Proposition d'une probabilité mesurant l'activité des Chiroptères sur un site de chasse

Hubert Lallement
6 rue de l'orée de la forêt
10190 Prugny.

Introduction

Pour étudier l'activité des chauves-souris sur les étangs du Parc Naturel Régional de la Forêt d'Orient (PRNFO), B. FAUVEL et D. BECU ont relevé en un point donné, à l'aide d'un appareil à ultrasons détectant les sonars des chiroptères, les présences-absences de chaque groupe d'espèces pendant 20 périodes d'une minute.

Par exemple lors d'une prospection menée le 28/06/2004, l'activité du groupe Pipistrelles a été mesurée par le rapport :

$$p = \frac{nombre\ de\ présences}{20} = \frac{14}{20} = 0.7.$$

Se pose le problème de savoir ce que mesure cette proportion.

1) On peut d'abord la voir comme une estimation de la proportion P du temps de présence (supposé continu par morceaux) du groupe sur les 20 minutes.

Dans notre exemple le groupe Pipistrelles est donc présent 70% du temps. Mais il n'existe pas d'intervalle de confiance donnant une estimation de l'erreur .

Cependant pour améliorer cette estimation on peut noter les présences-absences toutes les 30 secondes, on divise alors par deux l'erreur d'estimation de P sans pour autant estimer cette erreur ; dans notre exemple on trouve :

$$p = \frac{nombre\ de\ présences}{40} = \frac{17}{40} = 0.425, \text{ soit } 42.5\% \text{ du temps.}$$

Compte tenu du matériel utilisé et des possibilités humaines on ne peut pas réduire les enregistrements à moins de 30 secondes, donc on ne pourra pas améliorer encore la précision de cet estimateur. Enfin remarquons qu'a priori on ne peut pas comparer des points d'écoute qui n'ont pas la même durée !

2) On peut aussi voir p comme estimation de la probabilité P de contacter au moins une fois le groupe pendant une minute, pourvu que celle-ci soit constante au cours du relevé. p = 0.7 signifie alors qu'on a 70% de chance de contacter au moins une fois le groupe Pipistrelles pendant une minute d'écoute.

Si de plus on fait l'hypothèse que les présences-absences sont indépendantes au cours du relevé, on dispose alors d'un intervalle de confiance à 95% pour P : 0.46-0.88 ; c'est-à-dire que la probabilité P de contacter le groupe Pipistrelles au moins une fois pendant une minute est comprise entre 46% et 88%. Si on veut améliorer la précision de l'estimation de P il suffit d'augmenter le nombre de périodes d'écoute, par exemple avec 60 minutes d'écoute on a grosso modo $P = p \pm 13\%$.

Remarquons que les relevés par périodes de 30 secondes deviennent inutiles, ce qui est une bonne chose car ils sont bien moins confortables que les relevés par périodes d'une minute, donc moins sujets à erreurs !

Observons qu'avec cette probabilité P, on peut comparer des points d'écoute qui n'ont pas la même durée !

«Remarquons que les relevés par périodes de 30 secondes deviennent inutiles, ce qui est une bonne chose car ils sont bien moins confortables que les relevés par périodes d'une minute, donc moins sujets à erreurs ! »

Donnons maintenant une définition plus précise de P.

Définition d'une probabilité

Etant donné un observateur situé en un point x d'un site, muni d'un appareil détectant les ultrasons, on appelle probabilité d'occurrence d'une espèce (ou d'un groupe d'espèces), la probabilité P_x que celle-ci soit détectée avec cet appareil au moins une fois pendant une minute.

En fait cette probabilité est liée à la portée de l'appareil, c'est-à-dire à une certaine région de l'espace qui peut varier en fonction du milieu puisque certains obstacles comme les arbres peuvent faire écran acoustique. P_x dépend bien sûr du lieu x, et du milieu dans lequel il se trouve, mais également des caractéristiques de l'appareil, de la date et de l'heure, des conditions météorologiques, du ou des types de milieux situés à proximité, et de l'observateur s'il n'est pas expérimenté.

Cependant on fait l'hypothèse que pour un jour donné, P_x ne dépend pas du temps pendant les 2 ou 3 premières heures suivant le coucher du soleil.

Nous avons vu dans l'introduction l'intérêt de P_x sur les étangs puisqu'il n'y avait qu'un point d'écoute.

P_x étant un paramètre ponctuel, dans un site homogène du point de vue de l'effet d'écran acoustique, il est intéressant de se demander si P_x est constant.

Méthode

Pour répondre à cette question on réalise un plan d'échantillonnage en tirant au hasard N points x_1, x_2,…., x_N , puis en chaque point x_i on relève les présences-absences durant T_i unités d'écoute d'une minute, et on fait l'hypothèse fondamentale que *toutes les unités d'écoute sont indépendantes.*

On appelle P_i la probabilité d'occurrence du groupe au point x_i . Puis nous confrontons les modèles suivants :

M_1 : il existe au moins deux probabilités P_i et P_j distinctes ;

chaque probabilité P_i est alors estimé par : $p_i = \dfrac{nombre\ de\ présences\ en\ x_i}{T_i}$.

M_0 : les probabilités P_i sont toutes égales à un même nombre P ;

P est alors estimé par : $p = \dfrac{nombre\ total\ de\ présences}{nombre\ total\ d'unités\ d'écoute}$.

Plusieurs choix sont possibles pour comparer M_0 et M_1, pour des raisons de facilité d'interprétation nous ne retenons que le facteur de BAYES dont on expliquera le principe plus bas.

Donc si M_0 l'emporte, P sera appelée : *probabilité d'occurrence du groupe sur le site*, dans le cas contraire il faudra faire une étude complémentaire pour voir s'il y a des regroupements possibles de points adjacents et alors scinder le site en plusieurs parties.

Application

Dans le cadre de l'étude d'impact associée à un projet d'implantation d'éoliennes sur les communes de Fontvannes et Dierrey-Saint-Julien, D. BECU a sondé plusieurs sites en une soirée.

On retient pour appliquer notre méthode deux transects effectués le long de deux chemins d'exploitation agricole situés dans la pleine céréalière, car ce sont les sites qui offrent le plus d'unités d'écoute (25 et 30) ; et n'existant pas d'écran acoustique, les probabilités d'occurrence peuvent être comparées. Enfin nous n'avons retenu que le groupe Pipistrelles car il est le mieux représenté dans les deux cas.

Les résultats sont résumés dans les tableaux suivants :

Chemin A			
Numéros de points d'écoute	Nombres d'unités d'écoute	Nombres de présences	Estimations des Pi
i =1	5	4	0.8 (= 4/5)
2	5	2	0.4
3	5	0	0
4	5	2	0.4
5	5	0	0
Total	25	8	0.32 (=8/25)

Chemin B			
Numéros de points d'écoute	Nombres d'unités d'écoute	Nombres de présences	Estimations des Pi et de P
i =1	5	0	0 (= 0/5)
2	5	0	0
3	5	1	0.2
4	5	0	0
5	5	1	0.2
6	5	1	0.2
Total	30	3	0.1 (=3/30)

Analyse avec le facteur de BAYES

Le facteur de BAYES est défini par : $K(x) = [\, p(M_1 \mid x) \prod p(M_1) \,] \div [\, p(M_0 \mid x) \prod p(M_0) \,]$,
où :

$p(M_1 \mid x)$ est la probabilité du modèle M_1 connaissant les observations x,
$p(M_0 \mid x)$ est la probabilité du modèle M_0 connaissant x ,
$p(M_1)$ est la probabilité du modèle M_1 et $p(M_0)$ celle du modèle M_0 .

Le rapport $p(M_1 \mid x) \prod p(M_1)$ mesure la variation relative de la probabilité de M_1 due à la connaissance des observations x ;
de même le rapport $p(M_0 \mid x) \prod p(M_0)$ mesure la variation relative de la probabilité de M_0 due à la connaissance des observations x.

Donc si K(x) est nettement plus grand que 1, M_1 sera le modèle le plus plausible, si K(x) est nettement plus petit que 1 ce sera M_0.
Il est remarquable que K(x) ne dépende pas des valeurs de $p(M_1)$ et de $p(M_0)$.
Enfin, pour ne privilégier aucune valeur des paramètres P_i et P on leur a attribué des lois a priori uniformes sur l'intervalle [0 ; 1].

Les résultats sont donnés par le tableau suivant :

Chemins	Facteur de BAYES K(x)
A	7.23
B	0.02

Pour évaluer le degré de certitude en faveur ou au détriment de M_0 apporté par les données, JEFFREYS a développé l'échelle suivante :

Facteur de BAYES K(x)	Inférieur à 1/100	Compris entre 1/100 et 1/10	Compris entre 1/10 et 1/3	Compris entre 1/3 et 1	Compris entre 1 et 3	Compris entre 3 et 10	Compris entre 10 et 100	Supérieur à 100
Inter-prétation	Certitude décisive en faveur de M_0	Certitude forte en faveur de M_0	Certitude substantielle en faveur de M_0	Certitude faible en faveur de M_0	Certitude faible en faveur de M_1	Certitude substantielle en faveur de M_1	Certitude forte en faveur de M_1	Certitude décisive en faveur de M_1

Pour le chemin A : K(x) = 7.23 ;
donc le rapport $p(M_1 \mid x) \prod p(M_1)$ est 7 fois plus grand que le rapport $p(M_0 \mid x) \prod p(M_0)$;
d'après l'échelle de JEFFREYS la certitude que M_1 soit vrai est substantielle ;
donc il faut admettre l'hétérogénéité des probabilités de détection sur le chemin A.

Pour le chemin B : $K(x) = 0.02 = \frac{1}{50}$;
donc le rapport $p(M_1 \mid x) \prod p(M_1)$ est 50 fois plus petit que le rapport $p(M_0 \mid x) \prod p(M_0)$;
d'après l'échelle de JEFFREYS la certitude que M_0 soit vrai est forte ;
donc il faut admettre l'homogénéité des probabilités de détection sur le chemin B.

On peut donc parler de la probabilité P d'occurrence des Pipistrelles sur ce chemin, elle est estimée par p = 0.1 d'après le tableau « chemin A ».
Un intervalle de confiance à 95% pour P est : 2%-26% , il est très large, mais c'est normal vu que le nombre d'unités d'écoute n'est que de 30 !

Conclusion

Cette probabilité d'occurrence a de nombreux avantages : simplicité d'utilisation et d'estimation ; elle permet de faire des comparaisons, de caractériser un lieu de chasse si elle est homogène sur celui-ci. Mais d'autres études seront nécessaires pour infirmer ou confirmer l'hypothèse qu'elle est constante dans le temps et (ou) dans l'espace ; cela nécessitera un effort d'échantillonnage important.

Il semble vraisemblable que son existence même soit liée à la variabilité plus ou moins importante de co-variables biologiques, environnementales et météorologiques.

Bibliographie

-Les calculs ont été faits à l'aide du logiciel libre R :
R Development Core Team (2006). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

-ROBERT, C.P. (2006), Le choix Bayésien, principes et pratique. Springer-Verlag.