

## Un outil statistique pour l'exploration de la sélection de l'habitat avec des données de radio-pistage

Clément Calenge\*, Daniel Maillard & Anne-Béatrice Dufour\*

\* CNRS-UMR 5558, Laboratoire de biométrie et biologie évolutive, Université de Lyon 1, 69622 Villeurbanne Cedex

### Contexte de l'étude

La définition d'une politique de gestion efficace pour une population d'animaux sauvages implique la compréhension de la nature des interactions entre les individus qui la composent et leur environnement. Afin de mettre en évidence ces interactions, de nombreuses études ont été entreprises par l'Office national de la chasse et de la faune sauvage, aussi bien au niveau populationnel (Darmon, 2003) qu'au niveau individuel (*e.g.* Maillard, 1996). Les études populationnelles ont pour objectif de caractériser la distribution spatiale des individus sur une zone et de mettre en évidence les variables environnementales (altitude, pente, etc.) qui l'influencent. Les études individuelles se concentrent quant à elles sur les variables environnementales qui déterminent la distribution du temps passé par les animaux dans les différentes zones qui leur sont disponibles. Les études individuelles se focalisent essentiellement sur l'aspect comportemental de la sélection de l'habitat. Ce type d'étude nécessite l'utilisation de techniques de suivi temporel des localisations des animaux, telles que le radio-pistage.

La plupart des animaux sont sédentaires et concentrent leur activité à l'intérieur d'une surface réduite appelée « domaine vital ». Les études individuelles ont alors pour objectif de déterminer les variables d'habitat qui expliquent la distribution de l'activité d'un animal au sein de ce domaine vital. Le principal problème de ce type d'étude est que tous les animaux suivis ne sont pas soumis aux mêmes contraintes. La composition de l'habitat disponible dans le domaine vital – c'est à dire le contexte environnemental dans lequel les animaux vivent – varie d'un animal à l'autre. Ainsi, face à un type d'habitat, les animaux ne feront pas nécessairement les mêmes choix (sélection/rejet).

Cette variabilité potentielle de la sélection de l'habitat doit être prise en compte lors des analyses. Nous présentons ici une méthode d'analyse factorielle que nous avons développée pour explorer la sélection de l'habitat au niveau individuel, l'analyse K-select (Calenge *et al.*, 2005). Cette analyse repose sur le concept de niche écologique, et permet d'établir une typologie des animaux en fonction de leur mode d'utilisation de l'habitat. Nous en décrivons ici le principe, puis nous en donnons une illustration avec le sanglier en milieu méditerranéen.

### Concept de niche écologique pour un animal

La formalisation du concept de niche écologique par Hutchinson (1957) a permis le développement de nombreuses méthodes d'analyse des relations entre une espèce et son environnement. Nous décrivons brièvement ici comment ce concept peut être utilisé comme moyen d'analyser la sélection de l'habitat par un animal à l'intérieur de son domaine vital (figure 1A).

Chacune des variables environnementales étudiées définit une dimension dans un espace multidimensionnel (*e.g.* l'altitude, la pente et la distance aux points d'eau définissent un espace à trois dimensions ; le principe mathématique du concept peut être étendu à un nombre quelconque

de variables). Cet espace multidimensionnel est appelé espace écologique. On suppose par ailleurs que la surface disponible à l'animal – son domaine vital – est composée d'un certain nombre « d'unités de ressources » discrètes, par exemple les pixels d'une carte raster. Chacun des pixels du domaine vital possède une valeur pour chacune des variables environnementales, on peut donc associer à chaque pixel un point « disponible » dans l'espace écologique.

Par ailleurs, on dispose d'une mesure de l'utilisation des pixels par l'animal. Il peut s'agir par exemple de la proportion de localisations de l'animal dénombrées dans chaque pixel du domaine vital. Cette mesure permet d'associer un poids d'utilisation à chacun des points disponibles de l'espace écologique. L'ensemble des points de

l'espace écologique pour lesquels le poids d'utilisation est non-nul définit la niche écologique de l'animal.

Deux paramètres peuvent être utilisés pour caractériser la niche écologique : la marginalité et la tolérance. Le vecteur de marginalité est le vecteur reliant le centre de gravité du nuage de points disponibles au centre de gravité de la niche. Le carré de la longueur de ce vecteur, appelée marginalité, est une mesure de la sélection de l'habitat. La tolérance est égale à la somme des carrés des écarts entre les points disponibles et le centre de gravité de la niche, ces écarts étant pondérés par les poids d'utilisation.

La marginalité mesure donc la position de la niche dans l'espace écologique, et la tolérance mesure le volume occupé par la niche dans cet espace. Une margina-

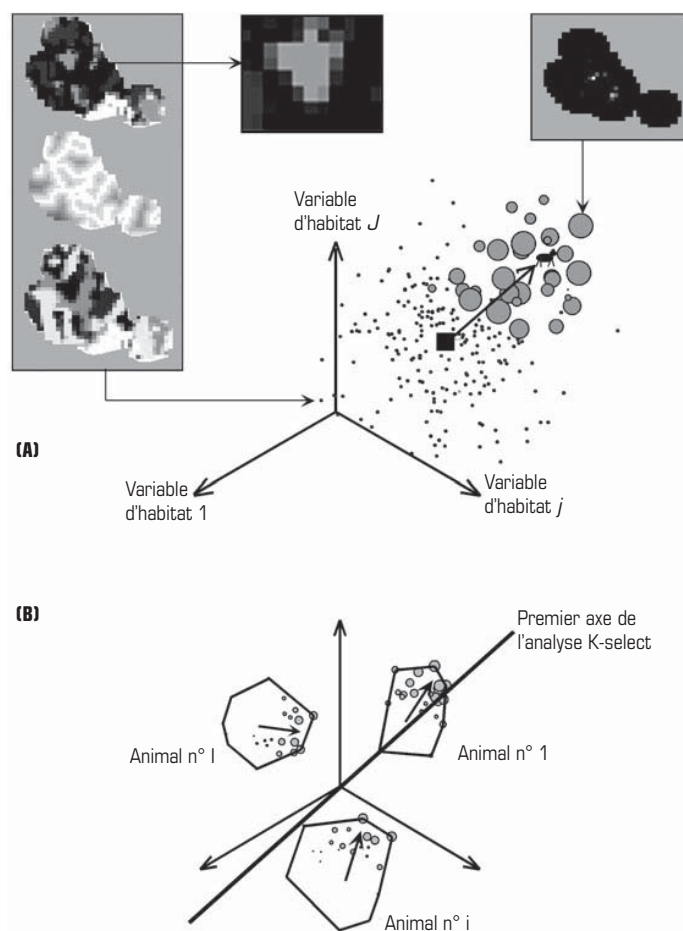
lité importante correspond à un caractère très excentré de la niche écologique par rapport à l'ensemble du nuage de points disponibles, donc à une sélection importante de la part de l'animal. Une tolérance faible correspond à un faible volume occupé par la niche dans l'espace écologique disponible donc à une exigence écologique plus forte.

### Cas de plusieurs animaux : l'analyse K-select

Dans le cas où l'on dispose de plusieurs animaux, les choses se compliquent. En effet, la forme et l'emplacement du nuage de points disponibles dans l'espace écologique varient d'un animal à l'autre (figure 1B). Comme nous l'avons noté précédemment, cette variabilité du contexte dans lequel les animaux vivent se traduit peut être par une variabilité des préférences de l'animal vis-à-vis de son environnement. En d'autres termes, les vecteurs de marginalité des animaux ne sont pas nécessairement tous orientés dans la même direction.

L'analyse K-select permet de prendre en compte cette variabilité inter-individuelle dans l'analyse de la sélection de l'habitat fondée sur des données de radio-pistage (Calenge *et al.*, 2005). Il s'agit d'une analyse factorielle très proche de l'analyse en composantes principales. Elle est conçue pour établir une typologie des animaux en fonction de l'importance et de l'orientation de leurs vecteurs de marginalité.

Quel est le principe de cette analyse ? Chacun des K animaux suivis possède un domaine vital qui lui est propre, représenté par un nuage de points dans l'espace écologique ayant une forme et une position particulières. Comme nous l'avons dit précédemment, la distribution des localisations de chaque animal dans son domaine vital permet d'associer des poids d'utilisation à chacun de ces points, et donc de définir une niche et un vecteur de marginalité particuliers au sein de ce nuage. L'analyse K-select recherche alors un axe dans l'espace écologique tel que la marginalité moyenne expliquée par cet axe est maximisée. Puis elle recherche un deuxième axe, orthogonal au premier, qui maximise à son tour la marginalité expliquée, et ainsi de suite. Le



**Figure 1 : (A)** Le concept de niche écologique utilisé pour analyser la sélection de l'habitat par un animal à l'intérieur de son domaine vital. Le domaine vital est constitué de pixels (*cf.* grossissement) possédant des valeurs pour chacune des variables d'habitat étudiées. Chaque pixel correspond à un point disponible dans l'espace écologique défini par les variables d'habitat (ici, les trois variables définissent un espace à trois dimensions). Le nombre de localisations dans chaque pixel (à droite) permet d'associer un poids d'utilisation à chacun de ces points disponibles (cercle gris de rayon proportionnel au pourcentage de localisations dans les pixels correspondants). L'ensemble des points utilisés définit la niche de l'animal. Le vecteur de marginalité (représenté par une flèche) relie le centre de gravité des points disponibles au barycentre de la niche.

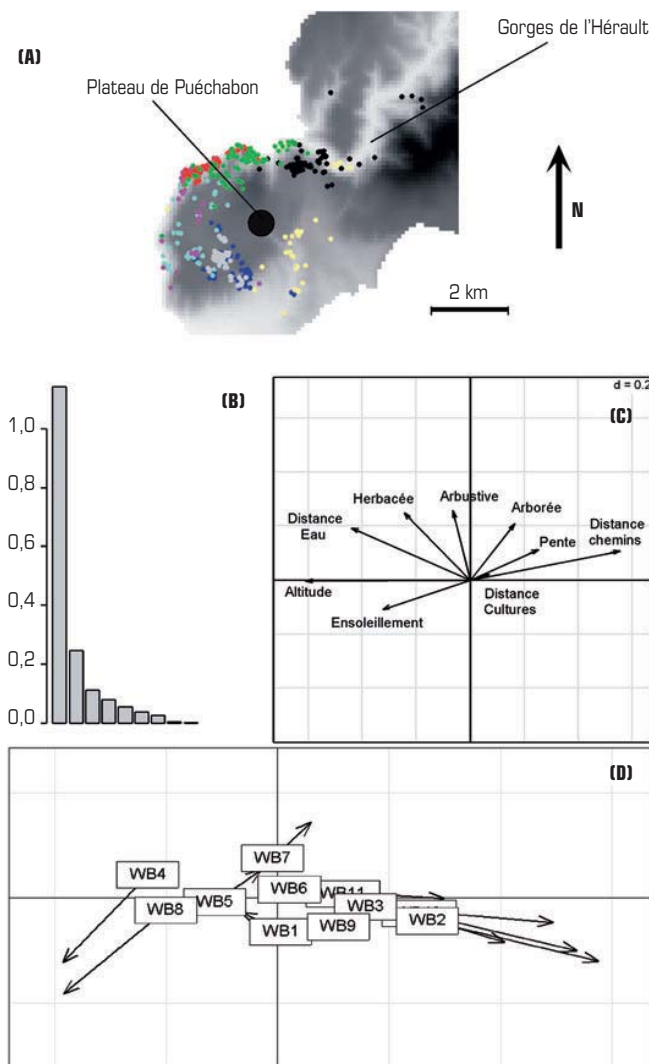
**(B)** Le principe de l'analyse K-select : pour chacun des animaux suivis, on peut définir un nuage de points disponibles (ici représenté par un polygone convexe, dans un souci de lisibilité), et une niche associée à ce nuage de points. Il est donc possible de définir un vecteur de marginalité pour chacun des animaux suivis. L'analyse K-select va rechercher un système d'axes orthogonaux qui maximisent la marginalité expliquée (*i.e.* un axe le plus « parallèle » possible à tous les vecteurs de marginalité).

premier plan factoriel de l'analyse est ainsi la meilleure « photographie » de la sélection de l'habitat dans l'espace écologique. Les scores des variables permettent ensuite d'attribuer une signification biologique à ces axes.

### Un exemple, la sélection des bauges par le sanglier

Dans le cadre d'une étude sur le sanglier en milieu méditerranéen, 11 sangliers

ont été suivis par radio-pistage en 1992 et 1993 (Maillard, 1996) sur le plateau de Puéchabon (zone d'étude de 4 500 ha, à 30 km au nord-ouest de Montpellier). Les gorges de l'Hérault structurent très fortement toutes les variables d'habitat sur la zone d'étude (les gorges sont à la fois pentues, ombragées, proche de l'eau, loin des chemins, etc.). Les sangliers ont été localisés quotidiennement à la bauge (site de repos diurne) pendant les mois de juillet et août (figure 2A). Nous nous concentrons ici sur la sélection des sites de bauge à l'intérieur du domaine vital. Les variables d'habitat étudiées sont



**Figure 2** : Résultats de l'analyse K-select appliquée pour analyser la sélection de l'habitat par les sangliers en milieu méditerranéen lors de l'établissement des bauges (sites de repos diurnes). (A) Distribution des localisations des sangliers suivis (une couleur par animal) sur une carte de l'altitude sur la zone d'étude. (B) Diagramme des valeurs propres. (C) Scores des variables sur les deux premiers axes de l'analyse. (D) Projection des vecteurs de marginalité des animaux sur les deux premiers axes de l'analyse.

l'altitude, la pente, l'ensoleillement, la distance aux points d'eau, aux chemins, et aux cultures ainsi que la densité de la végétation herbacée, arbustive et arborée. Nous considérons comme disponible à un animal l'ensemble des pixels situés à moins de 500 mètres d'une localisation. Nous avons appliqué la K-select à ces données afin de déterminer les composantes de l'habitat recherchées par le sanglier dans son domaine vital pour établir ses bauges au cours de cette période estivale.

Les résultats sont présentés sur la figure 2. Les valeurs propres de l'analyse mesurent la marginalité expliquée par chaque axe (figure 2B). Deux axes sont suffisants pour expliquer l'essentiel de

la marginalité des animaux suivis, avec un premier axe prédominant (58 % de la marginalité) qui oppose les zones proches des gorges de l'Hérault du côté des valeurs positives (proches de l'eau, fortes pentes, altitude faible, ensoleillement relativement faible, éloigné des chemins) à celles qui en sont éloignées, du côté des valeurs négatives (caractéristiques inverses, cf. figure 2C). Le deuxième axe oppose quant à lui les zones les plus denses en végétation herbacées, arbustives et arborées (valeurs négatives) aux zones plus ouvertes (valeurs positives).

Mais c'est la représentation des vecteurs de marginalité sur le premier plan factoriel qui apporte le plus d'informations (figure 2D). Rappelons que l'origine

de ces vecteurs correspond à la position du centre de gravité du nuage de points disponibles, et que leur extrémité correspond au centre de gravité de la niche. Trois types d'animaux peuvent être identifiés sur ce plan factoriel. Le premier groupe d'animaux est situé sur la droite du graphique ; il correspond aux animaux pour lesquelles les gorges de l'Hérault étaient très disponibles (e.g. WB2, WB3). Ces animaux sélectionnent les gorges de façon très importante. Le deuxième groupe, au centre du graphique, est constitué de sangliers avec une marginalité faible (donc pour lesquels la sélection de l'habitat était relativement faible ; e.g. WB1, WB7). Le dernier groupe est constitué des deux sangliers les plus à gauche du graphique (WB4 et WB8), c'est à dire pour lesquels les gorges de l'Hérault étaient peu disponibles. Ces animaux sélectionnaient un couvert végétal dense. Ainsi, nous mettons en évidence une sélection de l'habitat variable entre les animaux. Les sangliers pour lesquels les gorges sont disponibles sélectionnent ce type d'habitat pour établir leurs bauges. Ceux pour lesquels elles ne sont pas disponibles sélectionnent les zones avec un couvert végétal dense. La sécurité semble être le facteur déterminant pour l'établissement des bauges.

## Discussion

L'analyse K-select permet une exploration précise de la sélection de l'habitat par les animaux suivis par radio-pistage. Cette méthode nous a permis d'identifier une variabilité de la sélection de l'habitat parmi les sangliers en milieu méditerranéen : les zones les plus denses en fourrés sont en général préférées par les animaux, mais lorsque des zones plus sûres sont accessibles (e.g. les gorges de l'Hérault), alors le choix des animaux s'oriente vers ces zones. Ainsi, la sécurité est bien le facteur déterminant pour le choix des bauges, mais ce qui est considéré comme sûr par un sanglier dépend du contexte dans lequel il vit.

La prise en compte d'une éventuelle variabilité de la sélection de l'habitat entre les animaux est une force de l'analyse. En effet, la majorité des méthodes de la sélection de l'habitat suppose une sélection identique entre

les animaux. Bien sûr, certaines méthodes permettent de tester une différence de la sélection de l'habitat entre groupes d'animaux (*e.g.* entre mâles et femelles), mais ces méthodes supposent toutes que les groupes sont connus a priori. L'analyse K-select, au contraire, ne fait aucune hypothèse préalable sur les types d'animaux qui seront susceptibles d'être mis en évidence, et permet l'établissement de cette typologie (pour une discussion approfondie de ce sujet, cf. Calenge & Dufour, 2006).

Ce caractère extrêmement flexible de la méthode est une propriété des analyses factorielles. Le plus grand intérêt de cette famille de méthodes est qu'aucune hypothèse préalable n'est faite sur les données. Aucune hypothèse n'est faite sur la forme de la distribution des variables environnementales. On peut y inclure autant de variables environnementales qu'on le souhaite, et il n'y a aucune limite inférieure au nombre d'animaux pouvant être analysés. Notons toutefois qu'un nombre plus important d'animaux suivis permettra de donner plus de force aux conclusions biologiques tirées à la suite de l'analyse. En outre, ces méthodes permettent de prendre en compte le

caractère spécifiquement multivarié de l'habitat. Ces méthodes sont en conséquence l'outil idéal pour l'exploration des données écologiques.

Concrètement, l'application de l'analyse K-select est rendue possible grâce à la bibliothèque de fonctions *adehabitat* pour le logiciel R (Calenge, 2006). R est un logiciel coopératif, c'est à dire construit par ses utilisateurs et disponible gratuitement sur Internet (URL : <http://www.r-project.org>), qui offre une très large gamme d'outils statistiques pour l'analyse des données, notamment des données écologiques (Tufto & Cavallini, 2005). Les utilisateurs de R peuvent télécharger et installer un certain nombre de bibliothèques de fonctions directement par les menus du logiciel. Nous avons développé *adehabitat* sous R afin de permettre l'analyse de la sélection de l'habitat à l'aide d'un très grand nombre de méthodes. De nombreuses fonctions permettent l'interaction entre R et les systèmes d'information géographique, des fonctions spatiales de base (calcul de zones tampon, jointure spatiale, changement de résolution), des fonctions permettant l'estimation du domaine vital (polygone convexe minimum, méthode du noyau, etc.), des méthodes d'analyse

de la sélection de l'habitat (analyse K-select, ENFA, analyse de rapports de sélection, analyse compositionnelle, etc.). Une présentation de cette bibliothèque est donnée par Calenge (2006).

Les analyses factorielles ont prouvé leurs grandes potentialités en matière d'analyse des données écologiques, tout particulièrement en matière d'analyse de la sélection de l'habitat. Le récent développement des nouvelles technologies permettant un suivi plus fin des animaux (colliers GPS) et l'analyse des données spatiales (Système d'information géographique, logiciel R) vont permettre de répondre à beaucoup d'interrogations en termes de gestion des habitats de la faune sauvage, autrefois inabordable sur le plan technique. Le développement méthodologique dans ce domaine est en conséquence un véritable enjeu pour l'avenir.

## Remerciements

Nous remercions Daniel Chessel de son aide précieuse dans le développement de l'analyse K-select.

## BIBLIOGRAPHIE

- Calenge C., Dufour A. B. & D. Maillard (2005) – K-select analysis : a new method to analyse habitat selection in radio-tracking studies. *Ecological Modelling*, 186, 143-153
- Calenge C. (2006) – The package *adehabitat* for the R software : a tool for the analysis of space and habitat use by animals. *Ecological Modelling*, *sous presse*.
- Calenge C. & A. B. Dufour (2006) – Eigenanalysis of selection ratios from animal radio-tracking data. *Ecology*, *sous presse*.
- Darmon G. (2003) – Déterminants intra et interspécifiques de la sélection de l'habitat par les chamois et les mouflons du massif des Bauges. Rapport de D.E.A., Université Claude Bernard, Lyon 1.
- Hutchinson G. E. (1957) – Concluding remarks. In : Cold Spring Harbour Symposium pp. 415-427. *Quantitative Biology*.
- Maillard D. (1996) – Occupation et utilisation de la garrigue et du vignoble méditerranéens par le sanglier (*Sus scrofa* L.). In : *Biologie des populations et des écosystèmes* pp. 235 pp. Université de droit, d'économie et des sciences d'Aix-Marseille III.
- Tufto J. & P. Cavallini (2005) – Should wildlife biologists use free software ? *Wildlife Biology*, 11 : 67-76

## ABSTRACT

### A statistical tool to explore habitat selection from radio-tracking data

Clément Calenge, Daniel Maillard & Anne-Béatrice Dufour

■ We present here a new factorial method, the K-select analysis, allowing the study of habitat selection by animals using radio-tracking data and raster maps of several environmental variables. Each mapped variable defines a dimension in a multidimensional space, the ecological space, and each pixel in the map corresponds to a point in this space. For a given animal, the pixels defining its home-range delimit the space of resources available to the animal, and the pixels containing at least one relocation define the niche of the animal within this space. The K-select analysis computes the axes that maximises the mean marginality for all animals. The marginality is the mean squared distance between the centre of gravity of available resources and the centre of gravity of the niche, that is a measure of habitat selection. This analysis therefore provides the best possible image of habitat selection by the animals. We provide an application of this method to the study of habitat selection by the wild boar (*Sus scrofa*) at *Puéchabon* (South of France). This method can be applied using a free package for the R software named *adehabitat*.