A new insight for monitoring ungulates: Density Surface Modelling of roe deer in a Mediterranean habitat

Ana M. Valente[1], Tiago A. Marques[2,3], Carlos Fonseca[1], Rita Tinoco Torres[1]

[1]Departamento de Biologia & CESAM, Universidade de Aveiro, Campus de Santiago, 3810-193 Aveiro, Portugal

[2]Centre for Research into Ecological and Environmental Modelling, The Observatory, University of St Andrews, St Andrews KY16 9LZ, Scotland

[3]Centro de Estatística e Aplicações da Universidade de Lisboa, Faculdade de Ciências da Universidade de Lisboa, Bloco C6 - Piso 4, Campo Grande 1749-016 Lisboa, Portugal

Corresponding author:

Rita T. Torres

rita.torres@ua.pt

Telephone: +351234247138

Acknowledgments

We are grateful to all the people who provided valuable assistance in the field. Likewise, several institutions provided invaluable support: Nature and Forestry Conservation Institute and especially Núcleo Florestal de Bragança. We would like to thank University of Aveiro (Department of Biology) and FCT/MEC for the financial support to CESAM RU (UID/AMB/50017) through national funds and, where applicable, co-financed by the FEDER, within the PT2020 Partnership Agreement. TAM is partially funded by FCT, Fundação para a Ciência e a Tecnologia, Portugal, through the project UID/MAT/00006/2013.

Abstract

Ungulates are especially difficult to monitor and population estimates are challenging to obtain, nevertheless such information is fundamental for effective management. This is particularly important for expanding species such as roe deer (*Capreolus capreolus*), whose populations dramatically increased in number and geographic distribution over the last decades. In an attempt to follow population trends and assess species ecology, important methodological advances were recently achieved by combining line or point sampling with Geographic Information Systems (GIS). In this study, we combined density surface modelling (DSM) with line transect survey to predict roe deer density in northeastern Portugal. This was based on modelling pellet group counts as a function of environmental factors while taking into account the probability of detecting pellets and conversion factors to relate pellet density to animal density. We estimated a global density of 3.01 animals/100 ha (95% CI: 0.37 - 3.51) with a 32.82% CV. Roe deer densities increased with increasing distance to roads as well as with higher percentage of cover areas and decreased with increasing distance to human populations. This recently developed spatial method can be advantageous to predict density over space through the identification of key factors influencing species abundance. Furthermore, surface maps for subset areas will enable to visually depict abundance distribution of wild populations. This will enable the assessment of areas where ungulate impacts should be minimized, allowing an adaptive management through time.

Keywords: *Capreolus capreolus*, Iberian Peninsula, distance sampling, density surface models, GAM

Introduction

Large herbivores are particularly difficult to monitor (Schroeder et al. 2014) and ecologists are continuously searching more robust and precise techniques. Successful strategies for the management of wide-ranging species require reliable information on density and population trends (Marques et al. 2001). To cope with the dramatic expansion of ungulates in Europe and North America over the last decades, effective monitoring programs are pivotal (Rooney 2001; Apollonio et al. 2010). Throughout the last years, significant efforts have been made to improve the methods used for monitoring wild populations (Buckland et al. 2001; Hedley and Buckland 2004; Thomas et al. 2010). Distance sampling (Buckland et al. 2001) is recognised as one of the most robust methods for accounting for uncertain detection (Buckland et al. 2001; Marques et al. 2007) and it has been shown to be a reliable and robust method to estimate deer abundance (Marques et al. 2001; Acevedo et al. 2008; Valente et al. 2014). Basically, distance sampling methodology relies on the search for animals or animal signs from lines or points; for each observation the perpendicular distance from the transect is recorded and a detection function is estimated, enabling abundance and density estimation of the population of interest by accounting for undetected animals (or animals signs). With the fast advance of the spatial analysis techniques, the combination of spatial modelling with Geographic Information Systems (GIS) on population density estimation has been recently developed. This was firstly reviewed by Buckland et al. (2000), Hedley et al. (2004) and Hedley and Buckland (2004) who developed methods for improving abundance estimation of wildlife taking into account the population’s spatial distribution. This has allowed to include heterogeneity in the population spatial distribution while accounting for the probability of detecting the animal or its signs. An important output of such approach is a map with the spatial abundance distribution of a population, which is extremely useful to wildlife managers, particularly when communicating results to non-experts stakeholders (Katsanevakis 2007; Miller et al. 2013a). The recent development of density surface models (DSM) enabled the identification of meaningful ecological variables that can affect animal population's densities (Katsanevakis 2007; Miller et al. 2013a). DSMs offer a robust estimation of abundance (Katsanevakis 2007) and are simple to integrate within the line transect framework of distance sampling. Furthermore such models are less dependent on a random survey design or a uniform habitat coverage and allow the estimation of abundance in sub-areas of interest, through numeric integration under the section of the fitted density surface (Katsanevakis 2007). This spatial methodology can also improve management plans, since it makes possible to identify subtle impacts on species, by estimating spatial redistribution of animals as a result of a particular hazard (Petersen et al. 2011). DSMs are a model-based approach corrected for uncertain detection via a distance sampling framework (Hedley and Buckland 2004; Miller et al. 2013a), being typically implemented via generalized additive models (GAMs) (Hastie and Tibshirani 1990). DSMs have been successfully implemented in a few species, *e.g.* aquatic molluscs (Katsanevakis 2007), marine mammals (Henrys 2005; Burt and Paxton 2006), seabirds (Buckland et al. 2012) and only recently in ungulate species (Schroeder et al. 2014; La Morgia et al. 2015).

The European roe deer (*Capreolus capreolus*) is the most abundant and widespread cervid species in Europe, with an estimated population of 10 million individuals (Apollonio et al. 2010). In Portugal roe deer occurs at low densities (Valente et al. 2014) particularly when compared with central and northern Europe (Apollonio et al. 2010). Following the current European trend, roe deer density is expected to increase considerably in Portugal (Torres et al. 2015). It is therefore timely to implement management strategies that can prevent the potential negative impacts deer can have in the ecosystems, such as traffic car collisions, diseases transmission, impacts on commercial forestry and crop production, conflicts among deer and human populations, amongst others (for a review see Putman et al. 2011).

We combined line transect sampling with spatial analysis to predict the abundance of roe deer in northeastern Portugal. This was achievable taking into account a set of environmental variables relevant to the ecology of roe deer. The chosen variables were human disturbance (distance to the nearest road and distance to the nearest human settlement) which may be considered analogue to predation risk (Hewison et al. 2001; Torres et al. 2011) and availability of cover areas, which is particularly important since roe deer is one of the main prey for Iberian wolf (*Canis lupus signatus*). The abundance predictions were based on the relationship between pellet groups and environmental factors, taking into account the probability of detecting pellets while also using appropriate factors to convert pellet groups abundance into deer abundance. This was done through the collection of distance data regarding pellet groups along line transects covering the whole survey area. We expect that the use of such an approach will improve the accuracy of density and abundance estimates when compared with traditional distance sampling, since it models part of the spatial variability (Hedley et al. 2004).

Indirect methods have already been described in the context of deer populations (Marques et al. 2001; Acevedo et al. 2008; Valente et al. 2014), however they have never been used in conjugation with DSM. Although this type of approach have the main drawback of requiring production and decay rates to convert pellets density in animal’s density (which are not typically easy to obtain – for more details see *Discussion* section), they also provide several advantages since the field work is easy to carry out - it can be performed by park rangers to ensure a continuity of data - and results are unbiased even in woodland areas -

such as our study area, where direct methods are often not feasible or potentially biased (Marques et al. 2001; Scott et al. 2002; Anderson et al. 2012). DSM can be applied to other animals for which pellet group count methods are used to estimate their abundance. Examples include mountain hares (Newey et al. 2003), elephants (Barnes et al. 1995; Olivier et al. 2009) and a number of other large vertebrates (Hill et al. 1997; Acevedo et al. 2008; Carvalho et al. 2013). The methodology is equally applicable to surveys of nests or other signs for which production and decay rates can be estimated, *e.g.* apes are most easily monitored by surveying their nests (Plumptre 2000).

This study aims to (1) use an indirect methodology to model the density surface of roe deer in northeast Portugal; (2) estimate the density and abundance of this species, (3) to relate its density to environmental factors and (4) to compare the results of conventional distance sampling with density surface modelling.

Methods

Study area

The study was carried out in northeast Portugal (Montesinho Natural Park – MNP – and *Serra da Nogueira* – SN) (6º30'-7º12'W, 41 º43'-41 º59'N and 6º50'-6º56'W, 41º38'-41º48'N respectively), part of the European Union's Natura 2000 Network, covering an area of 63,500 ha (Fig. 1). The terrain consists of rolling hills with elevation ranging from 438 to 1,481m. The climate is mainly Mediterranean. The vegetation is diverse, characterized mainly by oak (*Quercus pyrenaica*, *Q. rotundifolia*, *Q. suber*), sweet chestnut (*Castanea sativa*) and maritime pine (*Pinus pinaster*). The shrub vegetation is dominated by heather (*Erica* spp.), gum rockrose (*Cistus ladanifer*) and furze (*Ulex europaeus* and *Ulex minor*). Other mammals present are the Iberian wolf (*Canis lupus signatus*), red fox (*Vulpes vulpes*), wild cat (*Felis silvestris*), wild boar (*Sus scrofa*) and red deer (*Cervus elaphus*), among others. The study area is crossed by some rivers and includes small villages with a low human presence (9.5 people per km^2).

Line transects and field work

The survey area was divided in 3 geographic strata: *Serra de Montesinho* (SM: 24,400 ha), Lombada National Hunting Area (LNHA: 20,800 ha) (both inside MNP) and *Serra da Nogueira* (SN: 18,300 ha) (Fig. 1). This was done to improve the precision of the final density estimate, taking into account a previous study (Valente et al. 2014), which includes a smaller sample of the same study area (without spatial modelling). This was also done for management purposes, since a large variation is expected in densities across strata. However, a common detection function was built pooling the data across the three regions. Transect location and orientation was randomly chosen, ensuring that they were representative of all habitat types in the study area. In total, 65 different transects were considered: 22 transects in SM, 16 in SN and 27 in LNHA. Each transect was 1,000m long: to maximize spatial coverage and to mitigate sampling dependence, sampling plots consisted of 4 100m on effort segments, each separated by 200m off effort segments, resulting in a total of 400m on-effort *per* transect. Later the transects were used to model the detection function and the segments to perform the density surface modelling. Given practical and logistic constraints precluding surveying the entire survey area in a single year, field work was conducted in 2012 and 2013 (2012: January and November; 2013: January, February and October), randomly carried among the three study areas. For modelling the detection function, distance data was pooled across years and regions. The transects were conducted on foot. A handheld Global Positioning System (GPS) unit and a compass were used to follow a straight line. A rope was used to facilitate the progress in a straight line, ensuring the scanning of 1 meter from each side of the line, and guaranteeing accurate distance measurements. The choice of 1 meter width (on each side of the rope) transects was based on Marques et al. (2001), where the use of long (>50 meters) and narrow transects was suggested to ease the search for pellets groups in low deer density areas, as is the case for our study area (Valente et al. 2014). The perpendicular distance from the centre of the group to the transect line was recorded for each pellet group detected. Additionally, three observation level covariates thought to influence detectability of pellets (Marques et al. 2007) were recorded: i) the size of the pellet group (medium, 10 - 40 pellets vs. large, > 40 pellets); ii) dispersion of the group (aggregated vs. scattered); and iii) type of habitat around the pellet group (open vs. closed). To minimize bias we considered only pellet groups with ten or more individual pellets (produced at the same defecation event, identified for similar size, shape, texture and colour). This practice reduces the risk of counting one spread pellet group as two (Marques et al. 2001).

A two-stage approach:

Modelling the detection function

Distance sampling allows uncertain detection of animals/objects (Buckland et al. 2001; 2004). A detection function, g(*x*), is used to model the decrease in detectability with increasing distance, from the observer (Buckland et al. 2001; Miller et al. 2013a). The detection function represents the probability of detecting an object given it is at distance *x* from the transect line. The probability of detection for the covered area is then given by:

$$P = \int_0^w g(x)\pi(x)dx$$

where w is a truncation distance and $\pi(x)$ represents the distribution of available distances, assumed to be uniform by design. Formally, this corresponds to the expected value of the detection function with respect to the available distances. In the first stage we used the *Distance* package (Miller 2014) in R (R Development Core Team 2013) to estimate roe deer density and abundance. The global density (D) estimate is obtained as a weighted average of stratum specific estimates, with stratum's areas as weights, *i.e.*

$$\widehat{D} = \frac{\sum_{i=1}^{3} \widehat{D_l}\, A_i}{\sum_{i=1}^{3} A_i}$$

Three key functions were tested: uniform, half-normal and hazard-rate with the three adjustment terms available (cosine, simple polynomial and hermite polynomial). The unit considered for analysis was 400m. The effect of observation level covariates in pellet group detectability was assessed through Multiple Covariate Distance Sampling (MCDS) analysis (Marques et al. 2007). Detection function choice was based on the Akaike information criterion (AIC, Akaike 1974), aided by visual inspection of the histogram of distance data and goodness-of-fit tests (Burnham et al. 2004). Distance data were right-truncated to remove 5% of the perpendicular distances as recommended by Marques et al. (2001), resulting in a maximum width of 95 cm of effective prospection. Density surface modelling results are based on the most parsimonious detection function obtained in this first stage.

Density surface modelling (DSM)

The second stage was also performed in R (R Development Core Team 2013) using the package *dsm* (Miller et al. 2013b). Modelling of density was implemented at the 100m segment level, totalling 260 segments. Four segment level spatial covariates were collected through ArcMAP (version 10.1) and used to model the density surface of roe deer in our study area: i) geographic coordinates (latitude and longitude); human disturbance variables

ii) distance to the nearest road – dist_road – and iii) distance to the nearest human settlement – dist_hum, and iv) percentage of cover areas (ca_perc: coniferous and deciduous forests). Geographic coordinates and human disturbance variables were collected in the center of the 100m segments. The percentage of cover areas was extracted in a 1.26 km radius around the center of each segment. This represents a home range scale calculated based on home range values for Portugal (Carvalho et al. 2008). We used GIS to build the buffers from the center of the 100m segments. Land cover information was obtained through CORINE Land Cover 2006 (CLC2006).

The count method of Hedley and Buckland (2004) was applied, using the number of pellet groups in each segment as the response variable in the density surface model, according to:

$$\mathrm{E}(n_j) = \hat{p}_j A_j \exp\left[\beta_0 + \sum_k f_k(z_{jk})\right]$$ (Miller et al. 2013a),

where z_{jk} is the value of covariate *k* in segment *j*, f_k represents the smooth function of the spatial covariate *k* and β_0 is an intercept term. A_j is the segment area and $\hat{p}_j$ the detection probability (if this parameter is constant throughout the segments it will simply be replaced by $\hat{p}$). The number of pellets (response variable) for each segment was related to the predictor variables through Generalized Additive Models (GAMs) (Hastie and Tibshirani 1990): a quasipoisson distribution and a logarithmic link function were used. The optimum degree of smoothing was defined through Generalized Cross Validation (GCV) score. By default *dsm* package applies a factor $\gamma = 1.4$ to model the effective degree of freedom in the GCV score to avoid overfitting (Miller et al. 2013b). The choice of the density surface model among the set of candidates was based on the lowest GCV value (Wood 2006), while accounting for the deviance explained by each model as well as the *p-value* of each spatial variable.

Abundance estimation

A prediction grid with 635 square cells of 100ha each was built in ArcMAP (version 10.1). The abundance of roe deer in the study area was estimated as the sum of the estimated abundance in each one of the grid cells, $E[\hat{n}_r]$, $\hat{N} = \sum_r E[\hat{n}_r]$, relying on the spatial model chosen for inference. Based on the predictions inferred by the density surface model, and taking into account the value of each variable in each grid cell, an abundance map for the survey area was drawn in R (R Development Core Team 2013). To estimate the abundance two conversion factors were used: i) the decay rate (*i.e.* number of days a pellet group takes to decompose – a pellet group was only considered to have less than six individual pellets), estimated by Torres et al. (2013) for our study area and species of interest (176 ± 31 days), and ii) the production rate (*i.e.* the number of pellet groups produced by an individual *per* day), calculated in the UK, which was considered to be 20 pellet groups *per* day (Mitchell et al. 1985). These values were embedded in the model through the use of an offset, to convert pellet groups density to animal density, accounting for the variance of the former via a bootstrap procedure and ignoring the non-available variance for the latter (see discussion), allowing a straightforward interpretation of the results. Variance for the abundance estimates of DSM analysis was obtained through the variance propagation method described by Williams et al. (2011). This approach enables a prompt variance estimate for both the global and sub-areas density estimates.

Results

The first stage: Modelling the detection function

Over the 26,000m on effort (SM – 8,800m; LNHA – 10,800m; SN – 6,400m) a total of 365 pellet groups were recorded. The detection function that better fitted the distance data among the set of candidates was the uniform key function with one cosine adjustment term (Fig. 2). As expected, the probability of detecting pellet groups decreased with increasing distance from the line, presenting however a broad shoulder (see discussion) with a surprisingly large number of observations very close to the transect line (Fig. 2). The three detection functions that included observation level covariates in the analysis had less support from the data, thus were discarded for the subsequent analysis (with the three covariates tested – habitat, size and shape with ΔAIC of 2.86, 2.65 and 2.81 respectively). The probability of detection for the chosen detection function was $\hat{p} = 0.623 \pm 0.026$ SE.

The second stage: Density surface modelling

From all the candidate density surface models, two were selected based on their GCV score (dsm 1 and dsm 3) (Table 1). The implementation of two DSM's was deemed necessary to fully exploit the data: a DSM for the analysis of environmental data (DSM without geographical variables – dsm 1 - with dist_hum, dist_road and ca_perc spatial covariates), and a DSM that enables a more robust estimate of abundance through the inclusion of geographical data (DSM with geographical variables – dsm 3 - with dist_hum, ca_perc, latitude and longitude spatial covariates). This division was merely practical, to ensure the identification of potential impacts of environmental variables, that could be hidden by the geographical data (taking into account the increase in explained deviance when these variables were included). Fig. 3 shows the smoothed spatial covariates used in the model without geographical variables, being dist_hum the most important variable in the analysis as revealed by p-values (Table 1).

Abundance estimation and uncertainty analysis

The conventional design based distance sampling density estimate was 3.53 animals *per* 100 ha (95% IC: 2.07 – 4.79), with $\hat{N}$ = 2, 233 animals, and a CV of 24.30% (Table 2). According to the best density surface model (DSM with geographical variables) the abundance of roe deer in our study area was estimated to be $\hat{N}$ =1,909 animals with a density of 3.01 animals *per* 100 ha (95% IC: 0.37 – 3.51) and a CV of 32.82%. In accordance with the DSM with geographical variables chosen for inference the distribution map of roe deer throughout the study area is shown in Fig. 4.

The values of abundance, density, 95% confidence intervals and coefficient of variation (%) of traditional distance sampling and density surface models are shown in Table 2.

Discussion

Wildlife managers and ecologists are continuously searching for accurate and unbiased methods to estimate species abundance, density and distribution. Such demand is particularly difficult for large herbivores (Schroeder et al. 2014) dwelling forested habitats (La Morgia et al. 2015). Density surface models, by combining animal density spatial variation with traditional line transect surveys open new possibilities for this (Schroeder et al. 2014). Estimating densities and relating them to meaningful ecological variables represents a step further on wildlife management. DSM allowed us to assess population ecological requirements through the predicted species distribution. Our results show that roe deer have higher densities in areas further away from roads. Previous authors have described a similar pattern for this species (Hewison et al. 2001; Torres et al. 2012a). Roads are known sources of disturbance and ultimately can lead to direct mortality events. Roe deer tendency to avoid roads may be related to the risk of collision, which can jeopardize individual's survival, as evidenced in red deer (*Cervus elaphus*) (Rowland et al. 2000). Our results evidenced that roe deer densities increase in areas near human settlements. This is contrary to previous studies elsewhere (Hewison et al. 2001; Coulon et al. 2008), but also for our study area (Torres et al. 2012b). Nevertheless, methodological differences might explain these on first sight puzzling differences. Torres et al. (2012b) used presence/absence of roe deer pellet groups as an index of habitat use while we estimate actual density for each grid cell, using additional information and hence potentially more accurate. The increasing density towards human settlements can be explained by rural depopulation in MNP throughout the last years (Afonso 2012), resulting in small villages with very low human density. Furthermore the rural depopulation experienced in MNP leads to land abandonment with consequent plant regeneration that represent new food resources to deer (Vingada et al. 2010). In our study area, higher roe deer densities correspond to areas with higher percentage cover. This hints towards the importance of these areas, particularly for a prey with a hiding strategy. Some studies (Mysterud and Østbye 1999) suggest that canopy cover functions as part of an anti-predator strategy, providing hiding places and reduced scent spreading, hence reducing detection by Iberian wolf.

Effectively, as noticed by Katsanevakis (2007) (with *Pinna nobillis*) density surface modelling - contrarily to the non-spatial conventional distance sampling - provided insights into ecological patterns that may be the first step to further studies regarding the studied species. In general, the underlying ecological assumptions of the density surface models, as well as the surface map predicted, fits the data observed during the field survey and previous studies (Torres et al. 2011; Valente et al. 2014). The survey was conducted over a two year period. Therefore, the estimated density, represents the average density over the corresponding time period. The detection function presented a broad shoulder and the expected decline with distance. With objects of interest like pellets, the main distance sampling assumptions naturally hold. Our only concern related to the surprisingly large number of very small distances, which could be due to some specific form of measurement error. Reassuringly, the estimated detection function appears to be fairly insensitive to these detections, largely due to the otherwise broad shoulder present. Regarding the CV of the chosen DSM, it showed an acceptable value, ensuring the predictive power of the survey method. The predictive power was boosted through the addition of geographical coordinates, which increased considerably the deviance explained by the spatial variables. The increased predictive power of the models allows the detections of trends in wild populations with less field data, which contributes to the feasibility of the methodology (La Morgia et al. 2015). Contrarily to what was a priori expected, due to accounting for part of the spatial variability, as suggested by Katsanevakis (2007), the inclusion of the spatial variables in the DSM did not decrease the variance of the estimate. Effectively, this has occurred in several studies considering DSMs (Cañadas and Hammond 2006; Katsanevakis 2007; Schroeder et al. 2014), suggesting that other spatial variables might have been helpful to explain spatial variation in our study area. This deserves further consideration in future studies, since it could potentially lead to more precise estimates. We should note that bias in density estimates will arise if the conversion factors considered (decay rate and production rate) are not valid for our survey. It is expected minimal bias from the decay rate since it was available from our survey region and species (Torres et al. 2013). Since decay can vary across habitats, the use of a site-specific value for each dominant habitat instead of a mean value could be assessed in future work. In fact, due to logistical constrains it was not possible to use the specific value in this work. Nevertheless, we do not believe that was a major limitation in our study. The key problem with our estimate is the use of a production rate obtained in the UK over 30 years ago (Mitchell et al. 1985). Furthermore, the value used does not have corresponding precision measures, which means that the reported density estimate variance ignores a potential source of variation. However, a clear advantage of the modular form of the estimator used is that, as soon as a production rate and corresponding standard error are obtained for our region, the density estimates could be easily updated. Obtaining such production rate should be a major goal for the effective management of these populations (Valente et al. 2014).

Moreover DSM results need to be carefully interpreted since GAMs model selection is still a research area under development (Williams et al. 2011; Miller 2014). Effectively, other indicators should be investigated during distance data spatial modelling (*e.g.* p-values associated with covariate coefficients). In our analysis, the p-value of the variables revealed the inexistence of a significant ecological variable ($p \leq 0.05$) for DSM's with geographical variables. Furthermore, the deviance explained in both models (dsm 1 with 7.17% and dsm 3 with 17.3%) was not satisfactory. These values lie far beneath other studies applying DSM (Cañadas and Hammond 2006; Katsanevakis 2007; Schroeder et al. 2014 with 48.7, 33.5 and 50.4 % respectively). This suggests future investigation of additional factors potentially influencing roe deer densities in our study area. Although slope is not heavily pronounced on our study area the influence of altitude/elevation on abundance distribution must be assessed in future works. Furthermore, as mentioned earlier, the interaction with the sympatric red deer (Torres et al. 2014) or with its main predator, the Iberian wolf would grant these species density to be a suitable predictor variable for roe deer. Additionally, an analysis incorporating sex and season should be assessed in the future, since differences in male and female roe deer ecological requirements, and differences in resource availability throughout the year as shown for other deer species (Thirgood 1995) and as seen by Schroeder et al. (2014) with *Lama guanicoe*, whose abundance showed a peak in summer, might be expected. These goals must be achieved with direct methodologies, which should be linked to DSM in a near future for ungulate populations in Iberian Peninsula.

We believe that the approach presented here could be easily applied in other studies, namely assessing interspecific sympatric relations using one species density as a spatial variable for the other. This paper presents a major advance due to the use of a promising methodology applied to an indirect approach widely used for ungulate populations. The use of these indirect methodologies enable the survey of large forested areas, enabling as well predictions for adjacent areas where there are no relevant differences. Actually, due to its simplicity, the field work can be carried out by park rangers ensuring the continuity of data collection. Furthermore, for an elusive species as roe deer, indirect methodologies potentially present more reliable results, since it is easier to fulfil all distance sampling assumptions. Data analysis is rather more complex, with results that however outweigh this drawback. Furthermore the graphic output of this methodology enables the non-experts to easily interpret the results through the abundance distribution maps. This will ease considerably the access to scientific information essential to management plans particularly useful for expanding species. This work is part of a continued long-term monitoring program and represents a step further in methodological optimization of recently developed distance sampling techniques, which aims to become the future in population size estimation and ecological assessment.

Funding

This study was co-supported by University of Aveiro (Department of Biology) and FCT/MEC financially supporting CESAM RU (UID/AMB/50017) through national funds and, where applicable, co-financed by the FEDER, within the PT2020 Partnership Agreement. TAM is partially funded by FCT, Fundação para a Ciência e a Tecnologia, Portugal, through the project UID/MAT/00006/2013.

Conflict of interest

The authors declare that they have no conflict of interest.

Compliance with Ethical Standards

This article does not contain any studies with human participants or animals performed by any of the authors.

References

Acevedo P, Ruiz-Fons F, Vicente J, Reyes-García AR, Alzaga V, Gortázar C (2008) Estimating red deer abundance in a wide range of management situations in Mediterranean habitats. J Zool 276(1):37–47

Afonso R (2012) O Parque Natural de Montesinho e a promoção do desenvolvimento local. MsC dissertation in Urban and Regional Planning, Aveiro University, Aveiro, Portugal

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Automatic Control 19:716–723

Anderson CK, Nielsen CK, Hester CM, Hubbard RD, Stroud JK, Schauber EM (2012) Comparison of indirect and direct methods of distance sampling for estimating density of white-tailed deer. Wildlife Soc B 37(1):146-154

Apollonio M, Andersen R, Putman R (2010) European ungulates and their management in the 21st century. Cambridge University Press, Cambridge

Barnes RFW, Blom A, Alers MPT, Barnes KL (1995) An estimate of the numbers of forest elephants in Gabon. J Trop Ecol 11(1):27-37

Borkowski J, Ukalska J (2008) Winter habitat use by red and roe deer in pine-dominated forest. For Ecol Manage 255:468-475

Buckland ST, Goudie IBJ, Borchers DL (2000) Wildlife population assessment: past developments and future directions. Biometrics 56(1):1-12

Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL, Thomas LT (2001) Introduction to Distance Sampling. Oxford University Press, Oxford

Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL, Thomas L (2004) Advanced Distance Sampling: Estimating Abundance of Wildlife Populations. Chapman & Hall, London

Buckland ST, Burt ML, Rexstad EA, Mellor M, Williams AE, Woodward R (2012) Aerial surveys of seabirds: the advent of digital methods. J Appl Ecol 49:960–967

Burnham KP, Buckland ST, Laake JL, Borchers DL, Marques TA, Bishop JRB, Thomas L (2004) Further topics in distance sampling Advanced Distance Sampling. University Press, Oxford

Burt ML, Paxton CGM (2006) Review of density surface modelling applied to JARPA survey data. 1–5

Cañadas A, Hammond PS (2006) Model-based abundance estimates for bottlenose dolphins off southern Spain: implications for conservation and management. J Cetacean Res Manag 8(1):13

Carvalho P, Nogueira AJA, Soares AMVM, Fonseca C (2008) Ranging behaviour of translocated roe deer in a Mediterranean habitat: seasonal and altitudinal influences on home range size and patterns of range use. Mammalia 72:89-94

Carvalho JS, Marques TA, Vicente L (2013) Population status of *Pan troglodytes* verus in Lagoas de Cufada Natural Park, Guinea-Bissau. PLoS ONE 8(8):e71527

Coulon A, Morellet N, Goulard M, Cargnelutti B, Angibault JM, Hewison AM (2008) Inferring the effects of landscape structure on roe deer (*Capreolus capreolus*) movements using a step selection function. Landsc Ecol 23(5):603-614

Hastie TJ, Tibshirani RJ (1990) Generalized Additive Models. Chapman and Hall, London

Hedley SL, Buckland ST (2004) Spatial models for line transect sampling. J Agric Biol Environ Stat 9:181–199

Hedley SL, Buckland ST, Borchers DL (2004) Spatial distance sampling models. In: Advanced distance sampling: estimating abundance of biological populations. Oxford University Press, New York, pp 67-89

Henrys BP (2005) Spatial distance sampling modeling of cetaceans observed from platforms of opportunity. MSc Dissertation. University of St. Andrews, Scotland

Hewison AJ, Vincent JP, Joachim J, Angibault JM, Cargnelutti B, Cibien C (2001) The effects of woodland fragmentation and human activity on roe deer distribution in agricultural landscapes. Can J Zool 79(4):679–689

Hill K, Padwe J, Bejyvagi C, Bepurangi A, Jakugi F, Tykuarangi R, Tykuarangi T (1997) Impact of hunting on large vertebrates in the Mbaracayu Reserve, Paraguay. Conserv Biol 11(6):1339-1353

Katsanevakis S (2007) Density surface modelling with line transect sampling as a tool for abundance estimation of marine benthic species: The *Pinna nobilis* example in a marine lake. Mar Biol 152:77–85

La Morgia V, Calmanti R, Calabrese A, Focardi S (2015) Cost-effective nocturnal distance sampling for landscape monitoring of ungulate populations. Eur J Wildl Res 61(2):285-298

Marques FFC, Buckland ST, Goffin D, Dixon CE, Borchers DL, Mayle BA, Peace AJ (2001) Estimating deer abundance from line transect surveys of dung: sika deer in southern Scotland. J Appl Ecol 38:349–363

Marques TA, Thomas L, Fancy SG, Buckland ST (2007) Improving estimates of bird density using multiple-covariate distance sampling. Auk 124:1229–1243

Mayle BA, Peace AJ, Gill RM (1999) How many deer? A field guide to estimating deer population size. Forestry Comission Publications, Edinburgh

Miller DL, Burt ML, Rexstad EA, Thomas L (2013a) Spatial models for distance sampling data: recent developments and future directions. Methods Ecol Evol 4(11):1001–1010

Miller DL, Rexstad EA, Burt ML, Bravingtion MV, Hedley S (2013b) dsm: Density surface modelling of distance sampling data. R package version 2.1.3. URL: http://CRAN.R-project.org/package=dsm. Accessed 25 November 2015

Miller DL (2014) Distance: A simple way to fit detection functions to distance sampling data and calculate abundance/density for biological populations. R package version 0.8.0. http://CRAN.R-project.org/package=Distance. Accessed 25 November 2015

Mitchell B, Rowe JJ, Ratcliffe P, Hinge M (1985) Defecation frequency in roe deer (*Capreolus capreolus*) in relation to the accumulation rates of faecal deposits. J Zool 207(1):1-7

Mysterud A, Østbye E (1999) Cover as a habitat element for temperate ungulates: effects on habitat selection and demography. Wildl Soc Bull 27(2):385-394

Newey S, Bell M, Enthoven S, Thirgood S (2003) Can distance sampling and dung plots be used to assess the density of mountain hares *Lepus timidus*? Wildlife Biol 3:185–192

Olivier PI, Ferreira SM, Van Aarde RJ (2009) Dung survey bias and elephant population estimates in southern Mozambique. Afr J Ecol 47(2):202-213

Plumptre AJ (2000) Monitoring mammal populations with line transect techniques in African forests. J Appl Ecol 37:356–368

Petersen IK, Mackenzie M, Rexstad EA, Wisz MS, Fox AD (2011) Comparing pre- and post-construction distributions of long-tailed ducks *Clangula hyemalis* in and around the Nysted offshore wind farm, Denmark : a quasi-designed experiment accounting for imperfect detection, local surface features and autocorrelation. Technical Report 2011–1, CREEM Technical Report, no. 2011-1, University of St Andrews, St Andrews

Putman R, Watson P, Langbein J (2011) Assessing deer densities and impacts at the appropriate level for management: a review of methodologies for use beyond the site scale. Mamm Rev 41(3):197–219

R Development Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing. http://www.R-project.org. Accessed 25 November 2015

Rooney TP (2001) Deer impacts on forest ecosystems: a North American perspective. Forestry 74(3):201–208

Rowland MM, Wisdom MJ, Johnson BK, Kie JG (2000) Elk distribution and modeling in relation to roads. J Wildl Manage 64(3):672-684

Schroeder NM, Matteucci SD, Moreno PG, Gregorio P, Ovejero R, Taraborelli P, Carmanchahi PD (2014) Spatial and seasonal dynamic of abundance and distribution of guanaco and livestock: insights from using density surface and null models. PloS ONE 9(1):e85960

Scott D, Bacon P, Irvine J (2002) Management of deer in woodlands - Literature reviews of woodland design, and techniques for assessing populations and damage. Report to the Deer Commission for Scotland, Centre for Ecology and Hydrology.

Thirgood SJ (1995) The effects of sex, season and habitat availability on patterns of habitat use by fallow deer (*Dama dama*). J Zool 235(4):645-659

Thomas L, Buckland ST, Rexstad EA, Laake JL, Strindberg S, Hedley SL, Burnham KP (2010) Distance software: design and analysis of distance sampling surveys for estimating population size. J Appl Ecol 47(1):5-14

Torres RT, Santos JP, Linnell JDC, Virgós E, Fonseca C (2011) Factors affecting roe deer occurrence in a Mediterranean landscape, Northeastern Portugal. Mamm Biol 76:491–497

Torres RT, Virgós E, Santos J, Linnell JD, Fonseca C (2012a) Habitat use by sympatric red and roe deer in a Mediterranean ecosystem. Anim Biol 62:351-366

Torres RT, Virgós E, Panzacchi M, Linnell JDC, Fonseca C (2012b) Life at the edge: Roe deer occurrence at the opposite ends of their geographical distribution, Norway and Portugal. Mamm Biol 77:140–146

Torres RT, Santos JP, Fonseca C (2013) Persistence of roe (*Capreolus capreolus*) and red (*Cervus elaphus*) deer pellet-groups in a Mediterrenean mosaic landscape. Mamm Biol 9:7–18

Torres RT, Santos JP, Fonseca C (2014) Factors influencing red deer occurrence at the southern edge of their range: A Mediterranean ecosystem. Mamm Biol 79(1):52-57

Torres RT, Miranda J, Carvalho J, Fonseca C (2015) Expansion and current status of roe deer (*Capreolus capreolus*) at the edge of its distribution. Ann Zool Fennici 52:1-14

Valente AM, Fonseca C, Marques TA, Santos JP, Rodrigues R, Torres RT (2014) Living on the edge: roe deer (*Capreolus capreolus*) density in the margins of its geographical range. PloS ONE 9(2):e88459

Vingada J, Fonseca C, Cancela J, Ferreira J, Eira C (2010) Ungulates and their management in Portugal. European Ungulates and their management in the 21st century. Cambridge University Press, Cambridge, pp 392–418

Williams ROB, Hedley SL, Branch TA, Bravington MV, Zerbini AN, Findlay KP (2011) Chilean blue whales as a case study to illustrate methods to estimate abundance and evaluate conservation status of rare species. Conserv Biol 25(3):526-535

Wood SN (2006) Generalized Additive Models: An introduction with R. Chapman and Hall/CRC, Florida

Fig. 1 Location of the study area in the Iberian Peninsula with transects location and prediction grid in the survey area (**SN** – *Serra da Nogueira*; **SM** – *Serra de Montesinho*; **LNHA** – Lombada National Hunting Area).

Fig. 2 Histogram of distance data of uniform detection function with cosine adjustment term. Observed distances were right-truncated to eliminate the largest 5% of the distances. The detection function was fitted to continuous data, not binned data, and hence the histogram bars cannot be interpreted as probabilities.

Fig. 3 Shape of the functional forms of smoothed spatial covariates with the DSM without geographical variables – **(a) dist_hum** representing the distance to the nearest human settlement; **(b) dist_road** representing the distance to the nearest road and **(c) ca_perc** representing the percentage of cover areas (coniferous and deciduous forests).

Fig. 4 Abundance distribution map of roe deer throughout our study area based on the DSM with geographical variables chosen for inference (dsm 3).

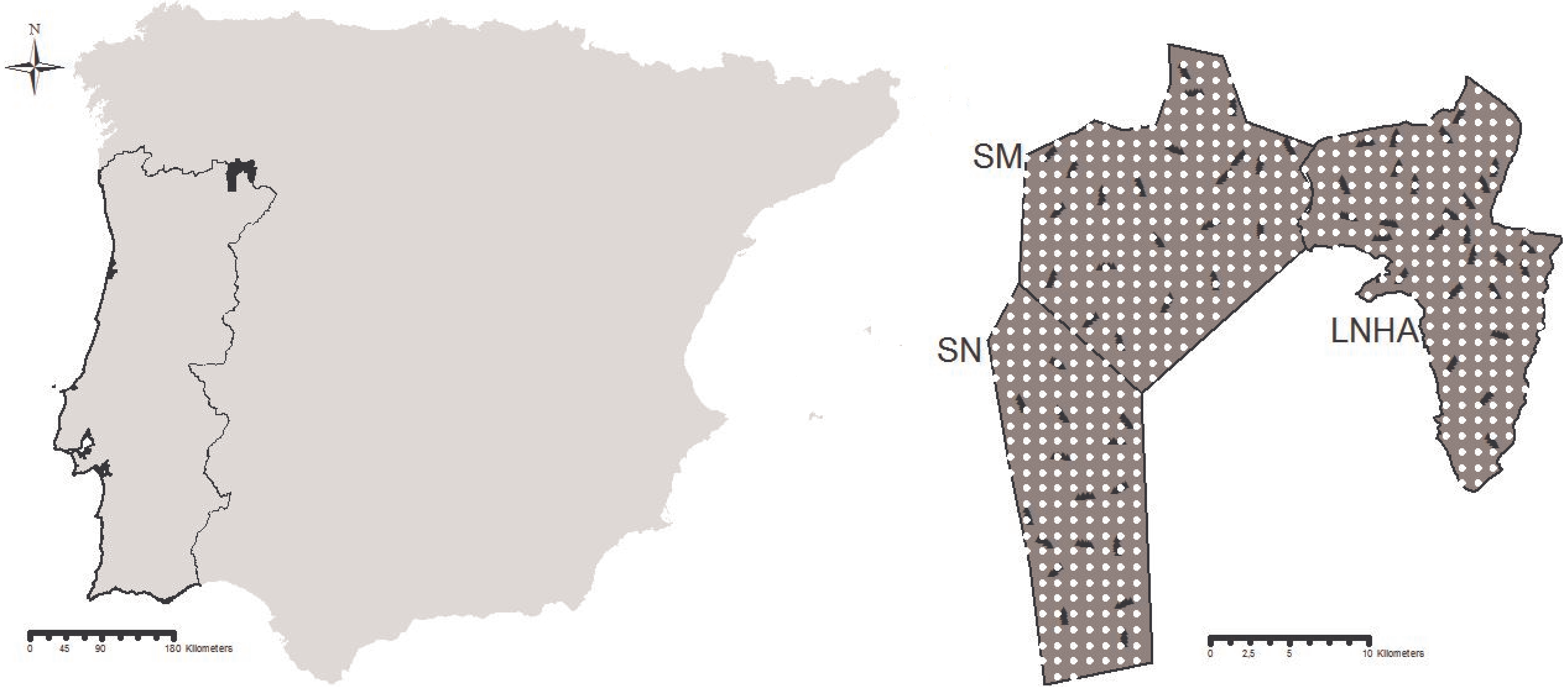
Table 1. Comparison between GCV score, R-square (adjusted), deviance explained, coefficient of variation (CV) and abundance among DSM's with and without geographical variables, with comparison of p-values and estimated degrees of freedom of each variable.

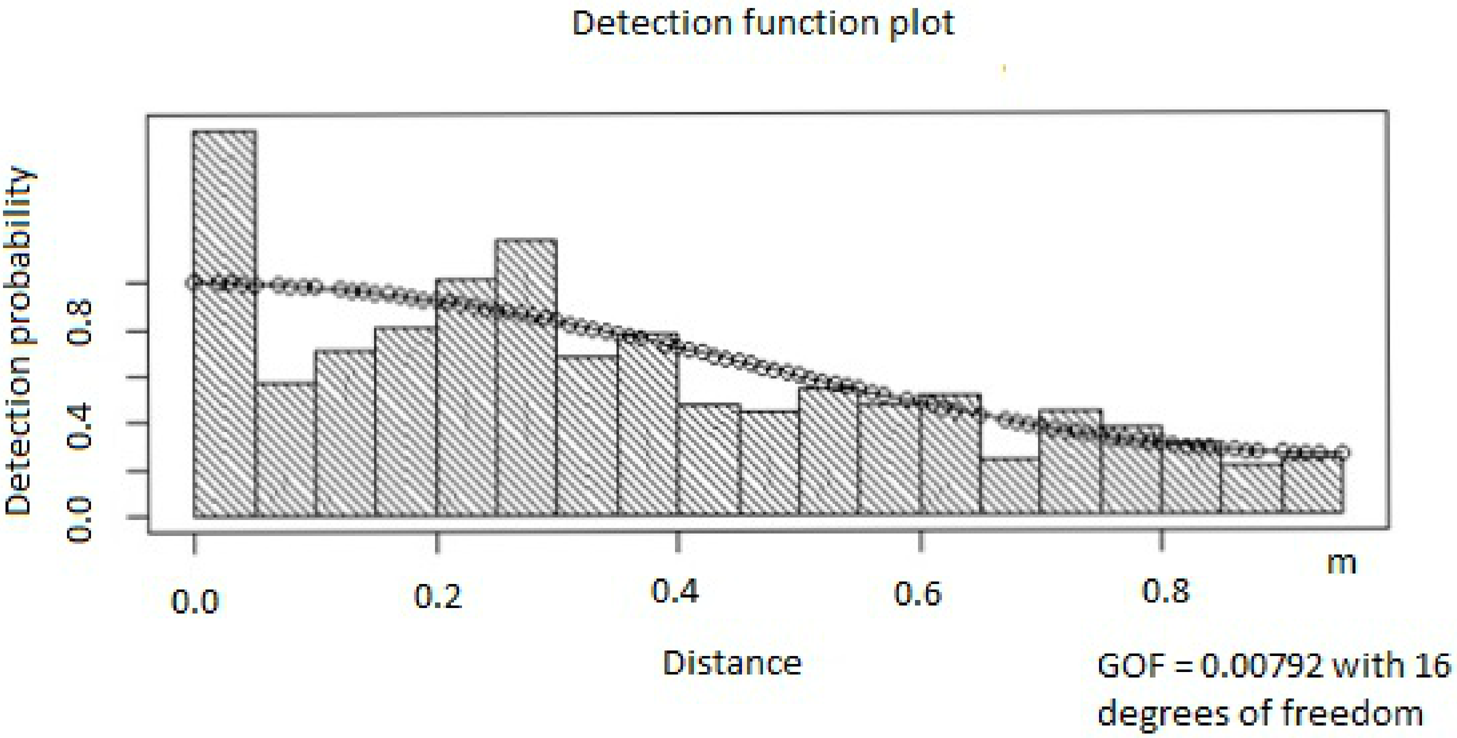
	p-value	Estimated d.f.	GCV	R-square (adjusted)	Deviance explained (%)	CV (%)	Abundance
Without geographical variables							
dsm 1 *			**2.694**	**0.047**	**7.17**	**30.45**	**1878**
dist_hum	**0.003**	1.661					
dist_road	**0.019**	1.000					
ca_perc	**0.022**	1.000					
With geographical variables							
dsm 2			2.561	0.106	17.4	36.07	1926
dist_hum	0.096	2.625					
dist_road	0.577	1.000					
ca_perc	0.084	1.000					
geographic	0.030	6.591					
dsm 3 *			**2.535**	**0.108**	**17.3**	**32.82**	**1909**
dist_hum	**0.105**	2.348					
ca_perc	**0.067**	1.000					
geographic	**0.008**	6.643					
dsm 4			2.554	0.082	13.9	32.22	1836
ca_perc	0.120	5.904					
geographic	0.003	6.571					
dsm 5			2.552	0.076	13.1	30.30	1846
geographic	0.002	6.190					

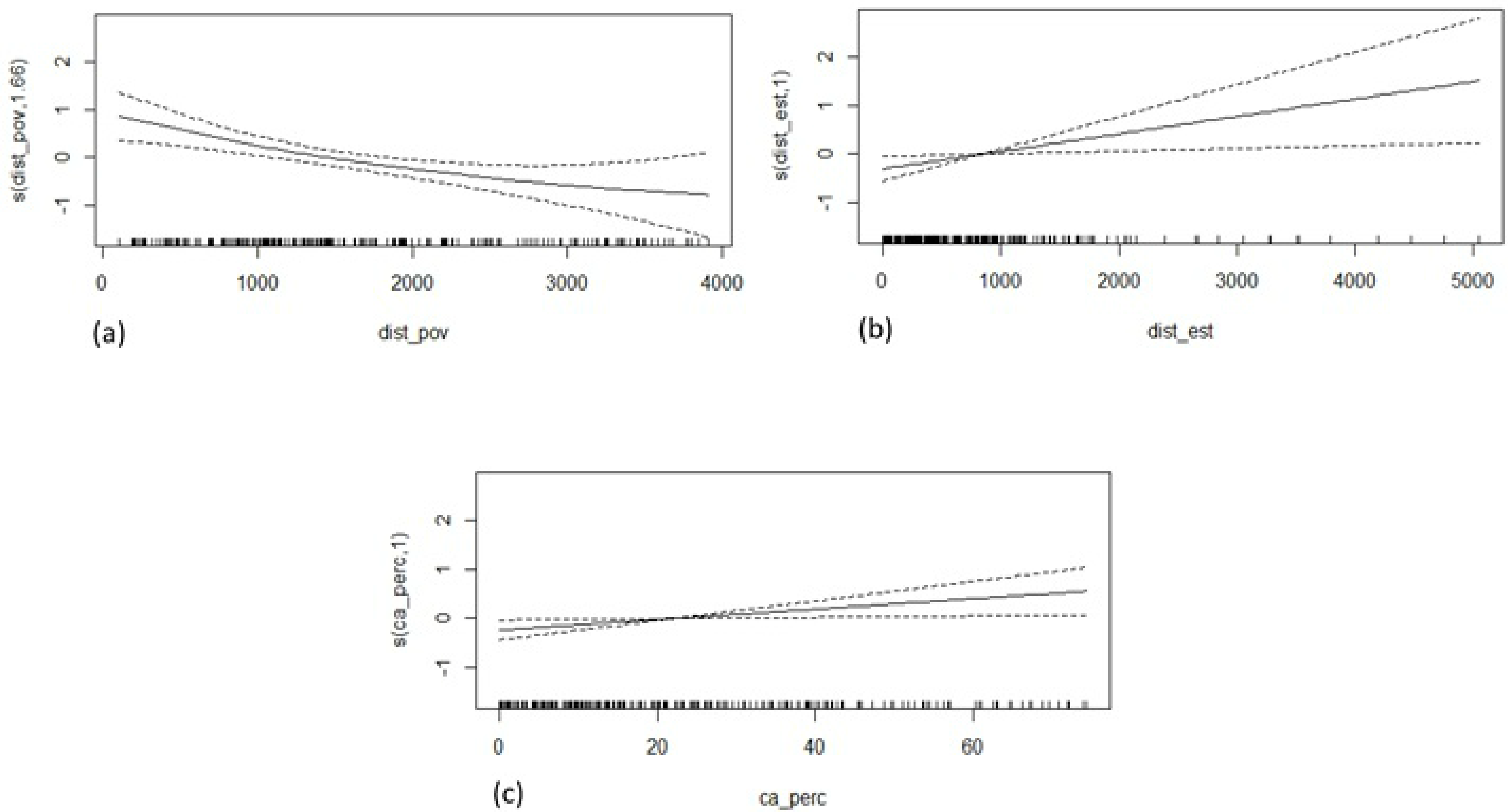
*dsm chosen for inference.

Table 2. Comparison between Density Surface Model and traditional distance sampling through analysis of abundance, density, 95% Confidence Interval and Coefficient of Variation (%) for the total area and for the three sub-areas: SN, SM and LNHA.

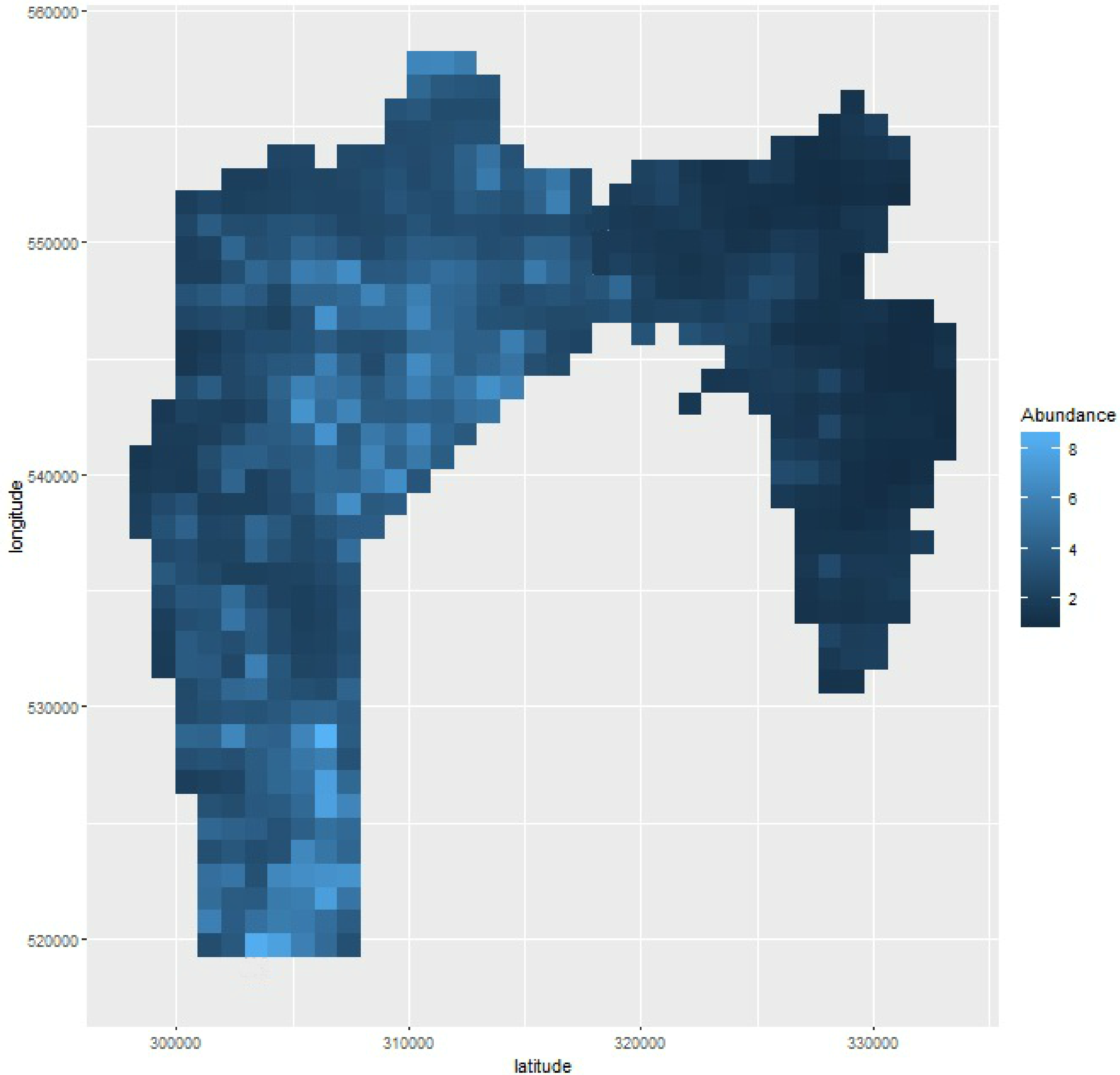
	Method							
	DSM (with geographical variables)	**DS**	**DSM (with geographical variables)**	**DS**	**DSM (with geographical variables)**	**DS**	**DSM (with geographical variables)**	**DS**
	Total area		**SN**		**SM**		**LNHA**	
Abundance	1,909	2,233	662	693	913	1,262	331	278
Density	3.01	3.53	3.62	3.79	3.74	5.17	1.59	1.34
Density - 95% Confidence Interval	0.37 – 3.51	2.07 – 4.79	0.50 – 4.04	2.10 – 6.52	1.67 – 4.40	3.56 – 6.74	0.47 – 3.31	0.82 – 2.51
Coefficient of variation (%)	32.82	24.3	27.90	28.50	27.40	22.54	58.47	32.33


N
0 45 90 180 Kilometers
SM
SN
LNHA
0 2,5 5 10 Kilometers


Detection function plot
Detection probability
0.0
0.4
0.8
0.0
0.2
0.4
0.6
0.8
m
Distance
GOF = 0.00792 with 16 degrees of freedom



s(dist_pov,1.66)
2
1
0
-1
0
1000
2000
3000
4000
(a)
dist_pov
s(dist_est,1)
2
1
0
-1
0
1000
2000
3000
4000
5000
(b)
dist_est
s(ca_perc,1)
2
1
0
-1
0
20
40
60
(c)
ca_perc



560000
550000
540000
530000
520000
longitude
300000
310000
320000
330000
latitude
Abundance
8
6
4
2