

Using zero-inflated models to predict the relative distribution and abundance of roe deer over very large spatial scales

Yaëlle Bouyer^{1,2,*}, Thibaud Rigot³, Manuela Panzacchi⁴, Bram Van Moorter⁴, Pascal Poncin¹, Roseline Beudels-Jamar², John Odden⁴ & John D. C. Linnell⁴

¹) *Unité de Biologie du Comportement, Université de Liège, Quai Van Beneden, 22, B-4000 Liège, Belgium (*corresponding author's e-mail: yaellebouyer@live.fr)*

²) *Section "Evaluation Biologique", Royal Institute of Natural Sciences of Belgium, Rue Vautier 29, B-1000 Brussels, Belgium*

³) *Pests and Diseases: Risk Analysis and Control (UR 106), Centre de coopération internationale en recherche agronomique pour le développement (Cirad), TA-A106/D, Campus International de Baillarguet, F-34398 Montpellier cedex 5, France*

⁴) *Norwegian Institute for Nature Research, P.O. Box 5685 Sluppen, NO-7485 Trondheim, Norway*

Received 13 Mar. 2014, final version received 20 July 2014, accepted 18 Aug. 2014

Bouyer, Y., Rigot, T., Panzacchi, M., Van Moorter, B., Poncin, P., Beudels-Jamar, R., Odden, J. & Linnell, J. D. C. 2015: Using zero-inflated models to predict the relative distribution and abundance of roe deer over very large spatial scales. — *Ann. Zool. Fennici* 52: 66–76.

In Norway, recovering populations of large carnivores commonly prey on roe deer (*Capreolus capreolus*). Understanding predator habitat use and ecology requires fine-scaled information on prey distribution and abundance. However, the massive spatial scales at which large carnivores use the landscape presents many practical and statistical challenges for developing functional prey distribution models. Pellet-count data from > 1000 km of transects gathered across southeastern Norway between 2005 and 2011 were used to derive a map of the relative prey abundance for roe deer. These data were modeled using zero-inflated hurdle models using both environmental and anthropogenic variables. Snow depth and agricultural fields were the most significant variables in explaining both presence and abundance. Internal *k*-cross validation of the model showed medium accuracy (Spearman's $r = 0.35$), whereas external evaluation carried out on the basis of independently collected snow-tracking data (Spearman's $r = 0.37$) and hunting statistics (Spearman's $r = 0.88$) showed higher accuracy. The map generated can facilitate both the study of broad scale processes linking predators and prey as well as roe deer management in southeastern Norway.

Introduction

Large carnivores, like Eurasian lynx (*Lynx lynx*) and wolves (*Canis lupus*), are returning to many

areas of Europe from which they have been absent for decades (Linnell *et al.* 2009, Linnell & Boitani 2012). Managing their recovery in the heavily human-modified landscapes of Europe

requires a detailed understanding of how they relate to a range of habitat and landscape features. Several studies have focused on issues like effects of human infrastructure and broad scale habitat availability (e.g. Schadt *et al.* 2002). However, knowledge of prey distribution and abundance is an important precondition to gain a better mechanistic understanding of carnivore demography, movements and habitat use (Karanth *et al.* 2004, Herfindal *et al.* 2005). Eurasian lynx and wolves have very large annual home ranges [100–1500 km² for lynx (Linnell *et al.* 2001, Mattisson *et al.* 2011), and 100–2000 km² for wolves (Mech & Boitani 2003, Mattisson *et al.* 2013)] as compared with their prey's home ranges [e.g. roe deer (*Capreolus capreolus*): 1 km² (Jeppesen 1989, Cederlund & Liberg 1995, Morellet *et al.* 2013)], and carnivore population processes occur over even larger areas. Models of prey habitat selection are often built at much smaller scales than the areas over which they need to be utilized for large carnivore research. Due to logistical reasons (lack of funds, time etc.) it is rarely feasible to intensively map ungulate abundance over very large areas. Such surveys will also only reflect a snapshot of animal distribution. When working with large carnivores, the use of predictive prey distribution models are therefore needed to extend the models to larger scales. When modeling issues related to lynx and wolves, this can imply study areas in the order of several tens of thousands of square kilometers (e.g. Basille *et al.* 2009, 2013). The logistical and financial constraints of all research projects impose limitations on how much sampling is possible. An additional challenge with sampling is imposed in study areas with low prey densities, which tends to give zero inflated data.

Species distribution models (SDMs) are empirical models relating species occurrences to eco-climatic variables on a correlative basis (Guisan & Thuiller 2005, Fitzpatrick & Hargrove 2009). They are mostly produced to map the potential geographic range of species within study regions, according to the hypothesis that their distribution is actually influenced by the factors investigated. While SDMs are commonly used for conservation purposes, few have used zero-inflated models. Zero-inflated models are useful tools when data sets exhibit more zeroes

than acceptable for a Poisson or negative binomial model (Zeileis *et al.* 2008), which is not unusual with ecological data sets such as counts of abundance or presence/absence. Statistically ignoring the existence of zero inflation can cause bias in parameter estimates as well as overdispersion (Zuur *et al.* 2009). When assessing the relationship between the abundance of a species and its environment, it has been shown that hurdle models performed better than Poisson or negative binomial models (Potts & Elith 2006), which justified their uses in this study.

Southeastern Norway has recently been recolonized by the Eurasian lynx and the wolf, two large carnivores regularly involved in conflicts with humans due to their predation on wild and domestic animals. For management programs to be effective, information on large-carnivore behavior and future abundance are needed. Roe deer are the main prey of lynx as well as a very common prey of wolf in this region (Wikenros 2001, Müller 2006, Odden *et al.* 2006). We mapped roe deer distribution and abundance in regions where lynx and wolf are present. Even though roe deer are known to be widely distributed throughout much of the country (Andersen *et al.* 2004), and their fine scale (habitat patch level) pattern of habitat use is well documented and understood (e.g. Torres *et al.* 2011), no abundance maps are currently available for any scale. For this reason, we used hurdle models based on pellet count data to investigate the importance of ecological and anthropogenic factors related to the species' abundance. In Norway, snow has a high impact on roe deer occurrence as it can affect both access to food and mobility, for example when escaping from predators (Jędrzejewski *et al.* 1992). Based on this, we expected roe deer to be negatively associated with increasing snow depth (Prediction 1). We also expected a selection for agricultural lands (Prediction 2) since human-dominated landscapes are known to provide high-quality ecotones and important forage resources (Panzacchi *et al.* 2010, Torres *et al.* 2012). Other variables that describe habitat, climatic and human disturbance characteristics were also included in an attempt to improve model predictions.

We created a large-scale abundance map that will then be available as a map layer to

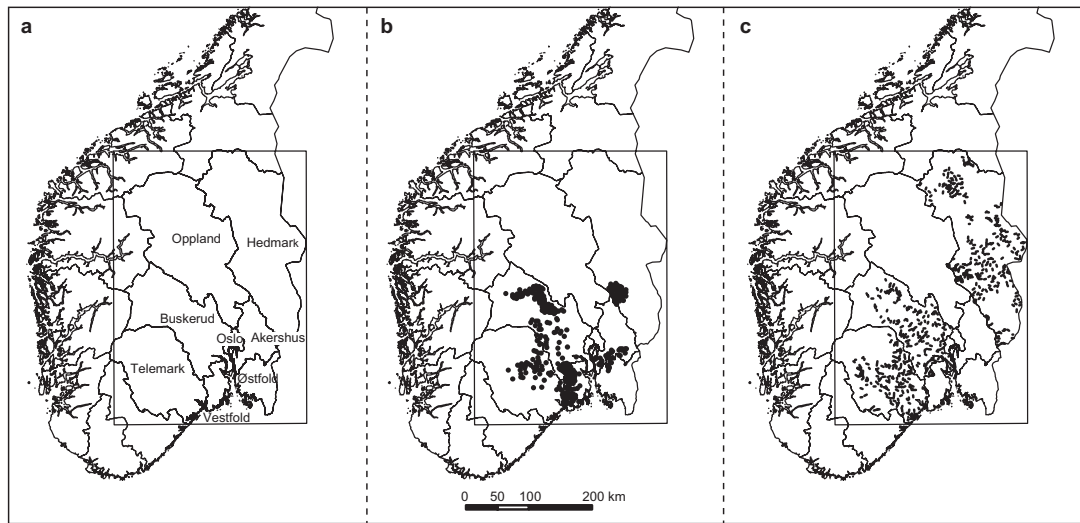


Fig. 1. (a) The five regions of the study area are situated in southeastern of Norway, (b) pellet group transects are represented by grey dots, and (c) snow-tracking transects are shown with black lines. The frame represents the extent of the predictive maps.

study large carnivore ecology (e.g. Gervasi *et al.* 2014). As the area to which the map was extrapolated was very large, the evaluation process was particularly important, and was conducted in two ways: firstly, using a spatial k -fold cross validation procedure and, secondly, by comparing the model results to two independent abundance indices. Based on the predictive map generated, we briefly discuss the implications for predator population distribution and dynamics.

Material and methods

Study site

The study took place in southeastern Norway across six counties (Telemark, Vestfold, Østfold, Buskerud, Hedmark and Akershus) (Fig. 1, extent area: 15 550 km²). The area spans a wide range of environmental conditions: from highly fragmented agricultural areas in the southeast (Østfold and Akershus) to forest dominated areas in the north (Hedmark). The topography is also highly variable: from flat or hilly areas in the east to higher altitudes with steep slopes in the west and north (elevation [mean \pm SD]: 1020 \pm 593 m). Overall, the forests are intensively exploited, and are mainly composed of Nor-

wegian spruce (*Picea abies*), Scots pine (*Pinus sylvestris*), hoary alder (*Alnus incana*) and birch (*Betula pubescens*). The study site is described in greater detail in other papers (Basille *et al.* 2009).

Sampling of pellet data

Data on roe deer pellet groups were accumulated for the whole study area during the years 2005, 2007, 2008 and 2011. Pellets were counted along 346 transects by different observers who all received the same training in pellet identification. Transects were stratified by altitude and distributed randomly inside and adjacent to the areas surrounding polygons representing the home-ranges of lynx that were equipped during the same period with GPS collars (Fig. 1). Each transect consisted of a 1 km \times 1 km \times 1 km triangle. A total of 30 circular 10-m² plots were surveyed for pellets, with one plot located at 100-m intervals along each triangle (Wahlstrom & Liberg 1995). This is logistically a very efficient field design with the starting point coinciding with the final point (Lindén *et al.* 1996). Each transect was chosen to fit within one 1 \times 1 km square on the UTM grid. If a transect fell on areas that could not be surveyed, such as a

lake, urban area, or an area totally dominated by agricultural fields, it was moved to the nearest 1×1 km square containing sufficient natural or semi-natural habitats. Within a transect, individual plots were only counted if they fell on natural or semi-natural habitats (e.g. forest, bogs, alpine tundra, moorland, rough meadows). Agricultural fields under cultivation were not counted since it was not possible to walk on fields during spring due to young crops growing and because many fields were ploughed in spring, which destroys the ability to count pellets accumulated during the autumn/winter. Although we can expect roe deer to be frequent users of fields when foraging, they rarely bed on fields and did not use them much during daylight, usually retreating to the surrounding forests to rest. Therefore, we did not expect the failure to count pellets on fields to greatly bias our results considering the spatial scale (1 km^2) of our analysis.

Following Mayle *et al.* (2000) a pellet group was counted when it contained six or more pellets, when it was considered to have been produced during the same defecation event (same color, texture, size, shape) and when the group center was within the plot. Other herbivorous species are present in the same area such as moose and red deer but assigning pellet group to the right species was possible due to differences in size and shape.

Data were collected in spring (mainly May and early June) to avoid high vegetation in summer and snow cover in winter. The pellets

accumulated thus mainly reflected the habitat use for a period extending from autumn, through winter, to spring. However, pellets decompose slowly in this northern environment, such that it is also possible that some summer pellets were counted as well. Given that the accumulation period covers periods when roe deer are in both winter and summer ranges (for the proportion of the population that migrates) the results should be viewed as reflecting year round distribution.

Explanatory variables for relative roe deer distribution and abundance

Explanatory data used to create the roe deer map were classified into two categories, relating to habitat (i.e. environmental data) and to human activity (Table 1). These data were chosen based on previous fine-scale studies of roe deer habitat selection (Torres *et al.* 2011, Torres *et al.* 2012) and on the availability of digital maps at the scale of the study area. The data were recoded into raster layers of 1 km^2 spatial resolution using a geographic information system (ArcGIS 9.3.1). In order to account for the broader landscape effect of each variable, we computed their respective average value around the center of each sampling plot (i.e. focal cell). This computation was carried out using moving windows of $3 \times 3 \text{ km}$, $7 \times 7 \text{ km}$ and $11 \times 11 \text{ km}$ pixels, which approximately correspond to buffers of 1.5, 3.5 and 5.5 km radii around the focal cell.

Table 1. Explanatory variables used in the analysis. All covariates were at a 1 km^2 scale.

Variables	Explanation
Habitat	
Forest	Proportion of forest within each pixel
Bogs	Proportion of bogs within each pixel
Snow	Mean snow depth (mm) during winter within each pixel
Slope	Mean of the slope (degrees) within each pixel
Hillshade	Index of the combination of the slope and aspect of the terrain in relationship to the azimuth and aspect of the sun within each pixel
Climatic gradient	Regional variation (gradient) from coast to inland and from oceanic/humid to continental each areas within pixel
Human activity	
Field	Proportion of agricultural fields within each pixel
Public roads	Total length (km) within each pixel
Private roads	Total length (km) within each pixel
Human density	Number of individuals within each pixel

The Global Land Cover 2000 database was used to calculate the amount of forest, fields and bogs within each pixel. Layers for slope and hillshade were obtained from the Norwegian Mapping Authority and interpolated snow depth maps were obtained from the Norwegian Meteorological Institute. Hillshade is a measure of the incoming solar radiation for the pixel at 12:00 on 20 June. The values ranged from 0 to 255, where pixels with the value 0 are in complete shadow whereas pixels with the value 255 are facing directly (perpendicular) towards the sun. Accordingly, it is an index representing a combination of the slope and aspect of the terrain in relationship to the azimuth and aspect of the sun. An additional environmental variable, called Climatic Gradient, was obtained from Bakkestuen *et al.* (2008). This variable is a principal component extracted from a set of 54 environmental variables and represents regional variation (a gradient) from coast to inland and from oceanic/humid to continental areas. This axis corresponds to one of the bioclimatic gradients used in expert classification of Norway into biogeographical regions (Bakkestuen *et al.* 2008). Data on public and private roads were obtained from the Norwegian Mapping Authority. Data describing human density came from Takle (2002). All the processed explanatory variables were checked for collinearity prior to analyses (filter = 0.7, Zuur *et al.* 2009). From this, collinear variables, such as snow, temperature and elevation, were filtered. Thereby, the most significant variables that remained may not be the ones with the more direct or intuitive mechanistic influence (e.g., if an animal does not occur at high elevation, it is more likely because of temperature-influenced vegetation than elevation itself). Instead, it is those with the highest statistical power that were retained, even if they may only reflect other factors that have a more direct mechanistic influence.

Statistical analysis

Because of the low densities at which boreal forest mammals like roe deer occur, many transects gave zero-count values (53% of the total number of transects for roe deer). This obvious

zero-inflation was instrumental in our analytical approach. Zero-inflated models, like hurdle models, are used when the response variable contains more zeroes than expected based on a Poisson or negative binomial distribution (Zuur *et al.* 2009). Developed by Cragg (1971), hurdle models consist of two parts. The first part uses a binomial distribution to model the presence and absence of a species. The second part is a count model with a Poisson, negative binomial or geometric distribution that can either accommodate zero observations or not (Barry & Welsh 2002) and models species abundance. The positive observations arise from crossing the zero-hurdle or probability threshold (Potts & Elith 2006, Zuur *et al.* 2009). This approach therefore recognizes the possibility that the mechanisms that determine presence/absence can be different from those that determine abundance (Ridout *et al.* 1998). The presence of over-dispersion appeared in the dataset thus justifying the use of negative binomial distributions instead of Poisson distributions (Zuur *et al.* 2009).

To initialize the model, the same set of environmental and human-related covariates were included in both parts of the hurdle (i.e. the presence/absence part and the abundance part). This was achieved in parallel, and independently for the covariates computed within the 3×3 km, 7×7 km and 11×11 km moving windows. To select which buffer size was more accurate, three hurdle models with the same covariates, but collected at different scales were applied. The model with the more accurate buffer size was evaluated using the stepAIC function in R (package MASS, Venables and Ripley, 1999) and $AIC_c < 2$. Finally, the best model was constructed with the covariates that explained a significant ($p < 0.05$) amount of variation of roe deer data. In addition, in order to account for spatial autocorrelation, we computed and added a residual autocovariate (RAC) to our best model based on the method provided in Crase *et al.* (2012). In essence, this approach consists of constructing an autocovariate based on the spatial autocorrelation found in model residuals. In addition, the method has been shown to increase the accuracy of parameter estimates and therefore further limits the potential bias in parameter estimation linked to spatial autocorrelation (Dormann *et al.* 2007).

In order to evaluate the distribution map, two different approaches were used. In the first approach, a spatial k -fold cross-validation (Nilsen *et al.* 2009b) was applied as an internal evaluation, with a k -fold partition of five groups representing sub-regions within the study area (Fig. 1). The model was trained iteratively on four of the five areas using the hurdle procedure. The remaining testing set was used as a validation set. As a second approach, external evaluations were carried out based on two independent datasets representative of relative prey densities. These datasets were compared with prey abundances predicted by the distribution map. The first dataset represents the number of snow tracks recorded along a total of 587 transects, each 3 km in length, collected during winters 2008–2011 in our study area. The correlation between the average number of tracks for all years and per transect (corrected for the number of days since snowfall) and the average abundance of animals predicted by the map in a 3 km buffer zone (diameter equivalent to the length of the snow tracks transects) from the center of each transect was assessed using Spearman's rank correlation. The second dataset consisted of hunting statistics (hunting bag) at the municipal level (Statistics Norway, www.ssb.no), which recorded the number of shot roe deer during the period 2009–2011 in each municipality of the whole study area. It is expected that hunting statistics give only a coarse representation of species abundance, but previous studies have

shown that a very good concordance was present between hunting statistics and other independent indices of roe deer abundance (Grøtan *et al.* 2005, Nilsen *et al.* 2009a). The fact that harvest abundances varied by several orders of magnitudes across the study area also reflects the fact that abundance differences are dramatic.

All analyses were carried out using R (R Development Core Team 2011) and the package *pscl* (Zeileis *et al.* 2008). The package *raster* (Hijmans & van Etten 2011) was used to generate the predictive maps. The spatial extent of the predictive map was selected in order to include the area of lynx and wolf distribution in southeastern Norway (Basille *et al.* 2009). Spatial autocorrelation was investigated using Moran's I statistic from the package *spdep* (Bivand *et al.* 2008).

Results

Hurdle models

The most parsimonious hurdle model (i.e. which minimized AIC values) is presented in Table 2. The best model was obtained with the 3×3 km (1.5 km radius) buffer and identifies factors associated with both roe deer presence (binomial part), and roe deer abundance (negative binomial part) within this area of presence.

Agricultural fields and snow depth were the only variables retained by the model. Presence and abundance of roe deer were positively asso-

Table 2. Response variables retained after hurdle models selection for roe deer: parameter estimation, standard error (SE) and level of significance. The left part of the table represents the best model (buffer scale: 1.5 km) without the autocovariate (RAC). The right part represents the best model with the autocovariate. The model in the right part of the table was kept to predict the map.

	Without RAC			With RAC		
	Estimate	SE	p	Estimate	SE	p
Negative binomial part						
Fields	2.613	0.540	< 0.001	2.209	0.393	< 0.001
Log(Snow)	-0.674	0.164	< 0.001	-0.510	0.147	< 0.001
RAC	-	-	-	0.178	0.039	< 0.001
Binomial part						
Fields	5.242	1.542	< 0.001	5.050	1.561	0.001
Log(Snow)	-0.823	0.191	< 0.001	-0.833	0.193	< 0.001
RAC	-	-	-	0.316	0.180	0.079

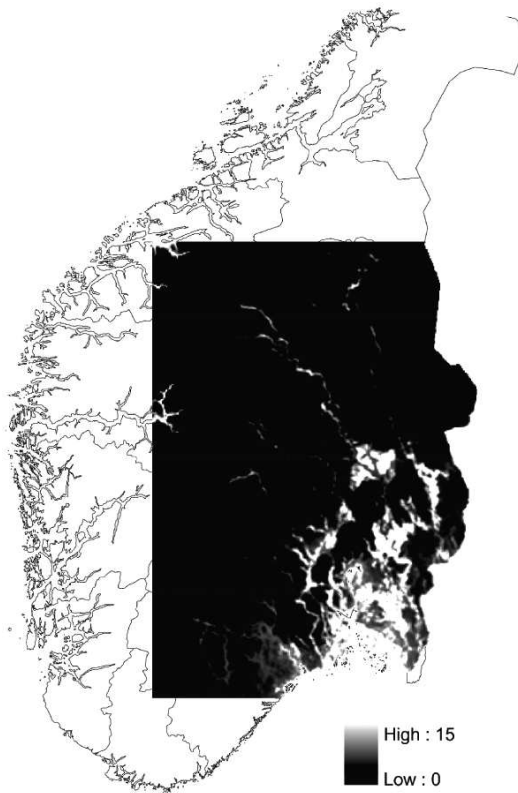


Fig. 2. Roe deer distribution and pellet-count abundance in southeastern Norway.

ciated with the area of fields within the buffer but negatively associated with snow depth.

Model evaluation

On average, Spearman's correlation for the internal evaluation was $r_s = 0.35$ but there were differences between the regions (Buskerud: $r = 0.19$, $p < 0.05$; Telemark: $r_s = 0.27$, $p < 0.001$; Hedmark: $r_s = 0.66$, $p < 0.0001$; Vestfold: $r_s = 0.31$, $p < 0.0001$; Ostfold/Akershus: $r_s = 0.34$, $p < 0.05$). The resulting map indicated that roe deer reach the highest abundances in the southeastern part of the investigated area. The range of predictions for the pellet-counts was between 0 and 15.3 pellets. The predictive map (Fig. 2) was compared with external independent datasets. The Spearman's correlation between snow-tracking transects and model prediction values was significant ($r_s = 0.37$, $p < 0.001$) and

at a coarser spatial scale (i.e. municipality level), the correlation between model predictions and the hunting bags was highly significant ($r_s = 0.888$, $p < 0.001$) (Appendices 1 and 2).

Discussion

Habitat preferences

The factors that were retained from the model selection procedure to explain the observed variation in pellet density are generally consistent with what is known about the species' ecology. The spatial and temporal impact of snow depth on roe deer distribution (Holand *et al.* 1998), habitat use (Myserud *et al.* 1997, Ratikainen *et al.* 2007) and dynamics (Grøtan *et al.* 2005, Nilsen *et al.* 2009a) is well documented, as is their preference for edge habitats at the forest-field interface (Kjøstvedt *et al.* 1998, Torres *et al.* 2011). In Norway, roe deer harvest is subject to only fairly loose management protocols and harvest quotas are rarely filled. This implies that the broad scale species abundance reported in this study is mainly influenced by biological drivers, rather than specific management decisions. Therefore, variables that were found to influence roe deer abundance (snow and fields) very likely depict natural ecological factors influencing roe deer.

Pellet counts, zero inflation and data resolution

When pellets were counted, the observer error was likely to be small because of the small size of the plots which facilitated careful searching and the fact that pellets from roe deer could be readily discriminated. Young red deer pellets can be similar to roe deer pellets but red deer are present only in a small part of the study area which limited the possible error and in the event of doubt, pellets were brought in from the forest for comparison with a reference collection.

Pellet group counts have been shown to be effective for other habitat selection studies (*see* Kurki *et al.* 2000, Shorrocks *et al.* 1998) and have been a technique widely used to assess

ungulate habitat use and density since the 1940s (Neff 1968, Tsaparis *et al.* 2008). Although pellet group surveys are an established method, criticism has been made concerning their use in the estimation of population densities, distribution and habitat use (Neff 1968, Collins & Urness 1981, Rowland *et al.* 1984). With pellet group counts, major concerns come from spatial variation of pellet decay rates (e.g. between moist and dry habitats, Harestad & Bunnell 1987) and from the sensitivity of defecation rates to seasonal variation (Mitchell *et al.* 1985, Rogers 1987). However, use of such indirect count methods are often favored to direct count methods in ungulate research as they demand less effort, cost less and are sometimes more adapted for some habitats such as forest (Marques *et al.* 2001, Laing *et al.* 2003). Also, results from pellet group counts on broad scale habitat use have been shown to be comparable to those obtain by other methods such as radio-tracking (Guillet *et al.* 1995, Månsson *et al.* 2011). Defecation rate was not an issue in this analysis as we did not use it to estimate absolute roe deer abundance.

Hurdle models were used to gain information on the distribution and abundance of roe deer pellet groups, which can be interpreted as a relative distribution and abundance of roe deer. No test was made on decay rates of roe deer pellet groups in the different habitat types in the study area prior to analysis and this could have biased the results. However, in Norway, pellet groups can persist for many months because of the temperate climate, cold winters and general lack of dung beetles. Even if the decay rate is higher in some habitats than others, the difference is likely to be small and impact on pellet group counts should be limited.

Another difficulty we were confronted with in our analysis was the coarse spatial scale of the environmental variables to which we had access. Most of the variables available to use as covariates in the model were only available at the scale of a roe deer's home range (1 km²) and did not differentiate between types or age classes of forest habitats which have been shown to be important for explaining fine-scaled herbivore distribution and density (Torres *et al.* 2011). If the map data had allowed us to work at the individual habitat patch scale instead of the transect

scale we would probably have been better able to make predictions of roe deer distribution and relative abundance. However, such fine scale data are rarely available for large spatial scales.

Model evaluation

Our predictions were evaluated using both snow track transects and hunting bags. These internal evaluations supported our results, although there was much unexplained variation. Species Distribution Models never explicitly incorporate underlying processes, and this constitutes the first important limitation of our study. For instance, this could strongly limit the extrapolative ability of SDMs out of the training area and further explains why the accuracy of spatial *k*-fold cross-validations herein constructed are weak (roe deer, on average $r_s = 0.35$).

On the contrary, the external evaluation carried out with the hunting bag data showed good accuracy ($r_s = 0.88$, for hunting data on roe deer), although our predictions were made for an area 8 times larger than the initial area used for training (although the training areas were scattered throughout the prediction area and covered all parts of the ecological gradient). The higher accuracy of the coarse scale (external evaluation) as compared with that of the local scale (internal evaluation) is surprising since model extrapolation at large scales may potentially lead to significant errors (Guisan & Zimmermann 2000). However, a similar effect has been reported in studies using climate variables at different scales in which coarse scale data (such as NDVI) outperformed local scale climate variables (such as weather station data) when explaining variation in performance traits of large herbivores (Post & Stenseth 1999, Nielsen *et al.* 2012). Thereby, we can assume that the hurdle SDM built in the present study on the pellet-counts data was well suited for mapping broad scale roe deer distribution across southeastern Norway.

Previous studies emphasized the fact that roe deer distribution and density are important drivers for lynx habitat selection in Norway (Basille *et al.* 2009), most likely explained by the importance of this ungulate on lynx diet (more than 80% in winter, Odden *et al.* 2006).

Although there were a number of challenging issues, we believe that the map that we have generated is sufficient for the study of broad scale processes linking predators and prey, including the investigation of the impacts of predators through an increased ability to explain spatial patterns in predator density, movement and kill rates. This map gives information relevant for roe deer management in the face of lynx predation at a regional scale because it will allow a more accurate prediction of relative predator impact as well as a more detailed understanding of the mechanisms of lynx predation behavior. Concerning the external validation of our map with hunting bags, we believed that this roe deer map could be a good asset for roe deer management in Norway at a regional scale. While one could wish for more confident predictions and fine scaled resolutions, the realities (concerning the abundance of training data and the resolution of map data) of large scale studies imply that it is unlikely that it will be possible to improve on the level of accuracy that we obtained for similar species with broad distributions and generalist habits.

Acknowledgments

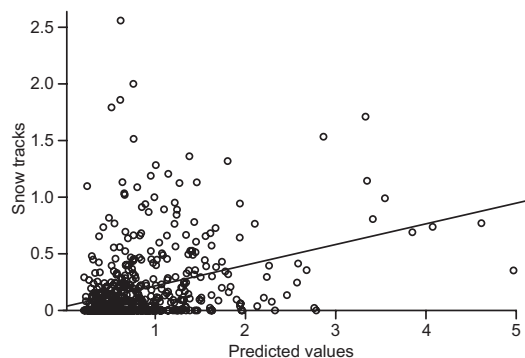
The collection of field data for this analysis was funded by the Norwegian Directorate for Nature Management, the Research Council of Norway, and the Offices of Environmental Affairs in the counties of Oslo & Akershus, Østfold, Vestfold, Telemark and Buskerud, as well as the municipalities of Flå, Gol, Hjartdal, Nes, Nore og Uvdal, Rollag, Sauherad, Tinn and Ål. The analysis was funded by an individual scholarship to YB provided by the FRIA (Fonds pour l'Encouragement de la Recherche Scientifique dans l'Industrie et l'Agriculture). The authors are especially grateful to Rita Torres for sharing her pellet data, as well as to all the other students and field assistants who have assisted in collecting the data. We thank the two anonymous referees for their constructive comments on this work.

References

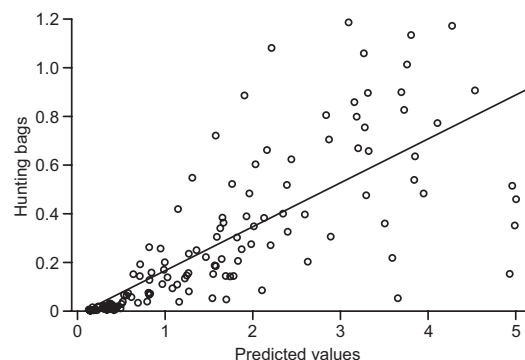
- Andersen, R., Herfindal, I., Sæther, B. E., Linnell, J. D. C., Odden, J. & Liberg, O. 2004: When range expansion rate is faster in marginal habitats. — *Oikos* 107: 210–214.
- Bakkestuen, V., Erikstad, L. & Halvorsen, R. 2008: Step-less models for regional environmental variation in Norway. — *Journal of Biogeography* 35: 1906–1922.
- Barry, S. C. & Welsh, A. H. 2002: Generalized additive modelling and zero inflated count data. — *Ecological Modelling* 157: 179–188.
- Basille, M., Herfindal, I., Santin-Janin, H., Linnell, J. D. C., Odden, J., Andersen, R., Høgda, K. A. & Gaillard, J.-M. 2009: What shapes Eurasian lynx distribution in human dominated landscapes: selecting prey or avoiding people? — *Ecography* 32: 683–691.
- Basille, M., Van Moorter, B., Herfindal, I., Martin, J., Linnell, J. D. C., Odden, J., Andersen, R. & Gaillard, J.-M. 2013: Selecting habitat to survive: the impact of road density on survival in a large carnivore. — *PLoS ONE* 8: e65493, doi:10.1371/journal.pone.0065493.
- Bivand, R. S., Pebesma, E. J. & Gómez-Rubio, V. 2008: *Applied spatial data analysis with R*. — Springer, New-York.
- Cederlund, G. & Liberg, O. 1995: *Rådjuret: viltet, ekologin och jakten*. — Svenska jägareförb., Solna.
- Collins, W. B. & Urness, P. J. 1981: Habitat preferences of mule deer as rated by pellet-group distributions. — *Journal of Wildlife Management* 45: 969–972.
- Cragg, J. G. 1971: Some statistical models for limited dependent variables with application to the demand for durable goods. — *Econometrica* 39: 829–844.
- Crane, B., Liedloff, A. C. & Wintle, B. A. 2012: A new method for dealing with residual spatial autocorrelation in species distribution models. — *Ecography* 35: 879–888.
- Dormann, C. F., McPherson, J. M., Araújo, M. B., Bivand, R., Bolliger, J., Carl, G., Davies, R. G., Hirzel, A., Jetz, W., Daniel Kissling, W., Kühn, I., Ohlemüller, R., Peres-Neto, P. R., Reineking, B., Schröder, B., M. Schurr, F. & Wilson, R. 2007: Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. — *Ecography* 30: 609–628.
- Fitzpatrick, M. C. & Hargrove, W. W. 2009: The projection of species distribution models and the problem of non-analog climate. — *Biodiversity and Conservation* 18: 2255–2261.
- Gervasi, V., Nilsen, E., Odden, J., Bouyer, Y. & Linnell, J. D. C. 2014: The spatio-temporal distribution of wild and domestic ungulates modulates lynx kill rates in a multi-use landscape. — *Journal of Zoology* 292: 175–183.
- Grøtan, V., Sæther, B. E., Engen, S., Solberg, E. J., Linnell, J. D. C., Andersen, R., Brøseth, H. & Lund, E. 2005: Climate causes large-scale spatial synchrony in population fluctuations of a temperate herbivore. — *Ecology* 86: 1472–1482.
- Guillet, B., Bergström, R., Cederlund, G., Bergström, J. & Ballon, P. 1995: Comparison of telemetry and pellet-group counts for determining habitat selectivity by roe deer (*Capreolus capreolus*) in winter. — *Gibier Faune Sauvage* 12: 253–269.
- Guisan, A. & Zimmermann, N. E. 2000: Predictive habitat distribution models in ecology. — *Ecological Modelling* 135: 147–186.
- Guisan, A. & Thuiller, W. 2005: Predicting species distribution: offering more than simple habitat models. — *Ecology Letters* 8: 993–1009.
- Harestad, A. S. & Bunnell, F. 1987: Persistence of black-

- tailed deer fecal pellets in coastal habitats. — *Journal of Wildlife Management* 51: 33–37.
- Herfindal, I., Linnell, J. D. C., Odden, J., Nilsen, E. B. & Andersen, R. 2005: Prey density, environmental productivity and home-range size in the Eurasian lynx (*Lynx lynx*). — *Journal of Zoology* 265: 63–71.
- Hijmans, R. J. & van Etten, J. 2011: *Geographic analysis and modeling with raster data*. — CRAN project <http://raster.r-forge.r-project.org/>
- Holand, Ø., Mysterud, A., Wannag, A. & Linnell, J. D. C. 1998: Roe deer in northern environments: physiology and behaviour. — In: Andersen, R., Duncan, P. & Linnell, J. D. C. (eds.), *The European roe deer: the biology of success*: 117–137. Scandinavian University Press, Oslo, Norway.
- Jędrzejewski, W., Jędrzejewska, B., Okarma, H. & Ruprecht, A. L. 1992: Wolf predation and snow cover as mortality factors in the ungulate community of the Białowieża National Park, Poland. — *Oecologia* 90: 27–36.
- Jeppesen, J. L. 1989: Activity patterns of free-ranging roe deer (*Capreolus capreolus*) at Kaloe. — *Danish Review of Game Biology* 13: 1–32.
- Karanth, K. U., Nichols, J. D., Kumar, N. S., Link, W. A. & Hines, J. E. 2004: Tigers and their prey: predicting carnivore densities from prey abundance. — *Proceedings of the National Academy of Sciences of the United States of America* 101: 4854–4858.
- Kjøstvedt, J. H., Mysterud, A. & Østbye, E. 1998: Roe deer *Capreolus capreolus* use of agricultural crops during winter in the Lier valley, Norway. — *Wildlife Biology* 4: 23–31.
- Kurki, S., Nikula, A., Helle, P. & Lindén, H. 2000: Landscape fragmentation and forest composition effect on grouse breeding success in boreal forests. — *Ecology* 81: 1985–1997.
- Laing, S. E., Buckland S. T., Burns R. W., Lambie D. & Amphlett A. 2003: Dung and nest surveys: estimating decay rates. — *Journal of Applied Ecology* 40: 1102–1111.
- Lindén, H., Helle, E., Helle, P. & Wikman, M. 1996: Wildlife triangle scheme in Finland: methods and aims for monitoring wildlife populations. — *Finnish Game Research* 49: 4–11.
- Linnell, J. D. C. & Boitani, L. 2012: Building biological realism into wolf management policy: The development of the population approach in Europe. — *Hystrix* 23: 80–91.
- Linnell, J. D. C., Breitenmoser, U., Breitenmoser-Würsten, C., Odden, J. & von Arx, M. 2009: Recovery of Eurasian lynx in Europe: What part has reintroduction played? — In: Hayward, M. S. M. (ed.), *Reintroduction of top-order predators*: 72–91. Wiley-Blackwell, Oxford, UK.
- Linnell, J. D. C., Andersen, R., Kvam, T., Andrén, H., Liberg, O., Odden, J. & Moa, P. F. 2001: Home range size and choice of management strategy for lynx in Scandinavia. — *Environmental Management* 27: 869–879.
- Månsson, J., Andrén, H. & Sand, H. 2011: Can pellet counts be used to accurately describe winter habitat selection by moose *Alces alces*? — *European Journal of Wildlife Research* 57: 1017–1023.
- Marques, F. F. C., Buckland, S. T., Goffin, D., Dixon, C. E., Borchers, D. L., Mayle, B. A. & Peace, A. J. 2001: Estimating deer abundance from line transect surveys of dung: Sika deer in southern Scotland. — *Journal of Applied Ecology* 38: 349–363.
- Mattisson, J., Persson, J., Andrén, H. & Segerström, P. 2011: Temporal and spatial interactions between an obligate predator, the Eurasian lynx (*Lynx lynx*), and a facultative scavenger, the wolverine (*Gulo gulo*). — *Canadian Journal of Zoology* 89: 79–89.
- Mattisson, J., Sand, H., Wabakken, P., Gervasi, V., Liberg, O., Linnell, J. D. C., Rauset, G. R. & Pedersen, H. C. 2013: Home range size variation in a recovering wolf population: evaluating the effect of environmental, demographic, and social factors. — *Oecologia* 1–13.
- Mayle, B. A., Putman, R. J. & Wyllie, I. 2000: The use of trackway counts to establish an index of deer presence. — *Mammal Review* 30: 233–237.
- Mech, L. D. & Boitani, L. 2003: *Wolves: behavior, ecology, and conservation*. — University of Chicago Press, Chicago.
- Mitchell, B., Rowe, J. J., Ratcliffe, P. & Hinge, M. 1985: Defecation frequency in Roe deer (*Capreolus capreolus*) in relation to the accumulation rates of faecal deposits. — *Journal of Zoology* 207: 1–7.
- Morellet, N., Bonenfant, C., Börger, L., Ossi, F., Cagnacci, F., Heurich, M., Kjellander, P., Linnell, J. D. C., Nicoloso, S., Sustr, P., Urbano, F. & Mysterud, A. 2013: Seasonality, weather and climate affect home range size in roe deer across a wide latitudinal gradient within Europe. — *Journal of Animal Ecology* 82: 1326–1339.
- Müller, S. 2006: *Diet composition of wolves (Canis lupus) on the Scandinavian peninsula determined by scat analysis*. — University of Munich, Germany.
- Mysterud, A., Bjørnsen, B. H. & Østbye, E. 1997: Effects of snow depth on food and habitat selection by roe deer *Capreolus capreolus* along an altitudinal gradient in south-central Norway. — *Wildlife Biology* 3: 27–33.
- Neff, D. J. 1968: The pellet-group count technique for big game trend, census, and distribution: a review. — *The Journal of Wildlife Management* 32: 597–614.
- Nielsen, A., Yoccoz, N. G., Steinheim, G., Storvik, G. O., Rekdal, Y., Angeloff, M., Pettorelli, N., Holand, O. & Mysterud, A. 2012: Are responses of herbivores to environmental variability spatially consistent in alpine ecosystems? — *Global Change Biology* 18: 3050–3062.
- Nilsen, E. B., Linnell, J. D. C., Odden, J. & Andersen, R. 2009a: Climate, season, and social status modulate the functional response of an efficient stalking predator: the Eurasian lynx. — *Journal of Animal Ecology* 78: 741–751.
- Nilsen, E. B., Gaillard, J. M., Andersen, R., Odden, J., Delorme, D., Van Laere, G. & Linnell, J. D. C. 2009b: A slow life in hell or a fast life in heaven: demographic analyses of contrasting roe deer populations. — *Journal of Animal Ecology* 78: 585–594.
- Odden, J., Linnell, J. D. C. & Andersen, R. 2006: Diet of Eurasian lynx, *Lynx lynx*, in the boreal forest of south-eastern Norway: The relative importance of livestock and hares at low roe deer density. — *European Journal of Wildlife Research* 52: 237–244.

- Panzacchi, M., Linnell, J. D. C., Melis, C., Odden, M., Odden, J., Gorini, L. & Andersen, R. 2010: Effect of land-use on small mammal abundance and diversity in a forest-farmland mosaic landscape in south-eastern Norway. — *Forest Ecology and Management* 259: 1536–1545.
- Post, E. & Stenseth, N. C. 1999: Climatic variability, plant phenology, and northern ungulates. — *Ecology* 80: 1322–1339.
- Potts, J. M. & Elith, J. 2006: Comparing species abundance models. — *Ecological Modelling* 199: 153–163.
- R Development Core Team 2011: *R: A language and environment for statistical computing*. — R Foundation for Statistical Computing, Vienna, Austria. [Available at www.R-project.org].
- Ratikainen, I. I., Panzacchi, M., Mysterud, A., Odden, J., Linnell, J. D. C. & Andersen, R. 2007: Use of winter habitat by roe deer at a northern latitude where Eurasian lynx are present. — *Journal of Zoology* 273: 192–199.
- Ridout, M. S., Demétrio, C. G. B. & Hinde, J. 1998: Models for count data with many zeros. — In: *Proceedings of the XIXth International Biometric Conference, Cape Town, December 1998*: 179–192.
- Rogers, L. 1987: Seasonal changes in defecation rates of free-ranging white-tailed deer. — *Journal of Wildlife Management* 51: 330–333.
- Rowland, M. M., White, G. C. & Karlen, E. M. 1984: Use of pellet-group plots to measure trends in deer and elk populations. — *Wildlife Society Bulletin* 12: 147–155.
- Schadt, S., Revilla, E., Wiegand, T., Knauer, F., Kaczensky, P., Breitenmoser, U., Bufka, L., Cervený, J., Koubek, P., Huber, T. C. S. & Trepl, L. 2002: Assessing the suitability of central European landscapes for the reintroduction of Eurasian lynx. — *Journal of Applied Ecology* 39: 189–203.
- Shorrocks, B., Albon, S., Kurki, S., Nikula, A. R. I., Helle, P. & Lindén, H. 1998: Abundances of red fox and pine marten in relation to the composition of boreal forest landscapes. — *Journal of Animal Ecology* 67: 874–886.
- Takle, M. 2002: *Befolkningsstatistikk på rutenett*. — Notater 2002/22, Statistisk sentralbyrå, Norway.
- Torres, R. T., Carvalho, J., Panzacchi, M., Linnell, J. D. C. & Fonseca, C. 2011: Comparative use of forest habitats by roe deer and moose in a human-modified landscape in southeastern Norway during winter. — *Ecological Research* 26: 781–789.
- Torres, R. T., Virgós, E., Panzacchi, M., Linnell, J. D. C. & Fonseca, C. 2012: Life at the edge: Roe deer occurrence at the opposite ends of their geographical distribution, Norway and Portugal. — *Mammalian Biology* 77: 140–146.
- Tsapis, D., Katsanevakis, S., Stamouli, C. & Legakis, A. 2008: Estimation of roe deer *Capreolus capreolus* and mouflon *Ovis aries* densities, abundance and habitat use in a mountainous Mediterranean area. — *Acta Theriologica* 53: 87–94.
- Venables, W. N. & Ripley, B. D. 2002: *Modern applied statistics with S*, 4th ed. — Springer, New York.
- Wahlstrom, L. K. & Liberg, O. 1995: Contrasting dispersal patterns in two Scandinavian roe deer *Capreolus capreolus* populations. — *Wildlife Biology* 1: 159–164.
- Wikenros, C. 2001: *Wolf winter predation on moose and roe deer in relation to pack size*. — Faculty of Natural Resources and Agricultural Sciences (Ulna), Sweden.
- Zeileis, A., Kleiber, C. & Jackman, S. 2008: Regression models for count data in R. — *Journal of Statistical Software* 27: 1–25.
- Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A. & Smith, G. M. 2009: *Mixed effects models and extensions in ecology with R*. — Springer, New-York.



Appendix 1. Graphical representation of the correlation between the predicted values of pellet groups in a 3 km buffer around the snow track transects and the density of snow tracks observed in the transects. $r_s = 0.37$, $p < 0.001$.



Appendix 2. Graphical representation of the correlation between the predicted values of pellet groups in the different municipalities and the density of roe deer observed in each municipality. $r_s = 0.88$, $p < 0.001$.