Eur J Wildl Res (2011) 57:1115–1118
DOI 10.1007/s10344-011-0549-3

TECHNICAL NOTES

Comparing profile methods and site-occupancy modelling for the study of occurrence of an elusive species

Jonathan Rolland · Mathieu Basille · Éric Marboutin · Jean-Michel Gaillard

Received: 27 November 2010 / Revised: 9 May 2011 / Accepted: 13 May 2011 / Published online: 28 May 2011

Abstract Based on 1,053 signs of presence collected between 2002 and 2006 by a network of well-trained observers, we modelled the occurrence of the Eurasian lynx (*Lynx lynx*) in France using two methods. The Mahalanobis distance factor analysis (MADIFA) provided a measure of habitat suitability based on environmental covariates, and site-occupancy modelling provided estimates of both presence and detection probabilities over time. Environmental covariates included in the site-occupancy modelling markedly improved the fit of the lynx presence model. We found a strong correlation between habitat suitability scores estimated from the MADIFA and probabilities of presence estimated from the site-occupancy modelling, indicating that both methods provided a convergent assessment of lynx potential occurrence.

Communicated by C. Gortázar

J. Rolland
École Normale Supérieure,
45 rue d’Ulm,
75005 Paris, France

J. Rolland (✉) · M. Basille · J.-M. Gaillard
Laboratoire de Biométrie et Biologie Évolutive,
Université de Lyon, F-69000, Lyon;
Université Lyon 1; CNRS, UMR5558,
F-69622 Villeurbanne, France
e-mail: jonathanketoi@hotmail.fr

M. Basille
Chaire de Recherche Industrielle CRSNG-Université Laval en
Sylviculture et Faune, Département de Biologie, Université Laval,
1045 Av. de la Médecine, Pavillon Alexandre Vachon,
Québec, QC G1V 0A6, Canada

É. Marboutin
Office National de la Chasse et de la Faune Sauvage,
Z.I. de Mayencin, 5 allée de Béthléèm,
38 610 Gières, France

Keywords Detection probability · Eurasian lynx · *Lynx lynx* · MADIFA · Presence · Spatial distribution

Introduction

Elusive species, and among them large carnivores, are usually monitored based on the collection of indirect signs of presence (Linnell et al. 1998). “Profile methods” (sensu Pearce and Boyce 2006) were developed to explore presence-only designs without the need of generating “pseudo-absences” and aim at extracting the characteristics of the environment used by a given organism (Calenge et al. 2005). Among them, the ecological-niche factor analysis (ENFA, Hirzel et al., 2002) gained a high popularity due to its easy application. However, Calenge et al. (2008) demonstrated that the ENFA was not statistically appropriate for computing habitat suitability maps (HSM). They proposed instead the Mahalanobis distance factor analysis (MADIFA, Calenge et al. 2008) to this very purpose, a method built upon the same statistical framework as the ENFA (Calenge and Basille 2008). Studying the detection of presence using site-occupancy modelling (MacKenzie et al. 2006) offers an alternative to tackle presence-only designs. Following a framework similar to mark–recapture methods (e.g., Chao 1987), site-occupancy modelling provides the estimates of both presence and detection probabilities of signs from sites sampled over time (MacKenzie et al. 2006). Both profile techniques and site-occupancy modelling may be helpful when assessing a species’ range by relying on presence-only data (Austin 2007).

In the present paper, we compared the results obtained from both methods applied to occurrences of Eurasian lynx

(*Lynx lynx*) in the French Jura Mountains. After their disappearance from France in the 1920s, lynx re-colonized the French Jura in the 1970s after reintroductions in neighbouring Switzerland (Vandel et al. 2006). Lynx are cryptic large carnivores for which problems of variable detection have already been identified (e.g. Marboutin et al. 2006). We aimed here at assessing whether profile methods and modelling approaches provided similar patterns of lynx occurrence in the Jura Mountains.

Material and methods

Study area and data collection

Within the French species range, the Jura Mountains (~16,000 km^2, including the French departments of Doubs, Jura and Ain, see also Fig. 1) have been re-colonized early so that most of the suitable habitat is now occupied. A network of ca. 1,200 specifically trained volunteers (mostly state employees like game wardens, park rangers and foresters), set up by the French National Game And Wildlife Agency, collected signs of lynx presence in the field, later checked and validated by a lynx specialist (Vandel and Stahl 2005). Although particular landscape features (e.g. steep slopes in forested areas) might have led to some spatial heterogeneity in the detection probability of lynx signs, the high level of experience and training of observers are likely to have kept such heterogeneities at a minimum. As our analysis averages over any variation in p over time, we thus assumed a constant detection probability throughout the lynx distribution area in French Jura.

During the study period (2002–2006), 1,053 presence signs (i.e. faeces, hair, tracks, visual observations, wild and domestic preys, camera traps) have been validated in the Jura Mountains, which represent 65% of the signs validated over the entire country. This selection of presence indices mostly excludes possible false positives. A grid of fixed resolution (81 km^2, i.e. cells of 9×9 km, which corresponds to half an average female lynx home range, Vandel and Stahl 2005) was first superimposed on the sign distribution. The study period was then split into five 1-year periods, for which each cell was assumed to have been visited at least once. The presence of lynx was finally defined as the occurrence of at least one lynx sign in a cell during that period, while an absence of sign corresponded to no detection. On the same grid, we used the Corine Land Cover 2006 database (European Environment Agency 2006) to define a set of environmental variables (including railway density, road density, river density, distance to highway, distance to railway, distance to river and proportion of forest cover) that could potentially influence lynx distribution based on previous analyses (Basille et al. 2008).

Methods

The MADIFA is based on a decomposition of Mahalanobis distances, i.e. the distance from each grid cell to the species' optimum given the correlation structure of the niche (Calenge et al. 2008) into successive uncorrelated axes. Based on the associated eigenvalues, the first few axes are then selected to compute a score of habitat suitability for each cell leading to an estimated HSM. These scores are in fact Mahalanobis distances; the greater the score, the farther the cell is from the

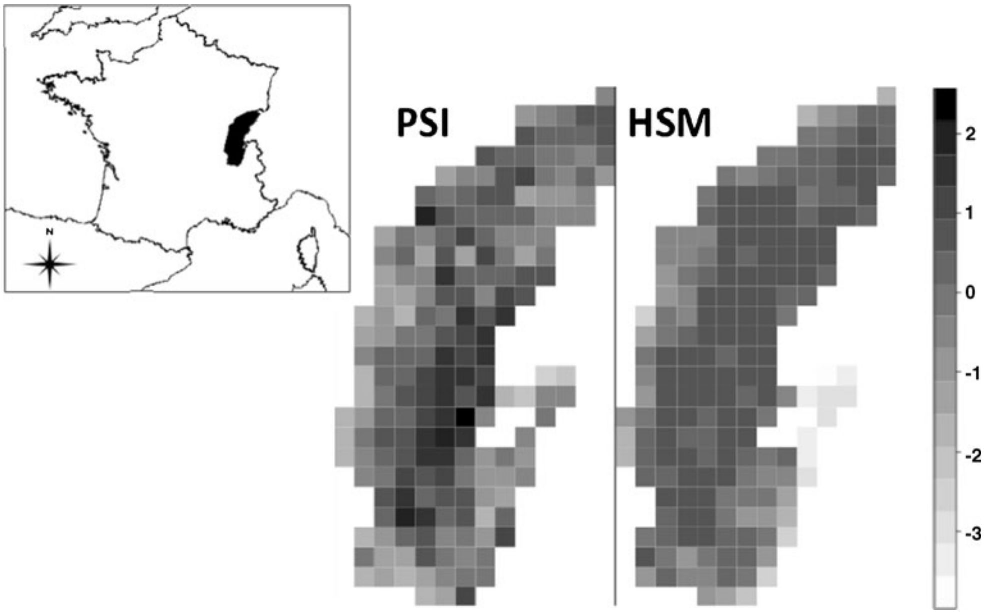

Fig. 1 Occupancy map estimated using the best site-occupancy model (*left*), i.e. ψ(forest) p(forest + PC1) and HSM drawn from the MADIFA (*right*). Both maps represent an estimation of the probability of presence of the lynx for each cell. To improve clarity, both maps were normalised and HSM scores were multiplied by −1; in both cases, *darker colours* correspond to a higher suitability. Each pixel is 9×9 km^2

species' optimum. All lynx signs (2002–2006) were used to build the HSM.

Using the site-occupancy modelling, we estimated the occupancy ψ in each cell (i.e. the probability of presence, given its history) and the detection probability p. The presence of lynx was defined as the occurrence of at least one lynx sign in a cell during 1 year. We fitted a model with both ψ and p constant ("$\psi(.)p(.)$") and a model with ψ constant and p varying over time ("$\psi(.)p(t)$") to take into account the potential variation of detection probability over time. We then tested for the effect of environmental covariates on both detection and occupancy probabilities at each cell using a logistic model as described in MacKenzie (2005). As environmental variables were intercorrelated, we computed a principal component analysis (PCA) with the seven variables. We used the first axis (PC1), which was mostly driven by river and road densities; and the proportion of forest cover that was the main variable involved in PC2 and thereby statistically independent of PC1.We used an information–theoretic framework based on Akaike information criterion (AIC, Burnham and Anderson 2002) to rank all models and select the one with the best compromise between deviance and complexity.

Finally, we calculated the correlation between MADIFA scores and occupancy probabilities using a Pearson's correlation test based on paired values for each cell. The site-occupancy modelling was computed using PRESENCE 2.0 (MacKenzie 2005); the MADIFA and all other statistical analyses were conducted using R 2.10.0 (R Development Core Team 2010) and the R package "adehabitat" (Calenge 2006).

Results

The first three axes of the MADIFA accounted for 67% of the Mahalanobis distances and were kept for the computation of the HSM. Both high road and river densities led to lower habitat suitability on the first two axes of the analysis, while higher proportions of forest indicated higher values of habitat suitability on the third axis.

The site-occupancy model with a constant detection probability received more support than the model including a time-dependent detection (Table 1), indicating that the detection can be considered as constant over the entire period. Models including covariates had also more support than the reference models $\Psi(.)p(t)$ and $\Psi(.)p(.)$, indicating that environmental covariates markedly influenced occupancy. Among them, two models had AIC values within 2 units, the one including both PC1 and the proportion of forest cover, and the one including only the proportion of forest cover. Based on parsimony rules, the latter model should be selected.

Occupancy maps (i.e. estimation of Ψ) from both models were correlated with the HSM drawn using the MADIFA ($R^2>0.25$, all $p<0.001$, Table 1). Therefore, both analytical approaches provided strongly convergent assessment of lynx occurrence (Fig. 1).

Discussion

In the case of rare or elusive species, "profile methods" (sensu Pearce and Boyce 2006) like the MADIFA allow computing habitat suitability maps that are biologically coherent with former knowledge of the species (e.g. Bryan and Metaxas 2007). Occupancy modelling provides a suitable alternative to the difficulties of estimating presence by taking into account detection probability less than one (MacKenzie 2005). The main factors structuring lynx occurrence identified in that study (i.e. road and river density, and proportion of forest) were consistent with previous findings on lynx ecological requirements (Vandel et al. 2006). The spatial patterns of lynx occurrence as assessed from the MADIFA and site-occupancy modelling were remarkably similar as illustrated in Fig. 1. In particular, the

Table 1 Modelling site occupancy with environmental covariates

Models	AIC	Delta AIC	AIC weight	Ranking of the combination of covariates	R^2 correlation between HSM and MADIFA
ψ(forest + PC1)p(forest + PC1)	1,015.10	0	0.531	1	0.280
ψ(forest)p(forest + PC1)	1,014.41	0.31	0.454	2	0.500
ψ(forest + PC1)p(forest)	1,022.90	7.80	0.0107	3	0.222
ψ(forest)p(forest)	1,025.40	10.30	0.0031	4	0.500
ψ(forest)p(PC1)	1,028.31	13.21	0.0007	5	0.502
$\psi(.)p(.)$	**1,058.12**	43.02	0		
$\psi(.)p(t)$	**1,060.90**	45.80	0		

Models without any environmental covariate are bolded as references

forest proportion of forest, *PC1* first axis of the PCA (see text for further details), *AIC* Akaike information criterion, *HSM* habitat suitability maps, *MADIFA* Mahalanobis distance factor analysis probability of lynx occurrence was lower at the periphery of the study area than in the centre, likely in relation with the presence of high human presence (and particularly roads) when we get closer to Lake Léman. The negative influence of road density highlighted by both approaches supports a negative effect of human presence on lynx occurrence. On the other hand, the probability of lynx occurrence steadily increases in the centre of the Jura Mountains, indicating that the well-trained observers were able to maintain a high detection probability even in remote areas.

Detection probabilities were also constant over time, likely because of a constant sampling effort. Such a constancy of detection probability confirmed that the study period considered corresponds to a closure period well suited for the use of the site-occupancy modelling. Our results underlined that both analyses, based on the same environmental covariates, yielded to convergent qualitative conclusions about the probability of presence. Therefore, our study provides the first empirical evidence that profile and modelling approaches may give similar assessment of the occurrence of an elusive species such as the lynx.

Profile methods provide a qualitative assessment of the most structuring factors of species distribution and can be used as explorative analyses. On the other hand, site-occupancy modelling uses a quantitative measure (i.e. AIC) to select the relevant environmental factors that shape the species distribution. Their concurrent use could lead to finer approaches which are able to estimate the presence of an elusive species like lynx with increased accuracy. We thus recommend performing first a MADIFA to disentangle the structuring covariates before using the site-occupancy approach for checking model robustness given the factors included.

Acknowledgements We thank Mark Kéry and two anonymous referees for their helpful comments that markedly improved our paper.

References

Austin M (2007) Species distribution models and ecological theory: a critical assessment and some possible new approaches. Ecol Modell 200:1–19

Basille M, Calenge C, Marboutin É, Andersen R, Gaillard J-M (2008) Assessing habitat selection using multivariate statistics: some refinements of the ecological-niche factor analysis. Ecol Modell 211:233–240

Bryan T, Metaxas A (2007) Predicting suitable habitat for deep-water gorgonian corals on the Atlantic and Pacific continental margins of North America. Mar Ecol Prog Ser 330:113–126

Burnham K, Anderson D (2002) Model selection and multi-model inference: a practical information-theoretic approach. Springer, Heidelberg, pp. 514

Calenge C (2006) The package ”adehabitat” for the R software: a tool for the analysis of space and habitat use by animals. Ecol Modell 197:516–519

Calenge C, Basille M (2008) A general framework for the statistical exploration of the ecological niche. J Theor Biol 252:674–685

Calenge C, Dufour A, Maillard D (2005) K-select analysis: a new method to analyse habitat selection in radio-tracking studies. Ecol Modell 186:143–153

Calenge C, Darmon G, Basille M, Loison A, Jullien J (2008) The factorial decomposition of the Mahalanobis distances in habitat selection studies. Ecology 89:555–566

Chao A (1987) Estimating the population size for capture–recapture data with unequal catchability. Biometrics 43:783–791

Hirzel AH, Hausser J, Chessel D, Perrin N (2002) Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? Ecology 83:2027–2036

Linnell J, Swenson J, Landa A, Kvam T (1998) Methods for monitoring large carnivores: a worldwide review of relevant experience. NINA Oppdragsmeld 549:1–38

MacKenzie DI (2005) Was it there? Dealing with imperfect detection for species presence/absence data. Aust N Z J Stat 47(1):65–74

MacKenzie DI, Nichols JD, Royle JA, Pollock KH, Bailey LL, Hines JE (2006) Occupancy estimation and modelling: inferring patterns and dynamics of species occurrence. Academic Press, New York

Marboutin E, Duchamp C, Rouland P, Leonard Y, Boyer J, Michallet D, Catusse M, Migot P, Vandel J-M, Stahl P (2006) Survey of the lynx distribution in the French Alps: 2000–2004 population status analysis. Acta Biol Slov 49(1):19–26

Pearce JL, Boyce MS (2006) Modelling distribution and abundance with presence-only data. J Appl Ecol 43:405–412

R Development Core Team (2010) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Vandel J-M, Stahl P (2005) Distribution trend of the Eurasian Lynx lynx populations in France. Mammalia 69:145–158

Vandel J-M, Stahl P, Herrenschmidt V, Marboutin E (2006) Reintroduction of the lynx into the Vosges mountain massif: from animal survival and movements to population development. Biol Conserv 131:370–385