

Spatial replicates as an alternative to temporal replicates for occupancy modelling when surveys are based on linear features of the landscape

Anaïs Charbonnel, Frank d'Amico, Aurélien Besnard, Frédéric Blanc, Laetitia Buisson, Mélanie Némoz, Pascal Laffaille

▶ **To cite this version:**

Anaïs Charbonnel, Frank d'Amico, Aurélien Besnard, Frédéric Blanc, Laetitia Buisson, et al.. Spatial replicates as an alternative to temporal replicates for occupancy modelling when surveys are based on linear features of the landscape. Journal of Applied Ecology, 2014, vol. 51 (n° 5), pp. 1425-1433. 10.1111/1365-2664.12301 . hal-01494958

HAL Id: hal-01494958
https://hal.science/hal-01494958

Submitted on 24 Mar 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Open Archive TOULOUSE Archive Ouverte (OATAO)
OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : http://oatao.univ-toulouse.fr/
Eprints ID : 13563

To link to this article : DOI : 10.1111/1365-2664.12301
URL : http://dx.doi.org/10.1111/1365-2664.12301

To cite this version : Charbonnel, Anaïs and D'Amico, Frank and Besnard, Aurélien and Blanc, Frédéric and Buisson, Laetitia and Némoz, Mélanie and Laffaille, Pascal *Spatial replicates as an alternative to temporal replicates for occupancy modelling when surveys are based on linear features of the landscape.* (2014) Journal of Applied Ecology, vol. 51 (n° 5). pp. 1425-1433. ISSN 0021-8901

Any correspondance concerning this service should be sent to the repository administrator: staff-oatao@listes-diff.inp-toulouse.fr

Spatial replicates as an alternative to temporal replicates for occupancy modelling when surveys are based on linear features of the landscape

Anaïs Charbonnel[1,2,3]*, Frank D'Amico[4,5], Aurélien Besnard[6], Frédéric Blanc[1], Laëtitia Buisson[2,3], Mélanie Némoz[1] and Pascal Laffaille[2,7]

[1]*Conservatoire d'Espaces Naturels Midi-Pyrénées, 75 voie du TOEC, BP 57611, 31076 Toulouse, France;* [2]*CNRS, UMR 5245, EcoLab (Laboratoire Ecologie Fonctionnelle et Environnement), 31326 Castanet-Tolosan, France;* [3]*Université de Toulouse, INP, UPS, EcoLab (Laboratoire Ecologie Fonctionnelle et Environnement), 118 Route de Narbonne, 31062 Toulouse, France;* [4]*Université de Pau et des Pays de l'Adour, UFR Sciences & Techniques Côte Basque, Campus Montaury, 64600 Anglet, France;* [5]*Université de Pau, Laboratoire de Mathématiques et de leurs Applications, LMAP CNRS-UMR 5142, Av. de l'Université, 64000 Pau, France;* [6]*Laboratoire de Biogéographie et d'Ecologie des Vertébrés, Ecole Pratique des Hautes Etudes, Centre d'Ecologie Evolutive et Fonctionnelle, CNRS UMR 5175, 1919, Route de Mende, 34293 Montpellier Cedex 5, France; and* [7]*Université de Toulouse, INP, UPS, EcoLab (Laboratoire Ecologie Fonctionnelle et Environnement); ENSAT, Avenue de l'Agrobiopole, 31326 Castanet-Tolosan, France*

Summary

1. Occupancy estimates can inform biodiversity managers about the distribution of elusive species, such as the Pyrenean desman *Galemys pyrenaicus*, a small semi-aquatic mammal that lives along streams. Occupancy models rely on replication within a sampling site and provide estimates of the probability of detection. However, we still do not know how occupancy and detection estimates obtained from spatial vs. temporal replications differ or the appropriateness of using one or the other when cost and logistics make one approach prohibitive. Recently, the Markovian occupancy model has been developed to analyse adjacent spatial replicates and to test for spatial dependence between them. This model has already been applied to large and highly mobile mammals using trails, but never tested for any species with linear home ranges.

2. We compared detection and occupancy estimates obtained from both temporal and spatial sampling designs that were subsequently organized into four data configurations (sites with both spatial and temporal replicates, adjacent spatial replicates only, temporal replicates only at the segment and site scales). From that, five occupancy models with different assumptions (the standard occupancy model, the standard multiscale model, the multiscale model with Markovian process for detection, the Markovian detection model and the Markovian occupancy model) were used. We also assessed which occupancy model was the most appropriate for each data configuration to determine whether it is necessary to incorporate correlation into models.

3. We found that the estimated detection probabilities were relatively high ($\geq$0·58) and similar when the same model was applied to each data configuration.

4. Spatial replication weakly underestimated occupancy. But when using this design, the Markovian occupancy model was the most supported and minimized the underestimation of occupancy, highlighting a spatial dependence between adjacent replicates.

5. *Synthesis and applications*. We show that a survey based on adjacent spatial replicates for a mammal living along linear features of the landscape is a good compromise between cost and occupancy estimates, while using the Markovian occupancy model to estimate detection and occupancy. Our finding may have wider applications for the monitoring of species especially

*Correspondence author. E-mail: anais.charbonnel@espaces-naturels.fr

when temporal replicates are difficult or unrealistic. Spatial design, for surveys based on sign detection, could thus be applied for species with linear home ranges or when surveys are constrained by linear habitats.

Key-words: autocorrelation, detection probability, *Galemys pyrenaicus*, Markovian model, multiscale model, sign survey

Introduction

Presence–absence data are widely used to determine the habitat requirements of species or predict their distribution in unexplored areas (Guisan & Zimmermann 2000). One widely acknowledged major issue is that false absences are unavoidable during fieldwork (Gu & Swihart 2004; Gibson, Barrett & Burbidge 2007; Lobo, Jiménez-Valverde & Hortal 2010). In species distribution modelling, failure to take imperfect detection into account can introduce a considerable degree of error into spatial and temporal distribution patterns (Yoccoz, Nichols & Boulinier 2001). For example, false conclusions can be reached about the suitability of areas for conservation actions (Rondinini *et al*. 2006). Correctly accounting for false absences is thus important, especially for the conservation of cryptic and secretive species (e.g. Gibson, Barrett & Burbidge 2007; Durso, Willson & Winne 2011), for which surveys often rely on recording indirect signs.

Over recent decades, many methods have been developed to deal with detection issues and to correct the biases they induce, for example distance sampling (Buckland *et al*. 2001) or capture–recapture (Pollock *et al*. 2002). More recently, MacKenzie *et al*. (2002, 2006) developed site occupancy modelling (hereafter referred to as 'standard site occupancy modelling'), a statistical method that models the probability that a species occupies some sites even though it has not been detected with any certainty. This method is increasingly being successfully used to model the distribution of various species (e.g. Kéry, Gardner & Monnerat 2010; Olea & Mateo-Tomás 2011; Kéry, Guillera-Arroita & Lahoz-Monfort 2013). Site occupancy modelling requires replication of detection–non-detection data at sampled sites. Replications at site level are usually temporal (i.e. the sites are monitored over several successive visits). They may, however, also be spatial (i.e. sampling several spatial subunits at each site during a single visit).

Standard site occupancy modelling (MacKenzie *et al*. 2002, 2006) relies on the assumption that there is no change in site occupancy during the survey (Rota *et al*. 2009). Violation of this assumption of closure leads to an underestimation of the detection probability and an overestimation of the occupancy. When temporal replicates are used, successive visits must therefore be conducted over a relatively short period. When spatial replicates are substituted for temporal ones, all the subunits must have a non-negligible occupancy probability. Thus, the sampling design must take into account the movement patterns of a species to adhere to the model assumptions (Kendall & White 2009; Guillera-Arroita 2011). Nichols *et al*. (2008) also developed a multiscale occupancy model (hereafter referred as 'standard multiscale occupancy model') which can be used when both spatial and temporal replicates are available for the same sites. This model relaxes the closure assumption by dealing with the situation in which the species is present at a site, but may be absent from some survey subunits.

Besides the closure assumption, for the standard site occupancy model (MacKenzie *et al*. 2002, 2006), the detection events of the species must be independent. In the context of temporal replicates, this assumption may be violated, for instance, if the same observer visits the same site several times and remembers where to look for signs of the species' presence (observer bias hereafter). In the case of spatial replicates, dependence could arise if the subunits are not randomly selected (e.g. if they are spatially close together or sequentially sampled; Hines *et al*. 2010; Anderson *et al*. 2012) and the observer becomes more attentive after the first detection event. In such situations, the spatial or temporal correlation between detection events can be explicitly modelled using a Markovian detection model (Hines *et al*. 2010). When data are collected using a spatial cluster of sampling subunits, spatial autocorrelation can also arise between adjacent replicates with regard to species' presence. To deal with such situations, Hines *et al*. (2010) developed a Markovian occupancy model.

To our knowledge, only one study has directly compared the parameter estimates of occupancy and detection of temporal vs. spatial replicates (Parry *et al*. 2013). Our work was prompted by a field survey of a small semi-aquatic mammal, the Pyrenean desman *Galemys pyrenaicus* E. Geoffroy Saint-Hilaire, 1811. Repeated visits to stream reaches in the Pyrenees to get temporal replicates are time-consuming and expensive, because of being frequently difficult to access, and they have to be sampled meticulously by wading. Spatial replicates therefore offer an appealing, cost-effective alternative. Since sites are river transects, the least expensive approach is to use successive portions of rivers as adjacent spatial replicates. The Markovian occupancy model, originally developed for large and highly mobile mammals using trails such as tigers *Panthera tigris*, may also be the most appropriate method to model data collected along river transects (Hines *et al*. 2010; Karanth *et al*. 2011). The goals of our study were to (i) directly compare occupancy and detection estimates obtained from temporal vs. spatial design, and (ii) evaluate whether the

Markovian occupancy model (with autocorrelation) is supported when spatial adjacent replicates are used, for a species with a small linear home range.

Materials and methods

STUDY AREA

The French Pyrenees (W1°40′–E3°10′, N43°08′–N42°23′) are a range of mountains approximately 400 km long, covering an area of 18 176 km² and reaching a maximum elevation of 3298 m (Fig. 1). Climate and topography are the main factors influencing the dynamics and flow regime of Pyrenean rivers.

BIOLOGICAL MODEL

The Pyrenean desman is a small, semi-aquatic mammal with a range restricted to the Pyrenees (Andorra, France and Spain) and some areas in northern and central Spain, and northern Portugal. It lives in mountain brooks, cold and well-oxygenated water courses, and its altitudinal range extends from 15 m (western French Pyrenees) or 450 m (eastern French Pyrenees) to 2700 m (Némoz & Bertrand 2008). The Pyrenean desman has declined substantially in its distribution, triggering several conservation regulations (Fernandes *et al.* 2008). Like many secretive species, searching for faeces is the most efficient and least cost- and time-consuming method for surveying this species across a wide geographic extent (e.g. Kindberg, Ericsson & Swenson 2009). Being elusive and cryptic, its ecology is still largely unknown. It is not yet understood whether this species is territorial or why it leaves its faeces on emergent items.

FIELD SURVEY PROTOCOL

Thirty-four sites (i.e. river transects) were surveyed for this study (Fig. 1). Previous sampling had reported presence of the Pyrenean desman in all these sites, as it is more efficient to select sites where the animal is known to be present than a random selection when modelling detection probabilities (MacKenzie *et al.* 2002). Skilled observers waded river transects to search for faeces. They meticulously inspected each emergent rock, tree root or branch in the stream. Pyrenean desman faeces were identified from their colour, size, position and smell. Surveys were conducted during the summer, when faeces seem to persist longest (Bertrand 1994), to maximize detectability. The same observer visited five times each site, between June 7th and October 5th 2012, to obtain temporal replication data over one season. For each site, visits were at least 1 week apart to limit observer bias. To maximize the abundance of available signs, we avoided surveys during or after a period of fluctuating water levels or heavy rainfall (Bertrand 1994).

Each site was a riverbed transect, 500 m long, which approximately matches the mean home range of the species determined by radiotracking (523 m; SE 50·85; Melero *et al.* 2012). Each sampling site had five subunits (i.e. segments) of equal length (100 m) that constituted the adjacent spatial replicates. As ecological knowledge for the Pyrenean desman is limited (Némoz *et al.* 2011), the optimum segment length was estimated. We chose a length of 100 m as the best compromise between the need to detect signs when animals are in fact present and that of having more than three spatial replicates, as recommended by Hines *et al.* (2010).

SITE OCCUPANCY ANALYSIS

The data configurations used in the analyses are shown in Fig. 2: the first had both spatial and temporal replicates at the 34 sites (Fig. 2a), the second had spatial replicates during a single visit (Fig. 2b), the third had temporal replicates for a single segment (Fig. 2c), and the fourth had temporal replicates at the site level (i.e. segments pooled; Fig. 2d). The multiscale models (Pavlacky *et al.* 2012) were fitted to both spatial and temporal replicates, that is the first data configuration. We fitted the standard site occupancy model (MacKenzie *et al.* 2002), the Markovian detection model and the Markovian occupancy model (Hines *et al.* 2010) to the last three data configurations.

Standard occupancy model – $\psi(.)$, $p(.)$

This modelling approach estimates two key parameters: p, the probability of detecting the species if the species is in fact present on the site (i.e. the probability of detection), and ψ, the probability that a site is occupied or used by a species (i.e. the probability of site occupancy).

Standard multiscale occupancy model – $\psi(.)$, $\theta(.)$, $p(.)$

This model has been used for presence–absence data collected at two spatial scales (Pavlacky *et al.* 2012). N sample sites are subsampled by R spatial segments (spatial replicates). All sampled

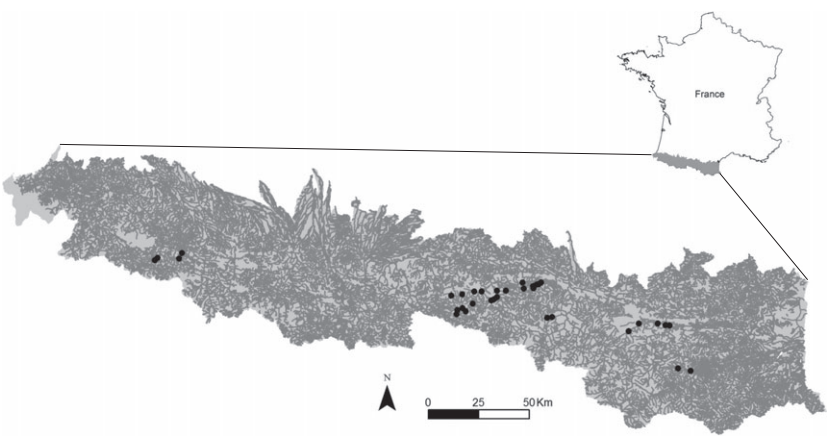


Fig. 1. Study area and location of the sampling sites (black dots, $N = 34$). Grey lines indicate the stream network.

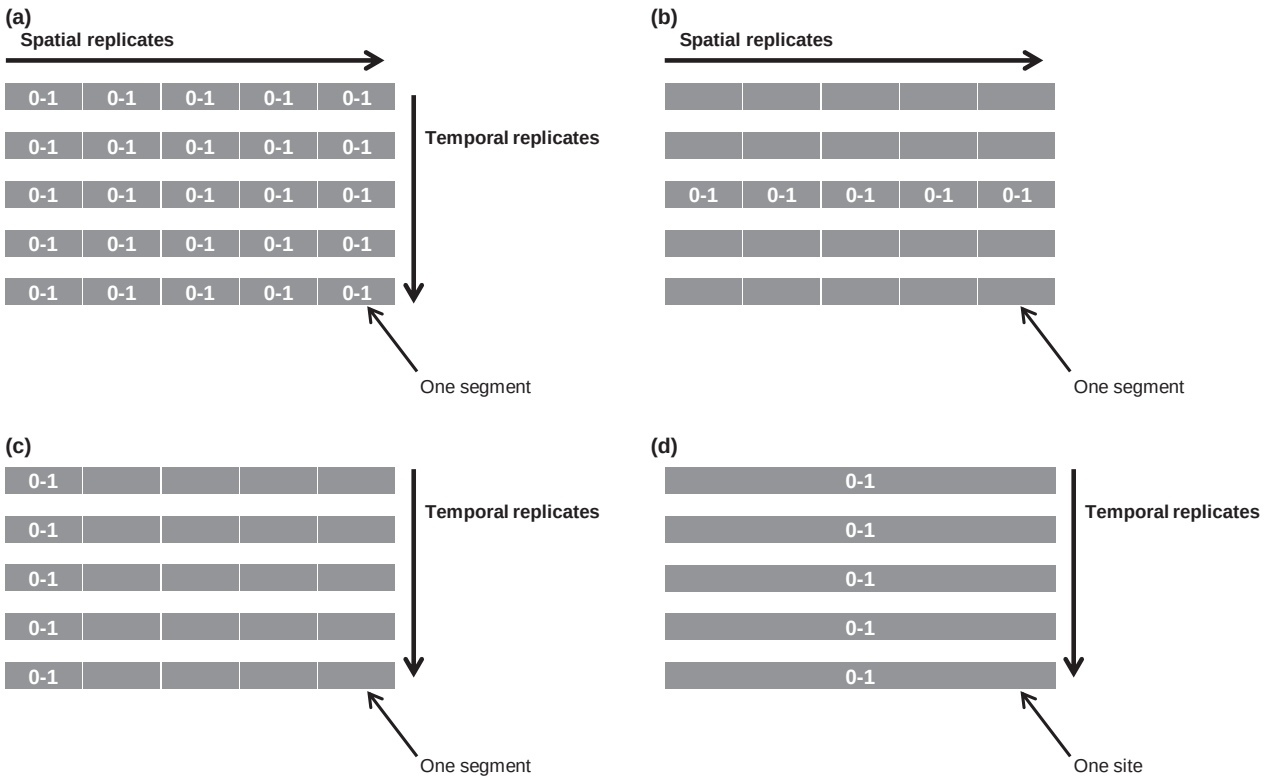


Fig. 2. Four different data configurations used in the analysis for one sampling site: (a) spatial and temporal replicates; (b) spatial replicates (i.e. only one sampling visit is studied, here the third one); (c) temporal replicates for segments (i.e. only one segment is studied, here the first one); and (d) temporal replicates for the site (i.e. data of the five segments are pooled). Data collected are detection (1) or non-detection (0) of faeces of the Pyrenean desman.

sites (and consequently all segments) are also repeatedly visited (temporal replicates). The parameters of the model are: p, the probability of detection on a segment at one visit if the sample site and the segment are in fact occupied; θ, the probability of occupancy for one survey segment if the sampling site is in fact occupied; and ψ, the probability of occupancy of a sampling site. The small scale occupancy, θ, is an availability parameter to account for situations in which the species is present in the sampling site and may be present in some survey segments but not in others.

Multiscale occupancy model with Markovian process for detection – $\psi(.)$, $\theta(.)$, $p_0(.)$, $p_1(.)$

This multiscale model takes into account the autocorrelation in detection events between spatial or temporal occasions. Two parameters are linked to detection: p_0, the probability of detection on one occasion if the site is in fact occupied and the species was not detected on the previous occasion, and p_1, the probability of detection on one occasion if the site is in fact occupied and the species was detected on the previous occasion.

Markovian detection model – $\psi(.)$, $p_0(.)$, $p_1(.)$

This model is derived from the standard occupancy model (MacKenzie *et al.* 2002) to account for autocorrelation in detection between occasions. Detection is modelled as an observable Markov process, which allows the detection probability for each occasion to depend on whether or not the species was detected on the previous occasion. The two parameters linked to detection, p_0 and p_1, are similar to those presented for the multiscale occupancy model with Markovian process for detection.

Markovian occupancy model – $\psi(.)$, $\theta_0(.)$, $\theta_1(.)$, $p(.)$

This model was developed for adjacent spatially autocorrelated selected replicates (i.e. segments; Hines *et al.* 2010). This model assumes spatial closure (i.e. occupancy of one spatial replicate assumes that the other spatial replicates are similarly occupied; Hines *et al.* 2010). That is why this model design was specifically developed for highly mobile species, and the spatial scale of the sampling sites must be the same as the home range of the species (or as close to it as possible). In addition to the occupancy probability measured at the site level (ψ) and the detection probability measured at the segment level (p), two probabilities are estimated: θ_0, the probability that a species is present on a segment if the site is actually occupied and the species was absent on the previous segment, and θ_1, the probability that a species is present on the segment if the site is occupied and the species was present on the previous segment. This model could also be useful in occupancy studies that use temporal replication to test for the temporal pattern of local occupancy (Hines *et al.* 2010).

MODEL COMPARISON AND EVALUATION

We used Akaike information criterion (AIC) to assess the relative fit of the models within each candidate data configuration.

Table 1. Occupancy and detectability estimates obtained with the five models (SE: standard error). The two multiscale models [ψ(.), θ(.), p_0(.), p_1(.)] and [ψ(.), θ(.), p(.)] were applied on both spatial and temporal replicates (Fig. 2a). The Markovian occupancy model [ψ(.), θ_0(.), θ_1(.), p(.)], the Markovian detection model [ψ(.), p_0(.), p_1(.)] and the standard occupancy model [ψ(.), p(.)] were applied on spatial replicates (Fig. 2b), temporal replicates of a single segment (Fig. 2c) and temporal replicates at the site level (i.e. segments pooled; Fig. 2d). The relative differences in AIC values compared with the top ranked model (ΔAIC) and AIC model weights (W) are given. For each data configuration, models are ranked according to their ΔAIC, the best model being the top ranked. No confidence intervals (CI) were derived when the estimate equalled 1

			Detection probabilities						Occupancy probabilities							
			p		p_0		p_1		θ		ψ		θ_0		θ_1	
Models	ΔAIC	W	Estimate (SE)	95% CI	Estimate (SE)	95% CI	Estimate (SE)	95% CI	Estimate (SE)	95% CI	Estimate (SE)	95% CI	Estimate (SE)	95% CI	Estimate (SE)	95% CI
Spatial and temporal																
Spatial correlation																
ψ(.), θ(.), p_0(.), p_1(.)	0·00	1·00	–	–	0·40 (0·03)	0·34–0·46	0·82 (0·02)	0·77–0·86	0·76 (0·04)	0·68–0·82	1·00	–	–	–	–	–
Temporal correlation																
ψ(.), θ(.), p_0(.), p_1(.)	44·92	0·00	–	–	0·41 (0·04)	0·34–0·48	0·75 (0·02)	0·70–0·79	0·77 (0·04)	0·68–0·84	1·00	–	–	–	–	–
Standard																
ψ(.), θ(.), p(.)	103·84	0·00	0·60 (0·02)	0·56–0·64	–	–	–	–	0·72 (0·03)	0·65–0·79	1·00	–	–	–	–	–
Spatial																
ψ(.), θ_0(.), θ_1(.), p(.)	0·00	0·60	0·93 (0·11)	0·33–0·99	–	–	–	–	–	–	0·85 (0·18)	0·26–0·99	0·24 (0·19)	0·04–0·70	0·85 (0·09)	0·59–0·95
ψ(.), p_0(.), p_1(.)	0·80	0·40	–	–	0·45 (0·08)	0·31–0·61	0·79 (0·05)	0·68–0·88	–	–	0·77 (0·08)	0·57–0·90	–	–	–	–
ψ(.), p(.)	12·56	0·00	0·64 (0·04)	0·55–0·72	–	–	–	–	–	–	0·74 (0·08)	0·57–0·86	–	–	–	–
Temporal (segment)																
ψ(.), p_0(.), p_1(.)	0·00	0·84	–	–	0·38 (0·08)	0·24–0·55	0·73 (0·06)	0·60–0·83	–	–	0·81 (0·10)	0·55–0·93	–	–	–	–
ψ(.), θ_0(.), θ_1(.), p(.)	3·43	0·15	1·00	–	–	–	–	–	–	–	0·83 (0·11)	0·51–0·96	0·34 (0·10)	0·17–0·55	0·73 (0·06)	0·60–0·83
ψ(.), p(.)	10·45	0·00	0·58 (0·05)	0·48–0·66	–	–	–	–	–	–	0·75 (0·08)	0·57–0·87	–	–	–	–
Temporal (site)																
ψ(.), p_0(.), p_1(.)	0·00	0·88	–	–	0·55 (0·06)	0·43–0·66	0·82 (0·04)	0·74–0·89	–	–	1·00	–	–	–	–	–
ψ(.), θ_0(.), θ_1(.), p(.)	3·97	0·12	1·00	–	–	–	–	–	–	–	1·00	–	–	–	–	–
ψ(.), p(.)	13·40	0·00	0·71 (0·03)	0·63–0·77	–	–	–	–	–	–	1·00	–	–	–	–	–

p: detection probability; p_0: detection probability without detection event on the previous occasion; p_1: detection probability with detection event on the previous occasion; θ: local occupancy probability (i.e. availability parameter); Ψ: occupancy probability; θ_0: presence probability with absence on the previous occasion; θ_1: presence probability with presence on the previous occasion.

Models were ranked, with the lowest AIC value indicating the best fit to the data. We did not evaluate the value of the AIC *per se*, but the difference between the value for the model of interest and that for the best-supported model (Δi; Burnham & Anderson 2002). A fitted model with a Δi value of less than two has more substantial support for the data than the other models. The relative differences in AIC values for the fitted models can be used to calculate the weight (Wi) of evidence of each model in the model set. We fitted all the models using the freeware PRESENCE v. 5.5 (Hines 2006).

Results

For spatial replicates and temporal replicates, we applied analyses successively to the five visits and segments, respectively. Given that the trends were similar, we only give here the results for the third visit (spatial replicates) and the first segment (temporal replicates) (see Tables S1–S3, Supporting Information for the other visits and segments).

MODEL SELECTION

In all cases, the standard occupancy model was the least-supported model according to the AIC, with a ΔAIC greater than 10 relative to the best model and with a null AIC weight (Table 1). The multiscale model, which took spatial dependence of detection into account, was the most-supported model (i.e. spatial and temporal replicates; W = 1; Table 1). For temporal replicates at both spatial levels (site and segment), comparison of AIC values indicated that the best-supported model was the Markovian detection model (W = 0·88 and W = 0·84 for site and segment, respectively), followed by the Markovian occupancy model (W = 0·12 and W = 0·15 for site and segment, respectively; Table 1).

This result suggests a possible observer bias between detection events, with p_0 (SE, standard error) = 0·55 (SE 0·06) and p_1 = 0·82 (SE 0·04) and p_0 = 0·38 (SE 0·08) and p_1 = 0·73 (SE 0·06) at the site and segment level, respectively (Table 1). For spatial replicates, the Markovian occupancy model was more highly supported than other models (W = 0·60; Table 1), suggesting that spatial autocorrelation existed in the presence of desman faeces between segments and that this should be modelled, with θ_0 = 0·24 (SE 0·19) and θ_1 = 0·85 (SE 0·09; Table 1).

PROBABILITY ESTIMATES

Site occupancy

The naïve site occupancy (i.e. the proportion of sampling sites with at least one detection event during the survey) was *a priori* known to be equal to one; hence, the site occupancy probability (ψ) estimated by all models applied to temporal replicates and to temporal and spatial replicates was, not surprisingly, equal to one. For the spatial replicates only, naïve occupancy was equal to 0·74, because some sites had no detection event during this sampling visit (the third one). Hence, estimated values of ψ were lower for all models applied to spatial replicates (Table 1). ψ was highest when fitted using the Markovian occupancy model (ψ = 0·85; SE 0·18). However, large standard error estimates made it impossible to make strong inferences.

Segment occupancy

On the local scale, estimates of segment occupancy (θ or ψ) were roughly similar between all analyses and models, with ψ or θ ranging from 0·72 (SE 0·03) to 0·83 (SE 0·11; Table 1). For spatial and temporal replicates, the most-supported multiscale model which accounts for spatial correlation in detection events estimated a θ value of 0·76 (SE 0·04). For temporal replicates, ψ was higher when correlation for detection [ψ(.), p_0(.), p_1(.)] or presence [ψ(.), θ_0(.), θ_1(.), p(.)] was accounted for, with ψ = 0·81 (SE 0·10) and 0·83 (SE 0·11), for each model, respectively (Table 1).

Detectability at the site level

For temporal replicates, the Markovian occupancy model gave an estimate of detection probability equal to one (Table 1). The standard model, which was the least-supported model according to the AIC, estimated a detection probability that was 29% lower (p = 0·71; SE 0·03).

Detectability at the segment level

The Markovian occupancy model applied to different data configurations (spatial vs. temporal replicates) gave roughly similar estimates. With this model, p ranged from 0·93 (SE 0·11) to 1·00, whereas the standard occupancy model gave estimates ranging from 0·58 (SE 0·05) to 0·64 (SE 0·04; Table 1) depending on data configurations. These estimates were comparable to those at the site level. Models that took into account the autocorrelation between detection events suggested that the detection probability was lower when there was no detection on the previous occasion than when the species had been detected on the previous occasion (Table 1). For example, for the Markovian detection model applied to spatial replicates, p_0 = 0·45 (SE 0·08) and p_1 = 0·79 (SE 0·05).

Discussion

OCCUPANCY AND DETECTABILITY ESTIMATES

When the same occupancy model was fitted to several different data configurations, we found that the estimated detection probabilities were relatively similar, whatever the replication design (i.e. spatial or temporal) and the scale (i.e. site or segment) used. This indicates that the detectability of the Pyrenean desman in this data set appears to be fairly constant when a single segment of

100 m was surveyed on five successive occasions, when a site of 500 m was surveyed on five successive occasions or when five adjacent 100-m segments were surveyed on one occasion. The most parsimonious model was the Markovian detection model with temporal replication, while the Markovian occupancy model was the most supported with only spatial replicates. Detection probabilities were high using the Markovian occupancy model (0·91–1), which is likely to be more accurate than the standard occupancy model (0·58–0·71). These results are consistent with those of Hines *et al.* (2010) showing that species detectability is underestimated in the presence of dependence between sampling occasions with the standard occupancy model.

The probabilities of detection estimated for the Pyrenean desman were relatively high ($P \geq 0{\cdot}58$), but their precision and inferences are limited by the small size of the samples collected during this study (34 sites). The Pyrenean desman seems to have moderate to high detectability when its faeces are present on emergent rocks, roots or branches in the stream. In comparison, the probability of detecting the Eurasian otter *Lutra lutra* on the basis of spatial replicates (600-m stream transects for sites and 50-m segments for subunits) has been estimated to be 0·26 on the Gower peninsula (UK; Parry *et al.* 2013). Given that the lowest estimated probability of detection for the Pyrenean desman equals 0·58, only five sampling occasions will be sufficient to get an almost perfect probability of detection ($P = 1-(1-0{\cdot}58)^5 = 0{\cdot}99$) when Pyrenean desman signs are in fact present on the sampled site. However, we acknowledge that the sites sampled in this study may contain a high abundance of signs of the Pyrenean desman since they were already known to be occupied prior to sampling. If sign abundance positively influences sign detectability, the latter may be overestimated in this study (McCarthy *et al.* 2012).

Our results show that temporal replication at site level is more accurate than spatial replication for estimating occupancy which is not surprising as all sites (but not all segments) have at least one detection event among the five visits. This is not necessarily the case when using spatial replication because some sites may have no detection event during the selected sampling visit (the third one here but see Tables S2–S3, Supporting Information). When using temporal replication at site level, all occupancy models (i.e. Markovian occupancy model, Markovian detection model and standard occupancy model) do not appear to underestimate occupancy (equalling 1), while temporal sampling restricted to a 100-m segment underestimates occupancy (0·81 for the Markovian detection model which is the most-supported model). Estimates of occupancy with spatial adjacent replicates are also underestimated in a magnitude comparable to those estimated when successively sampling a 100-m segment (0·77 for the Markovian detection model). However, when using adjacent spatial replicates, the Markovian occupancy model is the most appropriate and minimizes the underestimation of occupancy (0·85) which becomes even higher than temporal replicates at the segment level. Thus, there appears to be a trade-off between spending more time sampling with temporal replication to achieve a more accurate estimate of occupancy, or sampling with adjacent spatial replicates which is more cost-effective but with a slight underestimation of site occupancy.

For local occupancy (i.e. segment level), when signs of the Pyrenean desman are present at a site, they occur, on average, in 75% of the segments. The θ_0 and θ_1 estimates suggest that spatial autocorrelation between segments is very high. Indeed, when no species signs have been detected on a given segment, the next segment has only a probability of 0·24 signs being present. In contrast, the probability of occupancy of a given segment is 0·85 when signs of presence have been detected on the previous segment. It is still not understood why the Pyrenean desman leaves its faeces on emergent items and how it uses its home range. This finding suggests that signs of the Pyrenean desman are clustered within sites, with some adjacent segments more marked with faeces than others in a single 500-m transect. The choice of a segment length of 100 m seems to be appropriate for the Pyrenean desman when using the Markovian occupancy model, as it is long enough to detect signs when the animals are present and short enough to exhibit dependence in occupancy between adjacent segments.

Nevertheless, some caution is called for with regard to the closure assumption. In the temporal replicates design, we do not know the extent to which the Pyrenean desman shifts its home range. We kept the duration of the survey as short as possible (i.e. 4 months) to best meet this assumption (Rota *et al.* 2009). In the spatial replicates design, the sampling site was 500 m in length, which corresponds to the reported mean home range size for the Pyrenean desman (Melero *et al.* 2012). We cannot fully address how likely it is that we met the closure assumption. Preliminary findings have suggested that the Pyrenean desman is highly mobile within its home range (Y. Melero, pers. comm.).

The detectability of a species is rarely constant across space and time and often depends on factors such as environment or species biology (MacKenzie *et al.* 2002). Although the spatial variation of detectability is increasingly acknowledged (e.g. Royle & Nichols 2003; Royle 2006), little is known about the influence of unmodelled variation between sites on parameter estimates (MacKenzie *et al.* 2006). The difficultly in accessing and surveying our sampling sites mean that cost and time considerations have constrained the selection of these sites and limited their number. Adding covariates in occupancy models to a small data set might cause them to fail to converge, especially the Markovian occupancy model, which estimates five parameters even without any covariate (Hines *et al.* 2010). However, since the aim of this study was not to identify the environmental factors that influence the detectability of the Pyrenean desman, the model parameters were kept

constant. With a larger number of sites sampled at the scale of the Pyrenees range, for example, adding environmental covariates to the Markovian occupancy model (Hines *et al*. 2010) would help to identify important factors that influence both the detectability and occupancy of the Pyrenean desman.

IMPLICATIONS FOR THE CONSERVATION OF THE PYRENEAN DESMAN

The cryptic behaviour of the Pyrenean desman and its habitat heterogeneity clearly result in its detectability being variable in space and time, as well as being imperfect. Although it is much easier to detect the Pyrenean desman by looking for its faeces, many factors, such as the skills of the observer (e.g. Barber-Meyer *et al*. 2013), the configuration of the river, the fluctuations in water levels (natural and human-induced), the period when sampling is conducted or the meteorological conditions (e.g. Aing *et al*. 2011), may have a considerable impact on the chance of finding faeces. For example, it may be less easy to detect faeces of the Pyrenean desman in areas with higher annual rainfall or in rivers influenced by hydroelectric activities, as these factors may reduce the persistence of faeces or accelerate the process of deterioration. In this study, all the sampled stream segments had emergent items which must be present if Pyrenean desman signs are to be detected. However, the density of emergent items along the river varied between segments, and we assume that the presence and/or detectability of signs may be influenced by the proportion of the surface area covered with emergent items. A possible covariate describing, for instance, the density of emergent items along streams deserves to be further explored. It is indeed critical to understand how these different factors affect the probability of detecting the species for the design of efficient monitoring strategies in the future (Field, Tyre & Possingham 2005). This study provides further insights into possible survey design and ways of monitoring this endangered and elusive species. The results of a survey conducted at a larger scale (i.e. the whole Pyrenees) will enable to increase our knowledge about the geographical range of this species in order to identify any discontinuities within this range. This in turn should allow biodiversity managers and government organizations to make informed decisions regarding the conservation of the Pyrenean desman and its habitats in Andorra, Spain and France.

IMPLICATIONS FOR SURVEYS OF OTHER SPECIES

Survey areas sometimes present difficult field conditions (e.g. topography, climate) that hinder access and sign-based searches. On a large spatial scale and with a large number of sampling sites, temporal replicates could be very time- and money-consuming and consequently prohibitive. Moreover, for monitoring to be useful, it needs to be sustainable in the long term. Spatial replication seems more effective and straightforward to implement than temporal replication (Hines *et al*. 2010). Once observers have reached a site, the cost of performing additional survey replicates is low (Mackenzie & Royle 2005). This is supported by our results, suggesting that sampling with spatial replicates only slightly underestimates site occupancy compared to temporal replicates. Hence, adjacent spatial replication appears to be a good trade-off between costs of sampling (money and time) and accuracy of occupancy estimates. However, the choice between spatial vs. temporal design for the survey of other species will depend on the attention turned to each of these two aspects. Some general precautions should also be kept in mind when using the Markovian occupancy models to avoid violating the model assumptions and biasing the estimates of detection and occupancy. For instance, species behaviour, mainly in relation to the size of its home range and use (e.g. the distribution of signs within home range), must be considered when deciding the appropriate length of sampling sites and subunits. In this study, we showed that surveys based on adjacent spatial replicates along streams and the Markovian occupancy model parameter settings were adequate for a species with a small and linear home range. This approach may also be useful for a wider range of species which can be more easily detected by sign (e.g. burrows, tree scratches, dung, galleries, nests and exuviae) or use trails to move such as many large mammals (Karanth *et al*. 2011; Thorn *et al*. 2011; Barber-Meyer *et al*. 2013) or flightless birds (e.g. cassowaries). It may also be appropriate to restrict surveys to roads or trails, for the sake of improving sign detection, particularly for species living in remote habitats, such as mountains (e.g. bears) or tropical forests (e.g. orang-utangs). Additionally, it would be interesting to test the Markovian occupancy model on other species using linear home ranges, for which traditional resource selection analyses are often inappropriate (e.g. Slaght *et al*. 2013), such as species living along rivers (e.g. beavers, dippers, dragonflies), restricted to ecotones (e.g. edges, reeds, coastlines, continental slope) or living along hedges, cliffs or canyons.

Acknowledgements

We would like to thank V. Lacaze, E. Vas, A. Bodo, M.O. Durand, C. Fonty, V. Leenknegt, B. Le Roux & S. Perré for the field work. We also thank J.E. Hines for his helpful comments concerning the Markovian occupancy model, as well as M. Ghosh and S. Danflous who corrected the English text. The study is funded by ANRT (Cifre no. 2011/1018), EDF (Electricité de France) and European Union (FEDER) and is part of the French conservation Action Plan for the Pyrenean desman (2010–2015) supervised by DREAL MP (Direction Régionale pour l'Environnement, l'Aménagement et le Logement de Midi-Pyrénées) and coordinated by the CEN MP (Conservatoire d'Espaces Naturels Midi-Pyrénées).

References

Aing, C., Halls, S., Oken, K., Dobrow, R. & Fieberg, J. (2011) A Bayesian hierarchical occupancy model for track surveys conducted in a series of

linear, spatially correlated, sites. *Journal of Applied Ecology*, **48**, 1508–1517.
Anderson, G.B., Freeman, M.C., Hagler, M.M. & Freeman, B.J. (2012) Occupancy modeling and estimation of the holiday darter species complex within the etowah river system. *Transactions of the American Fisheries Society*, **141**, 34–45.
Barber-Meyer, S.M., Jnawali, S.R., Karki, J.B., Kanal, P., Lohani, S., Long, B. *et al*. (2013) Influence of prey depletion and human disturbance on tiger occupancy in Nepal. *Journal of Zoology*, **289**, 10–18.
Bertrand, A. (1994) *Répartition géographique et écologie alimentaire du desman des Pyrénées Galemys pyrenaicus* (Geoffroy, 1811) dans les Pyrénées françaises. Diplôme universitaire de recherche, Toulouse.
Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L. & Thomas, L. (2001) *Introduction to Distance Sampling – Estimating Abundance of Biological Populations*. Oxford University Press, Oxford.
Burnham, K.P. & Anderson, D.R. (2002) *Model Selection and Multi-model Inference: A Practical Information-Theoretic Approach*, p. 353. Springer, New York.
Durso, A.M., Willson, J.D. & Winne, C.T. (2011) Needles in haystacks: estimating detection probability and occupancy of rare and cryptic snakes. *Biological Conservation*, **144**, 1508–1515.
Fernandes, M., Herrero, J., Aulagnier, S. & Amori, G. (2008) *Galemys pyrenaicus*. In: *IUCN 2012. IUCN Red List of Threatened Species. Version 2012.2*. www.iucnredlist.org. Downloaded on 20 March 2013.
Field, S.A., Tyre, A.J. & Possingham, H.P. (2005) Optimizing allocation of monitoring effort under economic and observational constraints. *The Journal of Wildlife Management*, **69**, 473–482.
Gibson, L., Barrett, B. & Burbidge, A. (2007) Dealing with uncertain absences in habitat modelling: a case study of a rare ground-dwelling parrot. *Diversity and Distributions*, **13**, 704–713.
Gu, W. & Swihart, R.K. (2004) Absent or undetected? Effects of non-detection of species occurrence on wildlife–habitat models. *Biological Conservation*, **116**, 195–203.
Guillera-Arroita, G. (2011) Impact of sampling with replacement in occupancy studies with spatial replication. *Methods in Ecology and Evolution*, **2**, 401–406.
Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147–186.
Hines, J.E. (2006) *PRESENCE 2: Software to Estimate Patch Occupancy and Related Parameters*. U.S. Geological Survey, Patuxent Wildlife Research Center, Laurel, MD. http://www.mbr-pwrc.usgs.gov/software/presence.htmli
Hines, J.E., Nichols, J.D., Royle, J.A., MacKenzie, D.I., Gopalaswamy, A.M., Kumar, N.S. & Karanth, K.U. (2010) Tigers on trails: occupancy modeling for cluster sampling. *Ecological Applications*, **20**, 1456–1466.
Karanth, K.U., Gopalaswamy, A.M., Kumar, N.S., Vaidyanathan, S., Nichols, J.D. & MacKenzie, D.I. (2011) Monitoring carnivore populations at the landscape scale: occupancy modelling of tigers from sign surveys. *Journal of Applied Ecology*, **48**, 1048–1056.
Kendall, W.L. & White, G.C. (2009) A cautionary note on substituting spatial subunits for repeated temporal sampling in studies of site occupancy. *Journal of Applied Ecology*, **46**, 1182–1188.
Kéry, M., Gardner, B. & Monnerat, C. (2010) Predicting species distributions from checklist data using site-occupancy models. *Journal of Biogeography*, **37**, 1851–1862.
Kéry, M., Guillera-Arroita, G. & Lahoz-Monfort, J.J. (2013) Analysing and mapping species range dynamics using occupancy models. *Journal of Biogeography*, **40**, 1463–1474.
Kindberg, J., Ericsson, G. & Swenson, J.E. (2009) Monitoring rare or elusive large mammals using effort-corrected voluntary observers. *Biological Conservation*, **142**, 159–165.
Lobo, J.M., Jiménez-Valverde, A. & Hortal, J. (2010) The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, **33**, 103–114.
Mackenzie, D.I. & Royle, J.A. (2005) Designing occupancy studies: general advice and allocating survey effort. *Journal of Applied Ecology*, **42**, 1105–1114.
MacKenzie, D.I., Nichols, J.D., Lachman, G.B., Droege, S., Royle, J.A. & Langtimm, C.A. (2002) Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, **83**, 2248–2255.
MacKenzie, D.I., Nichols, J.D., Royle, J.A., Pollock, K.H., Bailey, L.L. & Hines, J.E. (2006) *Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence*. Academic Press, Burlington, MA.
McCarthy, M.A., Moore, J.L., Morris, W.K., Parris, K.M., Garrard, G.E., Vesk, P.A. *et al*. (2012) The influence of abundance on detectability. *Oikos*, **122**, 717–726.
Melero, Y., Aymerich, P., Luque-Larena, J.J. & Gosàlbez, J. (2012) New insights into social and space use behaviour of the endangered Pyrenean desman (*Galemys pyrenaicus*). *European Journal of Wildlife Research*, **58**, 185–193.
Némoz, M. & Bertrand, A. (2008) *Plan National d'Actions en faveur du desman des Pyrénées (Galemys pyrenaicus)*, 2009-2014. Société Française pour l'Etude et la Protection des Mammifères/Ministère de l'Ecologie, de l'Energie, du Développement Durable et de l'Aménagement du Territoire, Toulouse.
Némoz, M., Bertrand, A., Sourie, M. & Arlot, P. (2011) A French conservation action plan for the Pyrenean desman *Galemys pyrenaicus*. *Galemys*, **23**, 47–50.
Nichols, J.D., Bailey, L.L., O'Connell. Jr, A.F., Talancy, N.W., Campbell Grant, E.H., Gilbert, A.T., Annand, E.M., Husband, T.P. & Hines, J.E. (2008) Multi-scale occupancy estimation and modelling using multiple detection methods. *Journal of Applied Ecology*, **45**, 1321–1329.
Olea, P.P. & Mateo-Tomás, P. (2011) Spatially explicit estimation of occupancy, detection probability and survey effort needed to inform conservation planning. *Diversity and Distributions*, **17**, 714–724.
Parry, G.S., Bodger, O., McDonald, R.A. & Forman, D.W. (2013) A systematic re-sampling approach to assess the probability of detecting otters *Lutra lutra* using spraint surveys on small lowland rivers. *Ecological Informatics*, **14**, 64–70.
Pavlacky, D.C., Blakesley, J.A., White, G.C., Hanni, D.J. & Lukacs, P.M. (2012) Hierarchical multi-scale occupancy estimation for monitoring wildlife populations. *The Journal of Wildlife Management*, **76**, 154–162.
Pollock, K.H., Nichols, J.D., Simons, T.R., Farnsworth, G.L., Bailey, L.L. & Sauer, J.R. (2002) Large Scale Wildlife Monitoring Studies: Statistical Methods for Design and Analysis. *Environmetrics*, **13**, 105–119.
Rondinini, C., Wilson, K.A., Boitani, L., Grantham, H. & Possingham, H.P. (2006) Tradeoffs of different types of species occurrence data for use in systematic conservation planning. *Ecology letters*, **9**, 1136–1145.
Rota, C.T., Fletcher, R.J. Jr, Dorazio, R.M. & Betts, M.G. (2009) Occupancy estimation and the closure assumption. *Journal of Applied Ecology*, **46**, 1173–1181.
Royle, J.A. (2006) Site Occupancy Models with Heterogeneous Detection Probabilities. *Biometrics*, **62**, 97–102.
Royle, J.A. & Nichols, J.D. (2003) Estimating abundance from repeated presence-absence data or point counts. *Ecology*, **84**, 777–790.
Slaght, J.C., Horne, J.S., Surmach, S.G. & Gutiérrez, R.J. (2013) Home range and resource selection by animals constrained by linear habitat features: an example of Blakiston's fish owl. *Journal of Applied Ecology*, **50**, 1350–1357.
Thorn, M., Green, M., Bateman, P.W., Waite, S. & Scott, D.M. (2011) Brown hyaenas on roads: estimating carnivore occupancy and abundance using spatially auto-correlated sign survey replicates. *Biological Conservation*, **144**, 1799–1807.
Yoccoz, N.G., Nichols, J.D. & Boulinier, T. (2001) Monitoring of biological diversity in space and time. *Trends in Ecology & Evolution*, **16**, 446–453.

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Table S1. Model selection results for all temporal replicates and spatial replicates.

Table S2. Minimal occupancy and detectability estimates obtained for all temporal replicates and spatial replicates.

Table S3. Maximal occupancy and detectability estimates obtained for all temporal replicates and spatial replicates.