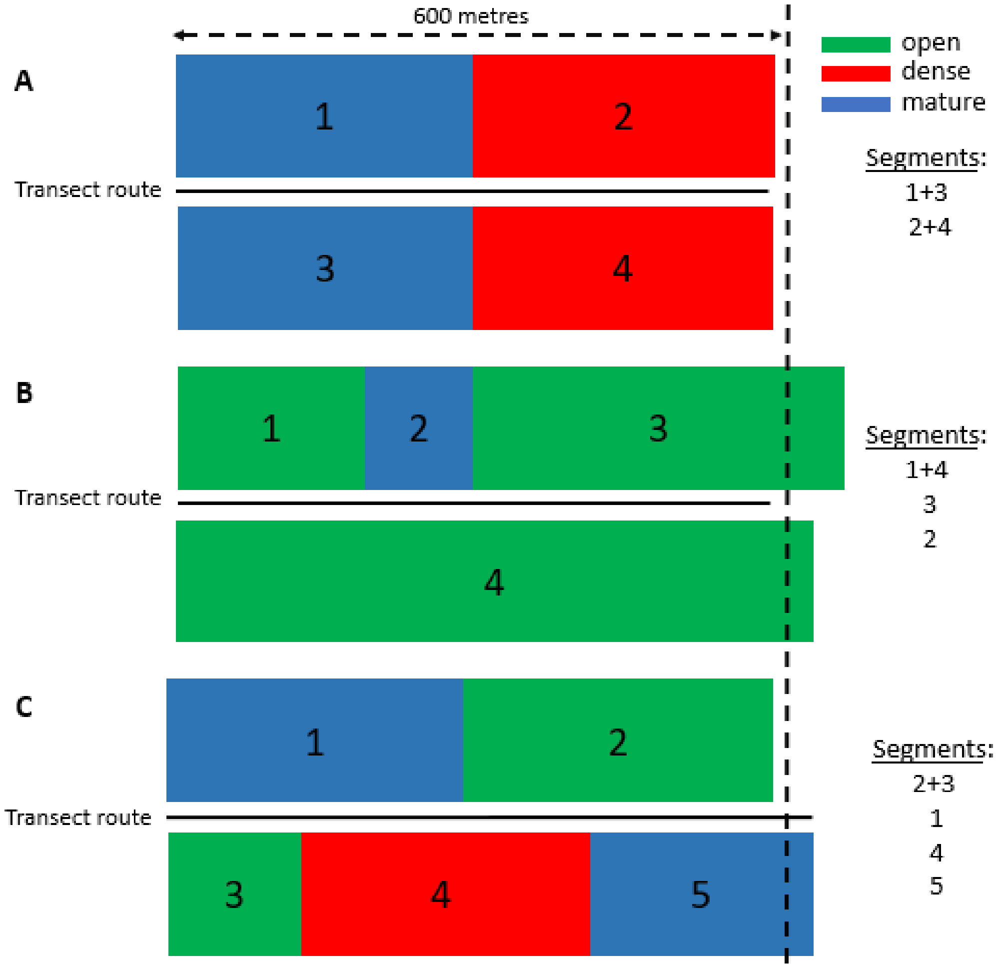

Figure 2S: Schematic illustrating combination of proximal distance transect 'sections' (defined by stands) into fewer, longer, 'segments' of common detectability class (denoted by colours). In A: one segment comprised two single-sided sections that sampled mature stands (1 and 3) on opposite sides of the transect centre-line (spatially linked, but separated by transect timing, first and second pass), and a second segment comprising two sections that each sampled dense stands (2 and 4). In B: one long section (4) exceeded 600m but was not subdivided (as it sampled a single stand), this was combined with a short section (1) on the opposite side of the transect centre-line, that also sampled a stand of the same visibility class (open). Although the remaining section also sampled a stand of this visibility class (open), this section was not combined in the same segment as it projected further beyond section (4);

section (2) sampled a different visibility class (mature) and was not combined with either of these segments (though would be a candidate for combination with other nearby section(s) further along the centre line if they also sampled a mature stand and lay within 600m). In C: we combined two sections (2 and 3) on opposite sides of the transect centre-line as both sampled stands of the same visibility class (open) and lay within 600m, while two sections (1 and 5) that sampled mature stands were not combined as they exceeded 600m from start to finish, section (4) sampled a different visibility class (dense) and was therefore not combined with any of these segments; again sections 1, 5, 4 would be candidates for combination with other nearby sections lying within 600m along the transect centre-line if theye sampled the same visibility class.

Detection function analysis

Species-specific models of detectability, the decreasing probability of seeing and recording an animal with increasing perpendicular distance from the transect line, comparing all available key detection functions (half-normal and hazard-rate, with adjustments: cosine, hermite, polynomial), examining whether detectability was lower for smaller groups at greater distance (group-size covariate), while controlling for forest stand detectability class (categorical covariate, three levels: open, dense, mature). The 5% most distal observation were excluded, following (Buckland et al. 2001); that with lowest value of Akaike's information criterion (AIC) was selected; when alternative models differed by less than 2 units then the most parsimonious model was chosen. Analyses were performed in R (R core Team 2018) using package "Distance" (Miller 2016).

Table 2S: Detection function, covariate and adjustment selection of muntjac and roe deer distance sampling data. For each model AIC and the difference in AIC value relative to the best supported model (ΔAIC) are shown.

Key function	Covariate or adjustment	AIC	ΔAIC
Roe deer			
Hazard-rate	~detectability class	-1670.91	0
Hazard-rate	~detectability class + group-size	-1670.77	0.13801
Hazard-rate	~hermite	-1652.77	18.13949
Hazard-rate	~cosine	-1652.77	18.13949
Hazard-rate	~polynomial	-1652.77	18.13949
Hazard-rate	~1	-1652.77	18.13949
Half-normal	~detectability class	-1647.48	23.42854
Half-normal	~detectability class + group-size	-1646.18	24.72175
Half-normal	~cosine	-1646.08	24.82675
Half-normal	~polynomial	-1639.41	31.49522
Half-normal	~hermite	-1630.22	40.68915
Half-normal	~1	-1629.96	40.9433
Muntjac			
Hazard-rate	~detectability class + group-size	-6887.05	0

Hazard-rate	~detectability class	-6876.29	10.75403
Half-normal	~detectability class + group-size	-6835.96	51.08545
Hazard-rate	~hermite	-6829.77	57.28012
Hazard-rate	~cosine	-6829.77	57.28012
Hazard-rate	~polynomial	-6829.77	57.28012
Hazard-rate	~1	-6829.77	57.28012
Half-normal	~detectability class	-6825.36	61.68807
Half-normal	~cosine	-6824.77	62.27543
Half-normal	~hermite	-6784.71	102.337
Half-normal	~polynomial	-6784.71	102.337
Half-normal	~1	-6784.71	102.337

Appendix C Spatially-explicit model of recreational forest use

A spatially-explicit layer of relative recreational intensity was constructed (Hornigold 2016) calibrated by data from field-based surveys of recreational activity. The spatially-explicit recreational model (Hornigold 2016) related observed recreational disturbance events (DE) per path element, to measures of the path network distance from the lowest impendence access point, path characteristics, and housing density (in a buffer around the lowest impendence access point), controlling for day in week and time of day. A DE was defined as a walker, dog walker or cyclist on a path element. The model was then used to extrapolate predictions of recreational intensity across the pathway network, allowing resampling within each deer distance-transect segment-buffer.

Methods

Visitor surveys

Visitor surveys were collected from April to October of 2007 (number of sampling points = 138), 2008 (n = 174), 2009 (n = 174); with further data collected in 2013 and 2014 to record the numbers of visitors at popular sites (n = 15, as these have greater car park capacity but were omitted from the original surveys). In 2013 and 2014 we re-sampled 11 of the original sampling points to control for potential between year effects. Points were randomly selected throughout the forest pathway network, constrained to be placed at intersections (junctions) of paths allowing data to be collected simultaneously from each converging pathway element (between consecutive junctions). In each year, each point was surveyed 3 times on different days at different times (between 06:00 – 18:00 hrs) to incorporate variability in the number of recreationists. At each visit, the number of DE in each path segment over a period of 1 hour was recorded, and for each DE, the number of recreationists, their activity (dog walking, walking or cycling) and the path element used to approach and leave the intersection were recorded. Additional methodological details can be found in (Dolman, Lake & Bertoncelj 2008; Dolman 2009).

Explanatory variables

Pathway elements were classified, using Forestry England GIS data, as forest roads (median verge width = 6 m, Wäber & Dolman 2015), fire routes (median verge width = 3 m) and narrow tracks (median verge width = 0.6 m, that may be grassy or muddy). Tracks were further classified based on the vegetation cutting regime as cut and uncut. As the forest is intersected by roads, links joining paths across the roads were classified according to the road crossed with links crossing 'A' roads (major arterial roads comprising the top tier of the UK roads classification system) excluded as too dangerous to cross, while links crossing 'B' roads (connecting A roads and smaller roads, comprising the second tier of the UK roads classification) and minor roads were retained for analysis with links crossing 'B' roads considered as a greater barrier than links crossing minor roads).

It was hypothesised that path type in combination with distance to nearest access point, and car capacity at access points may also affect number of DE. Access points were defined as either formal or informal car parks (gateways where fire routes or tracks meet roads). Car capacity at every access point (defined on GIS from the pathway network) was visually estimated from Google Earth following calibration during field visits. As it is unlikely that larger car parks are completely filled, car capacity was square-root transformed prior to modelling.

The larger the residential population in the vicinity of access points, the more disturbance events are recorded (Dolman, Lake & Bertoncelj 2008), thus the number of households (taken from the 2011 census of households; Office for National Statistics 2011 and linked to coordinates using the UK Postcode Directory UKDS, 2013) was initially extracted from a sequence of 15 concentric buffers built around access points (from 0 m to 5,000 m with 500 m radius increments, then from 5,000 m to 10,000 m using wider 1,000 m radius increments). Because of adjoining buffers being co-linear, buffers were merged based on Spearman correlation coefficient > 0.6 producing 9 distance 'bands', used as independent predictor variables to be used in subsequent modelling of DE. If the absolute or relative numbers of households in the vicinity of access points had changed between 2014 and 2018 (when deer distance sampling data were collected) this may have altered the overall number and relative spatial distribution of recreational DE. We therefore obtained data on areas designated for new housing and number of houses per development from Site Allocation Plans of relevant District Councils. Using Google Earth imagery, we verified that none of the proposed housing (within a buffer of 2000 m from access points) had been built in 2018.

Modelling disturbance events

The dependent variables were the number of walkers and the number of dog walkers per pathway element, modelled separately. Cyclists only made up 6% of observations and, as cyclists are more mobile than walkers we considered network distance from the nearest access point was unlikely to predict their frequency, therefore these were excluded from subsequent modelling. For walkers and dog walkers, the frequency of observations on each network element were modelled separately using a Generalised Linear Mixed Effects Model, with Poisson error incorporating point ID to control for spatial autocorrelation.

First, models of walkers and dog walkers were modelled in relation to day and time. Based on a process of model selection and parsimony, the final model incorporated day of week (categorical, Mon (reference level), Tue/Fri, Wed/Thu, Sat, Sun/national holiday), time of day (walkers: 6 (reference level), 7/8/18, 11/14/15/17, 13, 9/10/12/16; dog walkers: 6 (reference level), 8, 9, 11/13/14/16, 12/17, 7/10/15,), month (walkers: (April (reference), May, June-September, October; dog walkers: April (reference), May/June, July/October, August, September)) and a dummy variable indicating school holidays.

Local source population was extracted from the number of households around access points. First, the Closest Facility tool in ArcGIS 10.3 (Copyright © ESRI, USA) was used to find the nearest access point (measured along the path network) to the sampling point. Then, the local source population variables extracted from the 9 bands were added to models of walkers and dog walkers with the best fitting combination of time, day and month. Based on changes in AIC and coefficients similarity, distance bands were further simplified to 3 distance bands (of 0-2000m, 2000-6000m and 6000-10000m) both for walkers and dog walkers models used in subsequent modelling.

A weighted network distance between sampling points and access points was calculated using the Closest Facility tool in ArcGIS 10.3 (Copyright © ESRI, USA). Different combinations of weightings (corresponding to distance 'modified for the effect of landscape and behaviour' Adriaensen *et al.* 2003, 'effective distance') were tested for car park capacity, road crossings and path type while controlling for source population, time, day and month. As with different weightings the lowest impedance access point to a given sampling point can vary, household data were re-assigned accordingly each time. The weighting giving the lowest model AIC was selected.

Results

Weighted network analysis

The best-fitting model considered 'B' roads as barriers, while minor roads were classed as of no impedance for both walkers and dog walkers. Car park capacity differed in their impedance for walkers and dog walkers with car parks with lower capacity being of less impedance to dog walkers. As concerns path types, dogs preferred wider paths (impedance weightings; forest road = 1, fire route = 0.9, track = 2) than walkers (forest road = 1, fire route = 0.5, track = 1). Verges cutting regime didn't affect the impedance of paths.

Models of recreational disturbance

DE of dog walkers and walkers were both higher in weekends and school holidays but differed in time of day with DE of dog walkers being highest at 9:00 while walkers DE being highest at 13:00. A lower car park capacity and a route crossing the tracks instead of a fire route or forest road increased the effective distance for both walkers and dog walkers. Of the three distance bands used to extract number of households only the 2km buffer had a positive significant effect on DE. DE were lower in paths sections further (in terms of effective distance) away from access points.

Predictions

In order to analyse the response of deer densities to recreational activity we, a priori, selected dog walkers as our index of recreational activity as dog walkers are less subject to seasonal variation than walkers and deer have been reported to be more affected by dog walkers than walkers (e.g. Miller, Knight & Miller 2001). Dog walkers were predicted for all combination of time, day and month. All the combinations were created for school holiday and non-school holiday and predictions were multiplied by the ratio of school holiday or non school holidays in each month (to control for relative abundance of DE) and then averaged over spring summer and autumn (1st of April- 31st of October). Our distance sampling thermal imaging data were collected over the winter months, unlike the recreational data, however the spatial pattern of disturbance events is the same across seasons as it is determined by the distribution of access points and housing density.

Appendix D: Selection of buffer radius for deer distance transect segments

Local candidate variables were extracted from a symmetrical (two-sided) buffer around the transect centre-line of each segment (locality, see Appendix B). Buffer radius size were decided based on home-ranges values. Winter home-ranges based on earlier radio-telemetry in our study area (Chapman et al. 1993) measured 75ha for roe deer and 18ha muntjac (averaged across males and females individuals), corresponding to a circular home-range radius of 500 and 250m. For roe deer, a 500m radius only slightly exceeded the landscape grain (mean stand width = 355m, SD=60), and would result in bigger localities compared to the home-ranges (mean segment length = 462m, SD=297; corresponding localities built using a 500m buffer radius of mean area = 113ha SD = 18). We examined a series of full Density-Surface Models (DSMs), that included all candidate variables, geographical coordinates (as smoothing parameters), and the random effect of forest block, n=12) with variables extracted at buffer radii between a minimum of 100m to a maximum 650m (corresponding to the maximum roe deer monthly home-range reported of 136ha, southern Norway, Morellet *et al.* 2013) and examined the most informative radius at which to inform variable extraction. For roe deer, exploratory modelling showed a clear AIC minima at 300m (Fig 3S A, corresponding in localities with mean area = 48ha SD= 11.4), lower than the home-range based on historic VHF telemetry (that might overestimate the home-range, corresponding to a 500m circular home-range), therefore 300m was used in modelling. For muntjac, full DSMs across buffer radii from a minimum of 100m to a maximum of 700m (with 700m corresponding to a 153ha circular home-range, maximum home-range reported in Taiwan McCullough, Pei & Wang 2000), provided an AIC inflexion at 250m (Fig 3S B, corresponding in localities with mean area = 37ha SD= 9) that was used for subsequent modelling.

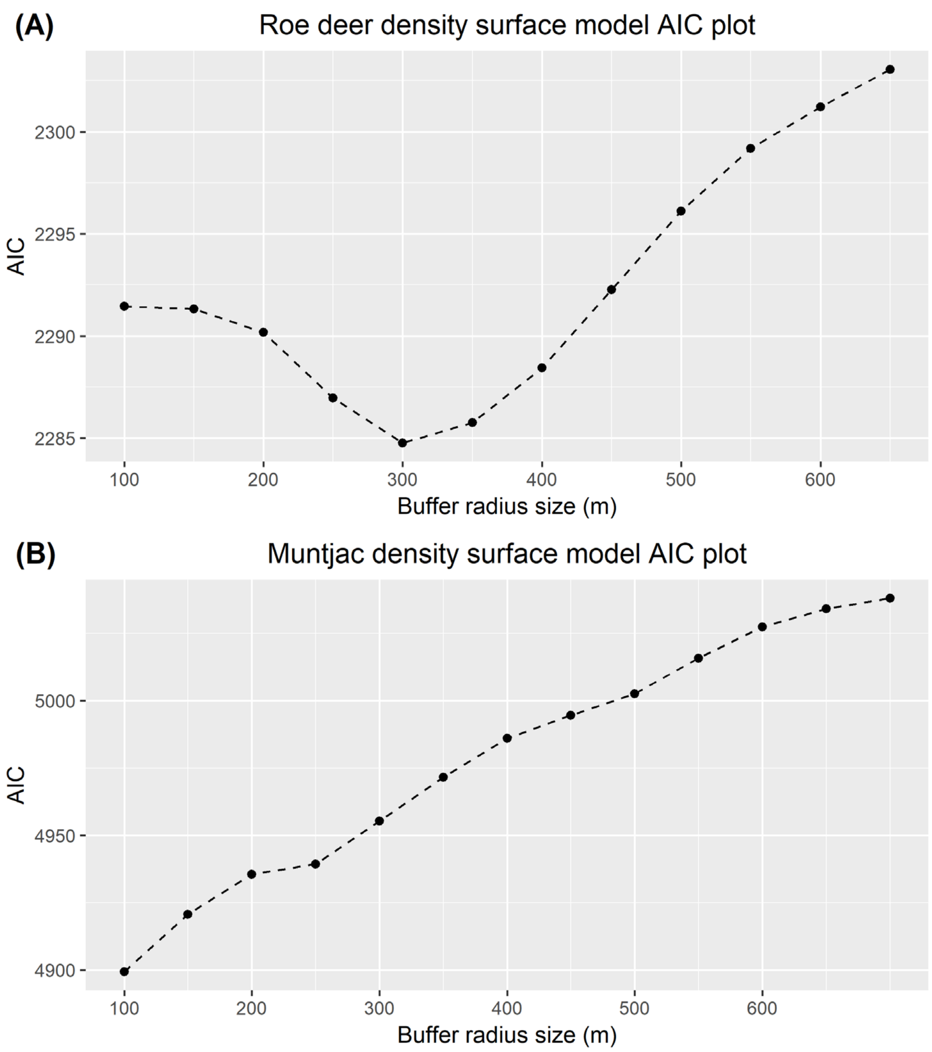

Figure 3S: Selection of buffer radii for which candidate variables were extracted to model local density surfaces (DSMs) of roe deer (A) and muntjac (B). The AIC of a full model (including all candidate variables, geographic coordinates as smoothing parameters and the random effect of forest block) is shown across sequential buffer radii, from 100 to 650 for roe deer and from 100m to 700m for muntjac with increments of 50m, buffering distance transect segments (n=1162, mean length=462m, SD=297).

References

Adriaensen F, Chardon JP, De Blust G, et al (2003) The application of 'least-cost' modelling as a functional landscape model. Landsc Urban Plan 64:233–247. https://doi.org/10.1016/S0169-2046(02)00242-6

Buckland ST, Anderson DR, Thomas L, et al (2001) Line transects. In: Introduction To Distance Sampling: Estimating Abundance of Biological Populations. Oxford University

Press, Oxford, pp 104–14

Chapman NG, Claydon K, Claydon M, et al (1993) Sympatric populations of muntjac (*Muntiacus reevesi*) and roe deer (*Capreolus capreolus*): a comparative analysis of their ranging behaviour, social organization and activity. J Zool 229:623–640. https://doi.org/10.1111/j.1469-7998.1993.tb02660.x

Dolman PM (2009) Woodlark and Nightjar Recreational Disturbance and Nest Predator Study 2008 and 2009 Final Report to Breckland District Council. University of East Anglia, Norwich

Dolman PM, Lake IR, Bertoncelj I (2008) Visitor flow rate and recreational modelling in Breckland. Report to Breckland District Council. Norwich. University of East Anglia

Hemami MR, Watkinson AR, Gill RMA, Dolman PM (2007) Estimating abundance of introduced chinese muntjac /*Muntiacus reevesi* and native roe deer *Capreolus capreolus* using portable thermal imaging equipment. Mamm Rev 37:246–254. https://doi.org/10.1111/j.1365-2907.2007.00110.x

Hornigold K (2016) Modelling nature-based recreation to inform land management. PhD thesis. University of East Anglia

McCullough DR, Pei KCJ, Wang Y (2000) Home range, activity patterns, and habitat relations of Reeves' muntjacs in Taiwan. J Wildl Manage 64:430. https://doi.org/10.2307/3803241

Miller DL (2016) Package 'Distance'

Miller SG, Knight RL, Miller CK (2001) Wildlife responses to pedestrians and dogs. Wildl Soc Bull 29:124–132

Morellet N, Bonenfant C, Börger L, et al (2013) Seasonality, weather and climate affect home range size in roe deer across a wide latitudinal gradient within Europe. J Anim Ecol 82:1326–1339. https://doi.org/10.1111/1365-2656.12105

R core Team (2018) R: A Language and Environment for Statistical Computing

Wäber K, Dolman PM (2015) Deer abundance estimation at landscape-scales in heterogeneous forests. Basic Appl Ecol 16:610–620. https://doi.org/10.1016/j.baae.2015.06.005

UKDS, 2013. Postcode Directory. http://borders.edina.ac.uk/html/pcluts_main.html