

**AMELIORATION DU MODELE VIA FAUNA,
UN MODELE D'IDENTIFICATION DES CORRIDORS
ECOLOGIQUES TERRESTRES EN MIDI-PYRENEES,
A PARTIR DE L'ETUDE DES SUIVIS GPS DES ESPECES
CHEVREUIL ET SANGLIER**



Source : ©INRAE.

Auteur : Sophie MILLET

Maitres de stage : Johan ROY, Lucie DEBEFFE

Tuteur académique : Claire MARSDEN

Lieu de stage : Fédération Départementale des Chasseurs de Haute-Garonne,
INRAE (CEFS)

Période de stage : 16 novembre 2020 au 16 mai 2021

SOMMAIRE

I- Introduction	4
A- Le projet Via Fauna pour améliorer les connaissances sur l'impact des infrastructures linéaires de transport sur la faune sauvage	4
B- Le modèle Via Fauna, un modèle d'identification des corridors écologiques des grands mammifères	4
C- Une étude empirique pour améliorer le modèle Via Fauna	5
II- Le modèle Via Fauna et les données factuelles utilisées pour l'améliorer	5
A- Principe général du modèle Via Fauna	5
1. Principe et étapes de la méthode	5
2. Les paramètres clés du modèle Via Fauna	6
B- Les données utilisées pendant le stage	7
1. Données FDC31 de suivi GPS des sangliers.....	7
2. Données CEFS-INRAE de suivi GPS des chevreuils	8
III- Amélioration de la définition des remises, les points de départ et d'arrivée des chemins de moindre coûts	10
A. Identification des données GPS localisées dans des remises	11
1. Mise en évidence d'une période de remise chez le sanglier grâce à des matrices de distance	11
2. Méthode se basant sur la distance entre deux localisations GPS successives	12
3. Méthode de segmentation comportementale des données GPS sur R.....	15
4. Comparaison des méthodes utilisées pour identifier remises	20
B. Identification des caractéristiques des remises	21
1) Sangliers.....	21
2) Chevreuils	22
C. Discussion et perspectives d'amélioration	27
IV- Amélioration des notes de friction attribuées à chaque occupation du sol	28
A) Principe de l'analyse de sensibilité	29
1. Les différents cas de friction étudiés.....	29
2. Lancement du modèle Via Fauna pour chaque cas de friction.....	29
3. Confrontation des trajectoires des CMC avec les trajectoires GPS réels	29
B) Résultats et interprétations	33
1. Comparaison sangliers/chevreuils.....	33
2. Sangliers.....	37
3. Chevreuils	37
C) Construction de tampons autour des éléments paysagers	38
1. Principe	38
2. Résultats et interprétations	39
D. Conclusion sur l'amélioration de la définition des notes de friction attribuées à chaque occupation du sol	40
E. Discussion et perspectives d'amélioration	41
V- Bibliographie	44
VI- Annexes	45

TABLE DES ANNEXES

Annexe 1	<i>Script R – Identification des données GPS situées dans des remises : méthode distance entre 2 localisations successives</i>	45
Annexe 2	<i>Script R – Identification des données GPS situées dans des remises : méthode segmentation comportementale segclust2d</i>	51
Annexe 3	<i>Valeurs du quantile à 2% pour les distances des remises-chevreuils (remises dans boisements et remises sur OS agricole) à la route, au bâti et au boisement, en fonction de la méthode d'identification des remises utilisée</i>	62
Annexe 4	<i>Nouvelle nomenclature des classes d'occupation du sol élaborée lors du stage</i>	62
Annexe 5	<i>Script R - Analyse sensibilité des frictions attribuées à chaque occupation du sol dans le modèle Via Fauna</i>	63
Annexe 6	<i>Guide méthodologique pour réaliser une carte d'occupation du sol intégrant les tampons répulsifs autour des routes et attractifs autour des boisements</i>	75

I- Introduction

A- Le projet Via Fauna pour améliorer les connaissances sur l'impact des infrastructures linéaires de transport sur la faune sauvage

D'après un rapport de la DREAL Occitanie l'Occitanie est une des régions métropolitaines les plus consommatrices de surfaces naturelles, agricoles et forestières, avec 33 000 000 de m² artificialisés par an en moyenne (DREAL, 2020). L'artificialisation progressive du territoire a des incidences fortes sur la biodiversité, par la fragmentation des écosystèmes et la destruction des habitats, mais aussi sur les paysages, la ressource en eau, etc. La préservation d'une trame verte et bleue, réseau écologique cohérent alliant réservoirs de biodiversité et espaces de liaison entre ces réservoirs, est un axe fort de la politique nationale en faveur de la biodiversité, que chaque région doit décliner à son échelle dans un schéma régional de cohérence écologique (SRCE). C'est dans ce cadre que s'inscrit le projet Via Fauna, porté par la Fédération des Chasseurs d'Occitanie.

Conduit et animé par les Fédérations des Chasseurs de la région Occitanie depuis 2017, Via Fauna est un projet régional rassemblant près de 40 partenaires techniques et financiers (gestionnaires d'infrastructures linéaires de transport (ILT), gestionnaires faune sauvage, organismes de recherche). Soutenu par l'Europe, l'Office français de la biodiversité (OFB), la Région Occitanie et la Fédération nationale des chasseurs, ce projet a pour objectif d'améliorer la connaissance sur les éventuelles ruptures de continuités écologiques engendrées par les infrastructures linéaires de transport existantes sur les 8 départements de l'ancienne région Midi-Pyrénées. A ce jour, le projet propose différentes méthodologies et outils permettant d'améliorer les connaissances sur les déplacements de la faune sur un territoire. Parmi eux, un outil cartographique pour identifier les secteurs potentiellement perméables et les éléments de ruptures des continuités écologiques vis-à-vis des ILT : le modèle Via Fauna.

B- Le modèle Via Fauna, un modèle d'identification des corridors écologiques des grands mammifères

Le modèle Via Fauna, basé sur une méthode dite de « friction-dispersion », modélise les occupations du sol les plus favorables aux déplacements de la grande faune. Il s'agit en fait de relier les uns aux autres les éléments à enjeux des grands mammifères (par exemple les grands boisements) grâce au trajet le plus court possible et le moins coûteux au regard des occupations du sol et de l'écologie des espèces.

Ainsi, le modèle constitue d'abord un outil de diagnostic rapide sur un territoire donné puisqu'il permet d'identifier les principaux corridors écologiques des grands mammifères, de localiser les principales entraves ou les secteurs perméables au regard des ILT. Outil d'aide à la décision d'une part (espaces à ne pas urbaniser, ouvrages majeurs pour la fonctionnalité écologique du territoire à préserver, ouvrages à adapter pour perméabiliser), il constitue d'autre part un outil d'évaluation des politiques publiques : en intégrant dans la modélisation la planification et les aménagements prévus puis en identifiant les incidences par rapport à l'état initial. Les cartographies issues du modèle Via Fauna peuvent par exemple s'avérer intéressantes dans le cadre de projets d'aménagement du territoire.

Cependant, cette modélisation reste théorique et des travaux de vérification du modèle doivent être envisagés, notamment en confrontant le résultat du modèle avec des données factuelles issues de suivis GPS d'Ongulés.

C- Une étude empirique pour améliorer le modèle Via Fauna

Depuis plus de 15 ans, le laboratoire CEFS de l'INRAE de Toulouse équipe des chevreuils à l'aide de colliers incluant un dispositif GPS. La Fédération des Chasseurs quant à elle pose des colliers GPS sur des sangliers depuis 2019. L'objectif général de ce stage INRAE/Fédération des chasseurs est d'utiliser ces informations de suivi afin de proposer un modèle Via Fauna amélioré à partir de données factuelles.

Objectifs du stage :

- **1/** Confronter les trajectoires prédites par la méthodologie existante basée sur un modèle théorique de connectivité (modèle Via Fauna utilisant les chemins de moindre coûts) avec des données empiriques de trajectoires réelles provenant de suivis GPS d'individus.
- **2/** Conduire une analyse de sensibilité du modèle existant afin d'identifier les paramètres clefs (caractéristiques de l'espèce fictive, définition des patches d'habitat favorables, seuils de friction des habitats composant la mosaïque paysagère...) et de tester la robustesse de ce dernier
- **3/** Affiner les paramètres du modèle théorique en utilisant les informations empiriques de connectivité qui pourront être déduites des suivis GPS des individus. Le produit de ce 3^{ème} objectif consistera alors en un modèle ViaFauna amélioré pour les espèces chevreuil et sanglier.

II- Le modèle Via Fauna et les données factuelles utilisées pour l'améliorer

A- Principe général du modèle Via Fauna

Nous l'avons vu, le modèle Via Fauna se base sur une méthode dite de « friction-dispersion » pour tracer le chemin le plus court et le moins coûteux en termes d'occupation du sol entre deux secteurs à enjeux utilisés par de grands mammifères, en s'appuyant sur leur écologie.

1. Principe et étapes de la méthode

Les étapes et le principe général du modèle Via Fauna seront ici décrits de manière très succincte afin de comprendre le travail réalisé durant le stage. Pour plus d'informations sur la méthodologie du modèle Via Fauna, on pourra se référer au rapport de stage d'Anaïs Sentenac (2018).

- **Choix de l'espèce fictive des Ongulés.** Le modèle Via Fauna fait le choix d'une démarche généraliste en définissant une "espèce fictive" rendant compte des exigences écologiques moyennes du groupe taxonomique des Ongulés (individu strictement terrestre, déplacements favorisés par linéaires boisés et peu contraints par espaces agricoles, etc.).

- **Réalisation de la carte d'occupation du sol.** Pour cela, l'Occupation du Sol à Grande Échelle (OCSGE) de l'IGN a été utilisée comme couche de base, à laquelle d'autres couches ont été ajoutées successivement afin d'améliorer et préciser les données de base (ajout des haies, des cours d'eau, des routes, etc.). L'ordre de fusion des couches a dû être minutieusement choisi afin de rendre compte autant que possible de la réalité du terrain.
- **Définition de « l'habitat » de l'espèce fictive.** Cette étape est essentielle puisque les « habitats » ont été utilisés comme points de départ et d'arrivées des trajets de moindre coût. Par défaut, il est fixé que tous les boisements supérieurs ou égaux à deux hectares constituent des habitats dans le modèle Via Fauna.
- **Définition des seuils de friction.** Le seuil de friction est une note, attribuée aux dires d'experts, représentant la difficulté d'une espèce à traverser une occupation du sol donnée. Une note de friction a été attribuée à chaque occupation du sol (OS) ainsi qu'à chaque ouvrage d'art (OA) dans le modèle Via Fauna. La carte d'occupation du sol a ensuite été rasterisée et chaque pixel s'est ainsi vu attribué, avec une résolution spatiale de 5 mètres, la valeur de friction de l'occupation du sol qu'il représente.
- **Calcul des chemins de moindre coût (CMC).** Grâce à l'interface Python de QGIS (GRASS), il est possible de calculer et tracer les CMC entre deux points. Un CMC correspond au chemin le moins coûteux, en termes de friction, entre un point de départ et un point d'arrivée et à travers une série d'occupations du sol donnée.
- **Calcul des couloirs de mouvement.** Les couloirs de mouvement correspondent à la somme des cartes de coûts cumulés de friction entre deux chemins inverses (A vers B puis B vers A). Là où les CMC permettent de visualiser qu'un seul passage possible, cette approche reflète une réalité de terrain plus importante puisqu'elle met en évidence un ensemble de corridors de déplacements potentiellement utilisables par la faune.

2. Les paramètres clés du modèle Via Fauna

Ainsi, le modèle Via Fauna possède trois paramètres clés, dont deux ont un lien direct avec les caractéristiques attribuées à l'espèce fictive aux dires des experts (Figure 1).

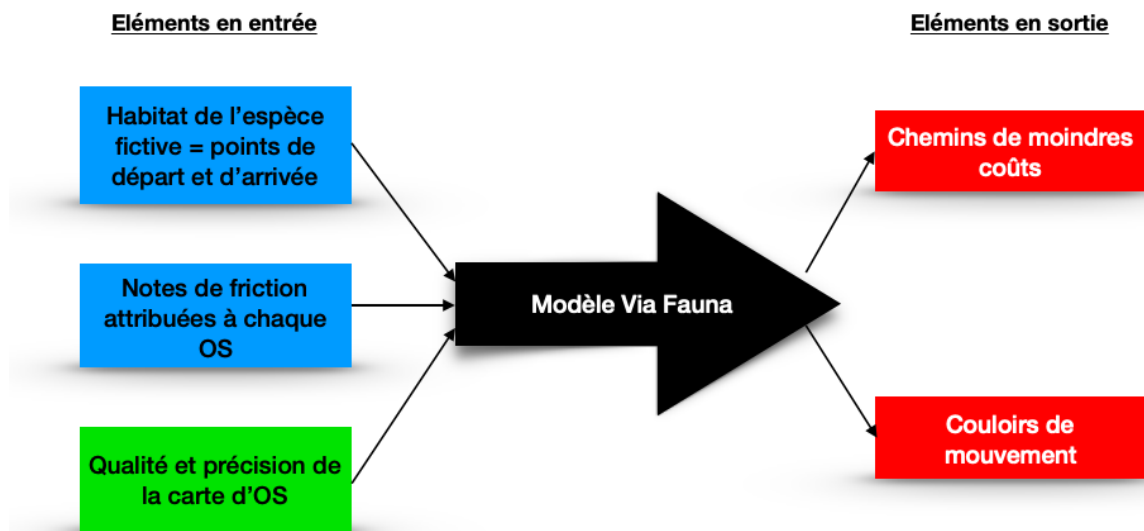


Figure 1 : Le modèle Via Fauna, trois paramètres en entrée et deux éléments en sortie.
En bleu : les éléments en entrée en lien direct avec les caractéristiques écologiques attribuées à l'espèce fictive.

Il est d'abord possible d'améliorer la précision de la carte d'occupation du sol du modèle : en intégrant une couche issue d'une base de données sur la végétation plus précise (IGN par exemple) que celle utilisée jusqu'alors.

Pour améliorer le modèle Via Fauna, il conviendra d'affiner successivement les deux paramètres clés liés aux caractéristiques écologiques de l'espèce fictive, en utilisant les informations déduites des suivis GPS des individus.

B- Les données utilisées pendant le stage

Pour répondre à l'objectif du stage, les données GPS de la Fédération Départementale des Chasseurs de Haute-Garonne (FDC31) et du CEFS ont été utilisées.

1. Données FDC31 de suivi GPS des sangliers

Depuis 2019, la FDC31 équipe des sangliers de colliers GPS de la marque Vectronic. Pendant le stage, les données de suivi GPS de 6 sangliers suivis toutes les heures pendant environ 1 an étaient à disposition (plus de 21 7000 données).

Les sangliers suivis se trouvent dans l'agglomération toulousaine, en zone péri-urbaine (Figure 2). Globalement, la zone d'étude des sangliers est une zone fortement anthropisée et urbanisée, avec de nombreuses infrastructures linéaires de transport relativement fréquentées et une densité de bâti importante.

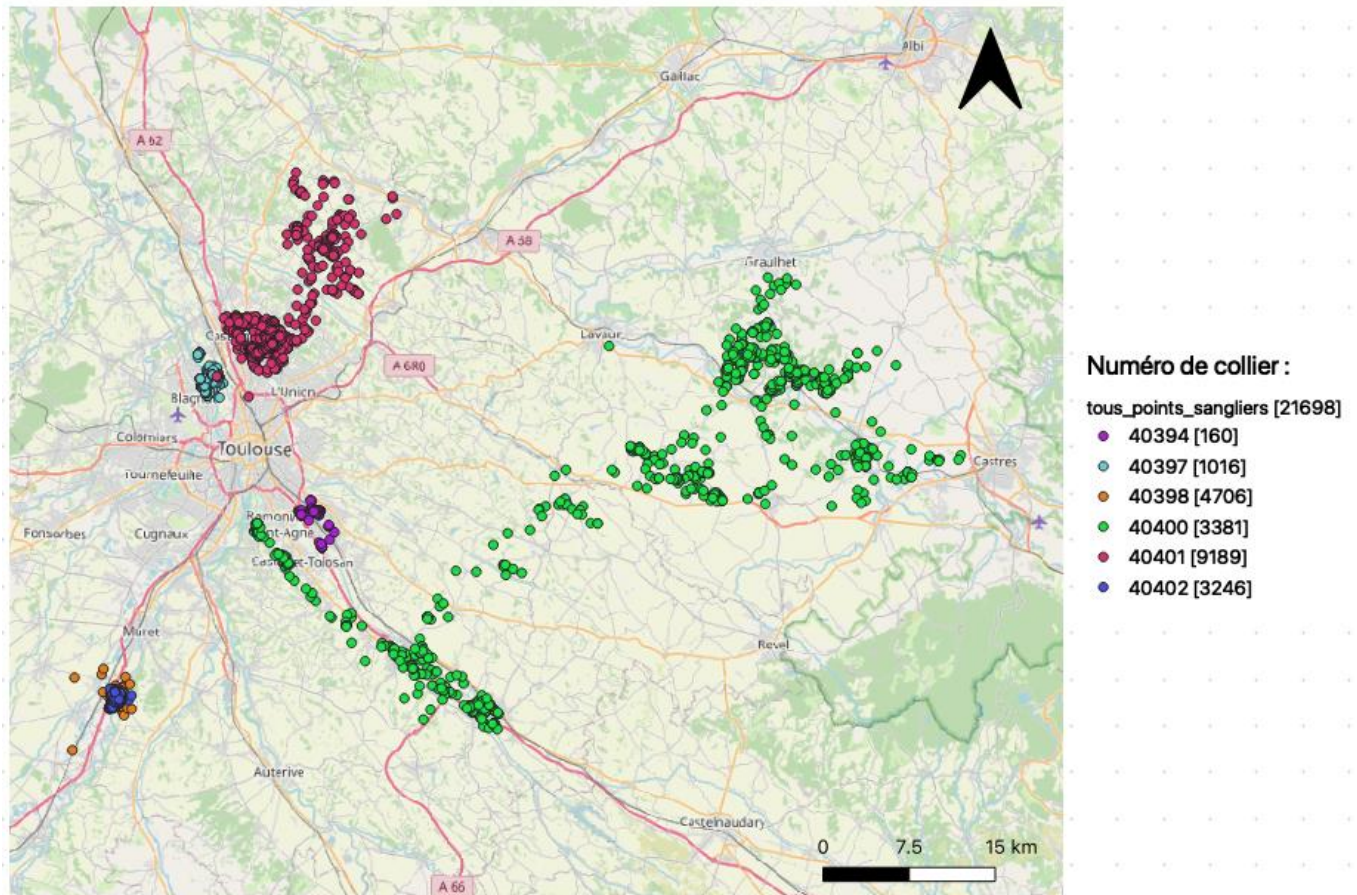


Figure 2 : Zone couverte par les relevés GPS des 6 sangliers suivis par la FDC31, chaque couleur représentant un sanglier.

2. Données CEFS-INRAE de suivi GPS des chevreuils

Depuis 2003, le CEFS capture, équipe de colliers GPS, puis relâche des chevreuils dans le canton d'Aurignac, à une soixantaine de kilomètres au sud de Toulouse. Le laboratoire dispose à ce jour de données GPS pour environ 500 individus. Au cours de la vie de chaque animal suivi, les relevés GPS sont effectués à des fréquences variables : toutes les 6 heures, toutes les heures lors des périodes clés de la vie de l'animal (dispersion natale, période de reproduction, mise-bas) et toutes les 10 minutes pendant 24 heures environ 1 à 2 fois par mois.

Les données GPS recueillies sont globalement de haute résolution : l'erreur médiane est de 7m en milieu ouvert, 23m en forêt de feuillus et 29m en forêt de conifères (Cargnelutti et al. 2007).

Les colliers avec lesquels les chevreuils sont équipés permettent également de mesurer leurs activités : grâce à un système de billes, les mouvements du cou de l'animal sont mesurés, dans le sens de la marche, ou perpendiculairement à celui-ci.

Durant ce stage, on ne se focalisera que sur les données GPS des individus disperseurs. La dispersion, définie comme l'ensemble des mouvements qui éloigne les organismes de leur lieu de naissance de façon permanente (Stenseth & Lidicker 1992), ne concerne presque uniquement que les jeunes de moins d'un an chez le chevreuil (Pettorelli et al. 2003). Ainsi, à

partir du moins d'avril, les faons deviennent plus mobiles et une partie d'entre eux disperse avant l'été (Linnell et al. 1998). Ce phénomène, permettant des flux de gènes conséquents, est un trait d'histoire de vie fondamental au maintien des populations.

Durant la dispersion natale les individus parcourent des distances pouvant être considérables et sont alors exposés à des problèmes de connectivité écologiques. Chez la majorité des individus chevreuils, sédentaires et fidèles spatialement, ces problèmes de connectivité au sein du domaine vital de quelques hectares n'apparaissent pas ou du moins sont bien moindres. Ainsi, il était plus pertinent d'étudier les individus disperseurs pour améliorer le modèle Via Fauna, modèle qui se focalise sur les continuités écologiques. De plus la résolution des cartes sur de petites surfaces pour étudier les trajectoires GPS des individus non-disperseurs ainsi que l'incertitude liée à la précision spatiale des colliers GPS pouvaient constituer un point bloquant de notre étude.

Une classification fournie et élaborée par Laura Benoit, sur la base des travaux de Delphine Ducros (Ducros et al. 2020.) a été utilisée pour identifier les individus disperseurs. A l'issue de ce travail, les données de 61 disperseurs (de 2005 à 2019) ont pu être isolées.

Les chevreuils suivis se trouvent dans un paysage hétérogène, où la densité de bâti est nettement inférieure à celle de la zone d'étude des sangliers. De la même façon, les infrastructures de transport de la zone sont beaucoup moins fréquentées et la superficie boisée ainsi que la matrice agricole représentent globalement une part importante du territoire en comparaison à la zone péri-urbaine où se trouvent les sangliers.

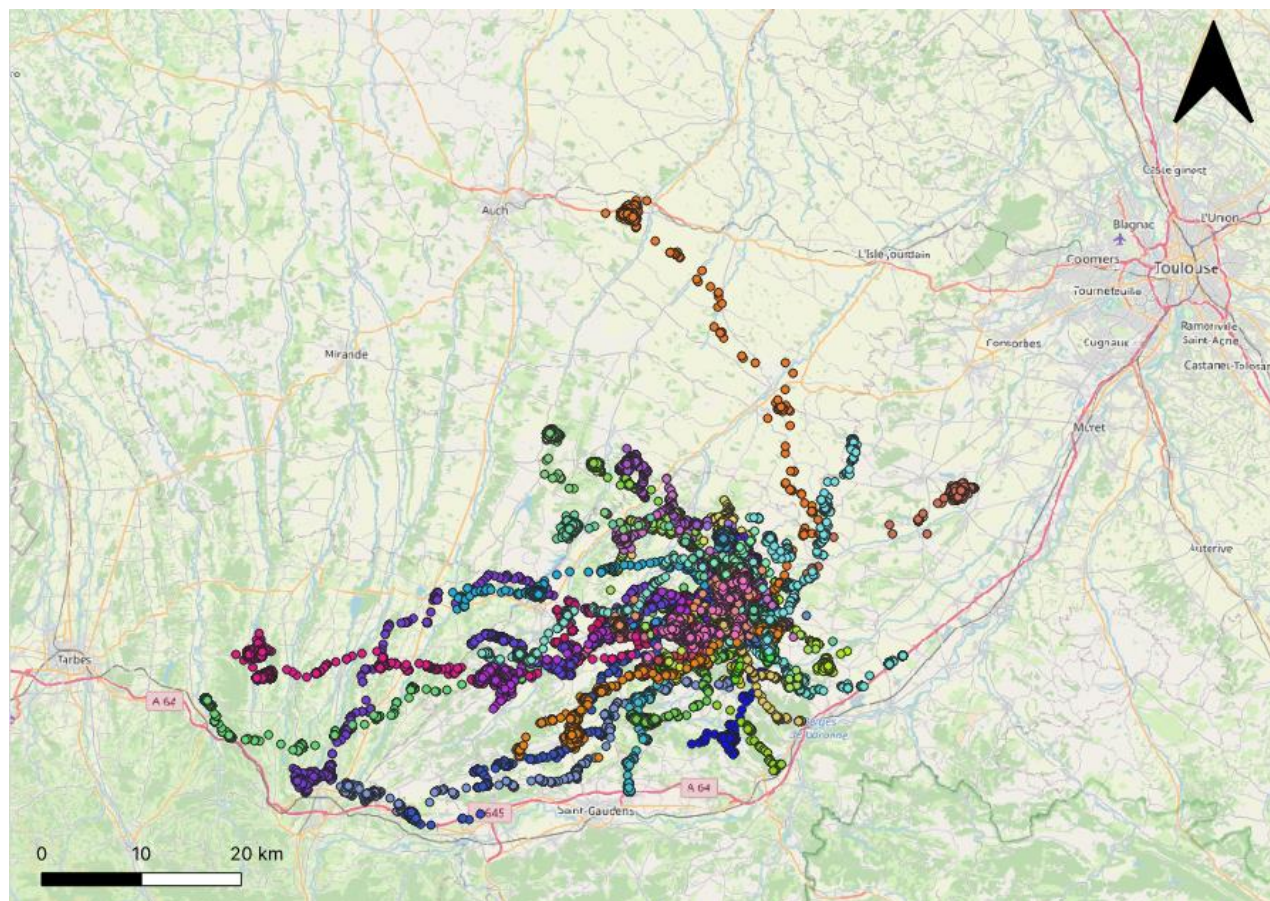


Figure 3 : Zone couverte par les relevés GPS des 61 dispersers suivis par le CEFS-INRAE, chaque couleur représentant un disperser.

III- Amélioration de la définition des remises, les points de départ et d'arrivée des chemins de moindre coûts

Nous l'avons vu, l'habitat est un paramètre clé du modèle Via Fauna puisque lors de son lancement, tous les habitats de la zone d'étude (par défaut les boisements d'une superficie supérieure à 2 hectares) constitueront les points de départ et d'arrivée des chemins de moindre coûts. Par soucis de compréhension, nous remplacerons le terme habitat par remise, en définissant une remise comme tout lieu où l'individu n'effectue que de petits mouvements et où son activité est moindre.

L'objectif premier a donc été d'améliorer la définition des remises dans le modèle Via Fauna. L'amélioration de la définition des remises s'est déroulée en deux étapes. Il a d'abord fallu, pour chaque espèce et à partir des données GPS disponibles, isoler celles se trouvant dans des remises. Une fois les données de remises isolées, il a été possible d'étudier leurs caractéristiques pour chaque espèce, ce qui a permis d'affiner la définition des remises dans le modèle Via Fauna.

A. Identification des données GPS localisées dans des remises

1. Mise en évidence d'une période de remise chez le sanglier grâce à des matrices de distance

Dans un premier temps, l'objectif était de déterminer si les espèces étudiées possédaient une certaine périodicité dans leurs activités quotidiennes. Plus spécifiquement, le but était de préciser s'il existait une période de la journée où les animaux étudiés étaient remisés, c'est-à-dire une période où ils n'effectuaient pas de grands mouvements et restaient principalement cantonnés à un lieu donné (appelé « habitat » dans le modèle Via Fauna). On nommera ces périodes « périodes de remises ».

Pour cela, en s'appuyant sur le principe du Net Square Displacement à la première localisation, très utilisé en écologie du mouvement (Bunnefeld et al. 2011., Borger et al. 2012.), des matrices de distance ont été générées sur QGIS pour chaque animal suivi : pour chaque relevé GPS on calcule sa distance par rapport à un point de référence (ici : la place du Capitole). Ainsi, en traçant un graphique de la matrice de distance en fonction du temps, il est possible de repérer des « plateaux », correspondant à des périodes où l'animal s'est très peu déplacé d'un point à l'autre.

Pour le sanglier, on remarque très vite la rythmicité des journées de l'animal : ces dernières se caractérisent par une période de remise la journée puis d'activité durant la nuit (Figure 4). A l'œil, il semblerait que la période de remise du sanglier se trouve entre 5h et 18h. Nous identifierons cette période de remise de manière plus précise dans la partie suivante.

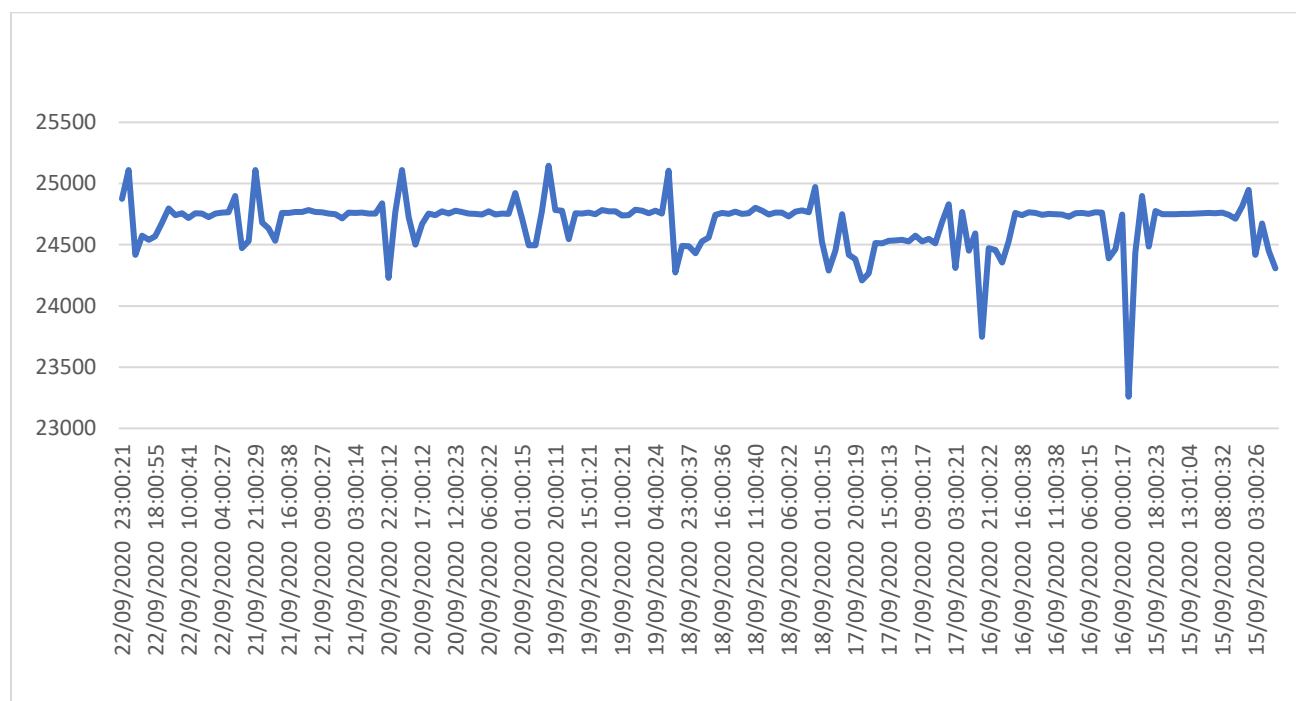


Figure 4 : Matrice de distance du sanglier équipé du collar40398 : distance de chaque localisation GPS par rapport à un point de référence (place du capitole) en fonction du temps (une semaine de suivi ici).

En revanche, pour le chevreuil, aucune tendance concernant la périodicité des journées de l'animal ne semble se dégager des graphiques générés. En effet, la forme des graphiques pour un même individu varie beaucoup d'une semaine à l'autre : si une tendance peut éventuellement être mise en évidence lors d'une semaine, la semaine suivante le même chevreuil peut passer plusieurs jours consécutifs dans la remise, sans effectuer de grands mouvements pour en sortir (Figure 5).

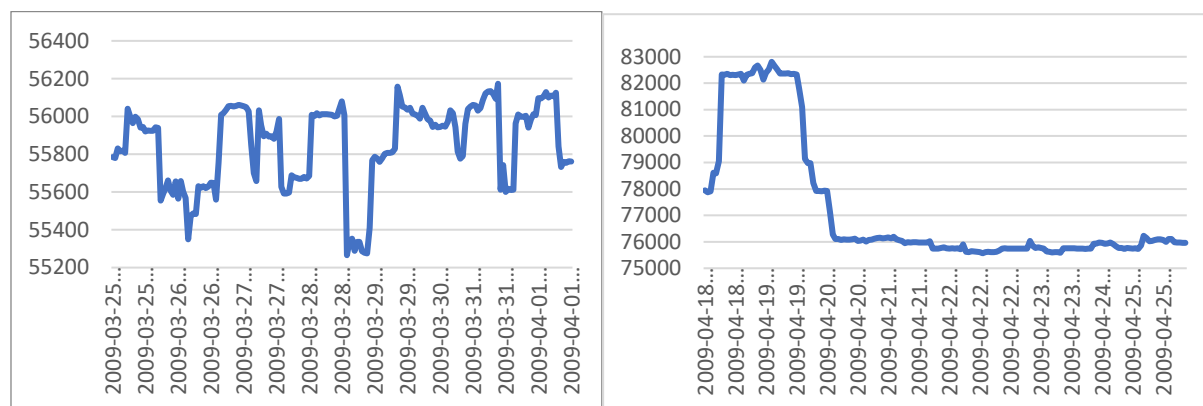


Figure 5 : Matrice de distance du chevreuil 510_09 : distance de chaque localisation GPS par rapport à un point de référence (place du capitole) en fonction du temps (une semaine de suivi ici). A gauche : semaine de suivi n°1. A droite : semaine de suivi n°4.

L'espèce sanglier possède donc clairement une période destinée à la remise durant sa journée, alors que ce n'est pas forcément le cas pour le chevreuil. Pour identifier les points GPS se trouvant dans des remises pour les sangliers et les chevreuils, deux méthodes ont été utilisées : l'une se basant sur la distance entre deux localisations GPS successives et l'autre utilisant un package R de segmentation comportementale des données.

2. Méthode se basant sur la distance entre deux localisations GPS successives

a. Principe

En partant de la matrice de distance générée précédemment, la différence de distance (toujours par rapport au point de référence) entre deux données GPS successives dans le temps a été calculée pour chaque animal suivi. Un compteur a ensuite compté, pour chaque heure de suivi, le nombre de fois où l'évènement A se produisait.

$A =$ « la distance entre 2 données GPS successives est supérieure ou égale à x mètres. »

Les données GPS pour lesquelles le compteur valait 0 ont ensuite été isolées, c'est-à-dire les données pour lesquelles la distance entre 2 points successifs est inférieure à x mètres. Elles correspondent aux remises ainsi définies dans cette méthode.

Grâce à cette méthode, on parvient donc d'une part à isoler les points GPS localisés dans des remises, et d'autre part à identifier la période de remise (si elle existe) de l'espèce étudiée en générant le diagramme en bâton du compteur de l'évènement A en fonction du temps.

Le script détaillé de cette méthode d'identification des remises figure en Annexe 1.

b. Résultats et interprétations

i. Sangliers

Pour isoler les points de remises des sangliers de toutes les données GPS disponibles, le script R a été exécuté pour différentes valeurs de x . Puis pour chaque cas de x , le pourcentage de points-remises se trouvant dans un boisement a été calculé (Figure 6). Comme il a été admis que les sangliers ne se remisaient exclusivement que dans des boisements (Gérard, J.F. & al., 1991) la valeur de x pour laquelle le pourcentage sera le plus grand correspondra à la distance optimale pour identifier les remises des sangliers.

On trouve ainsi que la valeur $x=200$ permet d'identifier au mieux les remises des sangliers. Avec cette méthode basée sur la distance entre deux localisations GPS successives, les remises des sangliers correspondent donc à toutes les données pour lesquelles les sangliers parcourent moins de 200 mètres en une heure.

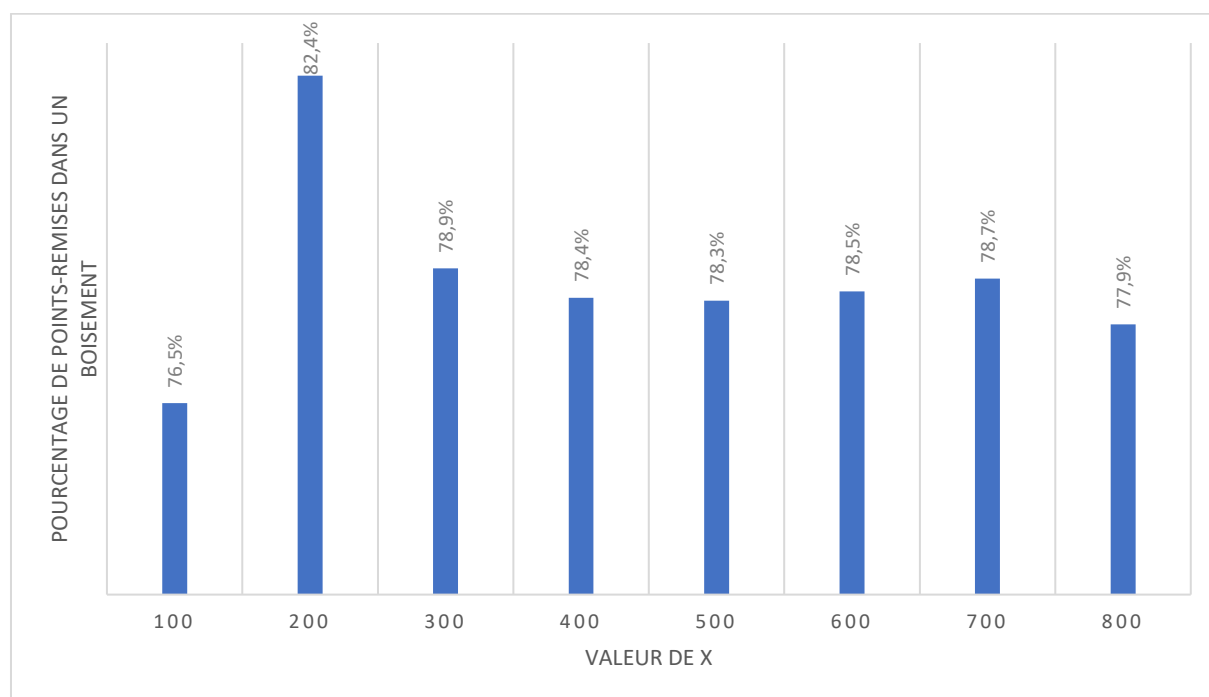


Figure 6 : Pourcentage de points-remises dans un boisement en fonction de la valeur de x

Pour chaque valeur de x testée, un diagramme en bâton donnant l'occurrence de l'évènement A a également été tracé en fonction de l'heure de la journée (Figure 7 pour $x=200$ mètres). Sur ces graphiques, quelle que soit la valeur de x testée, on observe toujours une période où les sangliers ont une activité moindre, soit une période de remise. Cette période se produit toujours au même moment de la journée et sur un intervalle de 13 heures : de 5h à 17h.

Une autre méthode pour isoler les remises des données GPS a donc été trouvée : on sélectionne parmi toutes les données seulement celles incluses dans l'intervalle de la période de remise des sangliers (5h-17h).

L'intervalle de 13 heures ainsi trouvé est en accord avec celui sur lequel figure les plateaux des graphiques des matrices de distance (Figure 4) et la bibliographie confirme également notre conclusion quant à la période de remise des sangliers. En effet, Mauget (1980) a montré que l'entrée en activité des sangliers était synchrone avec le coucher du soleil. J.F Gérard (1991) quant à lui montre que la fin de « la » phase d'activité peut intervenir à des

moments très variables de la nuit, ou plus rarement en début de matinée. Enfin, Janeau & Spitz (1984) ont prouvé que la période de remise des sangliers durait entre 13h et 16h.

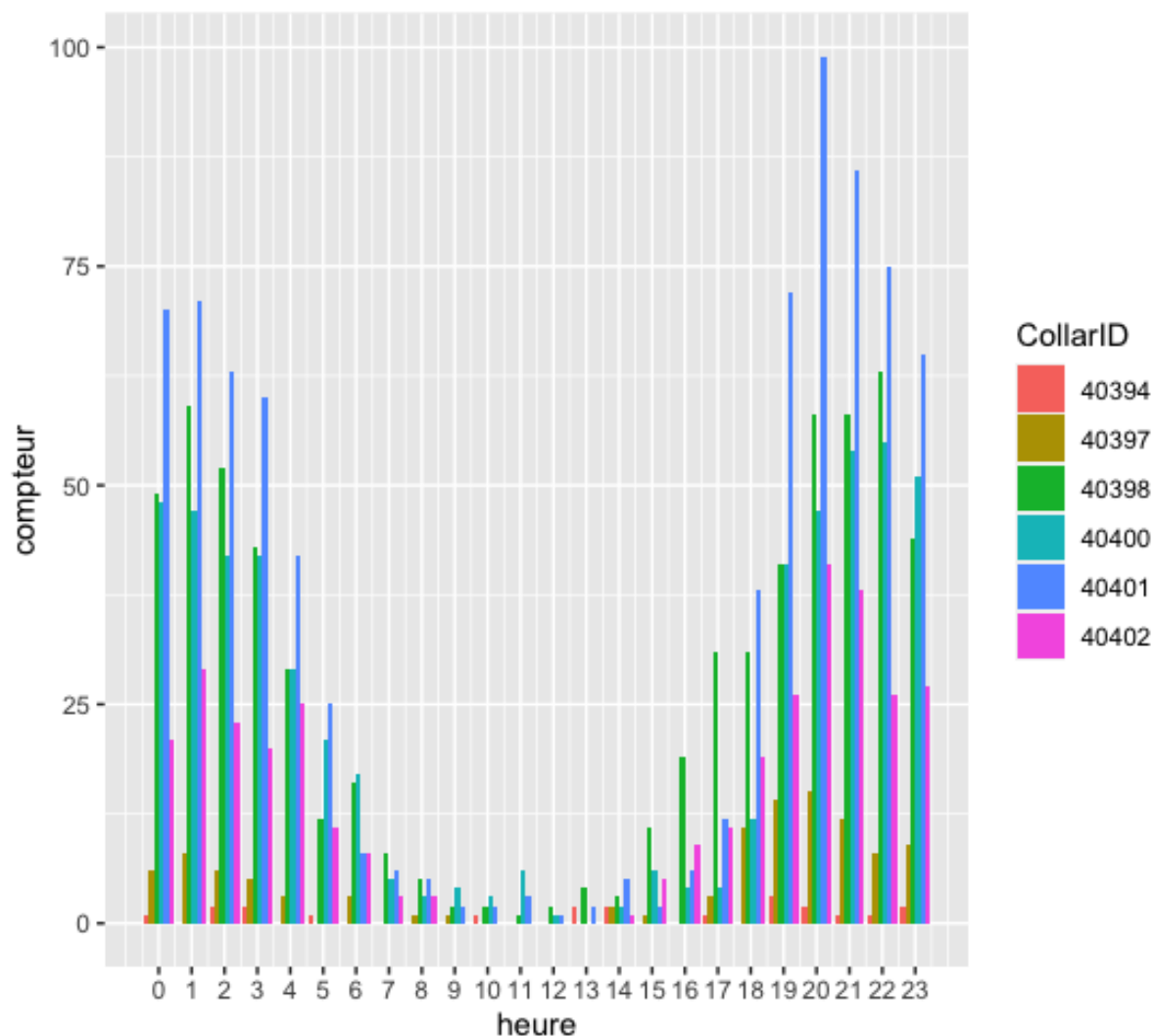


Figure 7 : Occurrence de l'évènement "la distance entre deux données GPS successives est supérieure ou égale à 200 mètres" en fonction de l'heure de la journée.

ii. Chevreuils

Pour l'espèce chevreuil, différentes valeurs de x ont également été testées pour isoler les points de remises. Pour chaque valeur de x, le pourcentage de points correspondant à une remise par rapport au nombre de points total a été calculé (Tableau 1).

Tableau 1 : Pourcentage de points-remises par rapport aux points GPS totaux, en fonction de la valeur de x testée

x (en mètres)	Pourcentage de points-remises par rapport aux points totaux
---------------	---

10	8%
20	23%
50	55%

N'ayant aucune information sur les remises des chevreuils, il a été impossible de choisir la valeur de x permettant de sélectionner les remises de manière optimale, comme cela a été fait précédemment pour le sanglier. La valeur de x pour cette méthode a donc été choisie de manière un peu plus arbitraire et en cela il mériterait de changer de méthode pour sélectionner les remises des chevreuils.

En procédant par élimination, $x=20$ a été retenu pour sélectionner les remises des chevreuils en se basant sur la méthode des distances entre 2 localisations successives. En effet, il serait trop risqué de choisir la valeur $x=10$ pour cette méthode puisque la résolution spatiale des GPS est de 7 mètres environ, d'autant plus que cette dernière augmente si le couvert végétal est dense (en forêt par exemple). Les remises extraites avec cette méthode en prenant la valeur $x=50$ correspondent à plus de la moitié des données GPS totales. Il convient donc d'abandonner la méthode pour cette valeur de x si l'on souhaite étudier les caractéristiques des remises de manière fine.

Ainsi, avec cette méthode basée sur la distance entre deux localisations GPS successives, les remises des chevreuils correspondent donc à toutes les données pour lesquelles les chevreuils parcourent moins de 20 mètres en 10 minutes.

Les diagrammes en bâtons de l'occurrence de l'évènement A en fonction de l'heure de la journée ont été générés pour l'espèce chevreuil mais aucune période de remise n'a pu être identifiée, comme le prédisaient les graphiques des matrices de distance en fonction du temps. Ainsi, il n'est pas possible de faire une sélection sur l'heure pour isoler les remises des chevreuils, comme il a été fait pour les sangliers.

3. Méthode de segmentation comportementale des données GPS sur R

a. Principe

La méthode utilisée se base sur un package R nommé `segclust2d` et élaboré par R. Patin. (Patin R et al., 2019.). Cette méthode est capable de segmenter une série chronologique bivariée et de regrouper des segments similaires (correspondant à des phases stationnaires) dans une classe commune (correspondant à un état donné). Selon la finalité recherchée, la méthode peut identifier deux types d'éléments (spatial ou comportemental) :

- Des phases stationnaires correspondant à des domaines vitaux temporaires lorsqu'elle est basée sur des séries chronologiques de localisation bivariées.
- Des modes de déplacement lorsqu'elle est basée sur des séries chronologiques bivariées de mesures telles que la vitesse et la tortuosité. C'est ce type de segmentation, que l'on nommera segmentation comportementale, qui a été utilisée pour identifier les remises des sangliers et chevreuils.

La segmentation comportementale se base sur le fait qu'en général, les animaux en recherche de nourriture alternent entre un mode de recherche intensive (zone concentrée en ressources), caractérisé par une tortuosité élevée et une faible vitesse, et un mode de recherche extensive (transit), caractérisé par une tortuosité faible et une vitesse élevée.

Ainsi, dans la méthode *segclust2d*, les longueurs de pas et les angles de rotation ont été calculés à partir des données GPS, puis un traitement de ces 2 métriques a permis de segmenter les données en fonction du comportement de l'animal.

Pour l'espèce sanglier, la segmentation comportementale a été effectuée pour chaque semaine de suivi et pour chaque individu, avec des données GPS d'une acquisition toutes les heures.

Pour l'espèce chevreuil, les données GPS étant disponibles selon différents temps d'acquisition (6h, 3h, 1h ou 10min), deux segmentations comportementales ont pu être faites pour chaque individu suivi :

- Une segmentation pour chaque semaine de suivi, avec des données d'une acquisition toutes les heures (comme pour le sanglier).
- Une segmentation pour chaque série de données acquises toutes les 10 minutes pendant 24 heures.

A l'issue de la segmentation comportementale, deux produits graphiques sont générés :

- Une carte représentant chaque point GPS de la couleur du segment auquel il appartient.
- Deux graphiques représentant soit la longueur du pas, soit l'angle, en fonction du temps et où chaque élément du graphique est représenté par une couleur, symbolisant le segment auquel appartient la série de données GPS.

Deux valeurs seuils de remise, basées sur la médiane et la variance de la longueur de pas ont ensuite été définies en étudiant un à un les résultats de la segmentation comportementale pour chaque espèce. Ainsi, toutes les données GPS en deçà de ces 2 valeurs seuils, correspondant donc aux remises, ont été extraites.

Le script détaillé de cette méthode figure en Annexe 2.

b. Résultats et interprétations

i. Sangliers

Pour la majorité des résultats de la segmentation, on remarque une alternance entre des segments caractérisés par une longueur de pas grande et très variable, correspondant à des grands déplacements (période d'activité) et des segments caractérisés par de faibles longueur de pas et de petite variance (période de remise). Sur la carte du résultat de la segmentation, les segments correspondant à des périodes de remise représentent en effet des données GPS agglomérées entre elles.

Avec cette méthode de sélection des remises, on identifie 67 segments, soit 67 remises. Ces 67 remises correspondent à 3,90% des données GPS disponibles pour les sangliers.

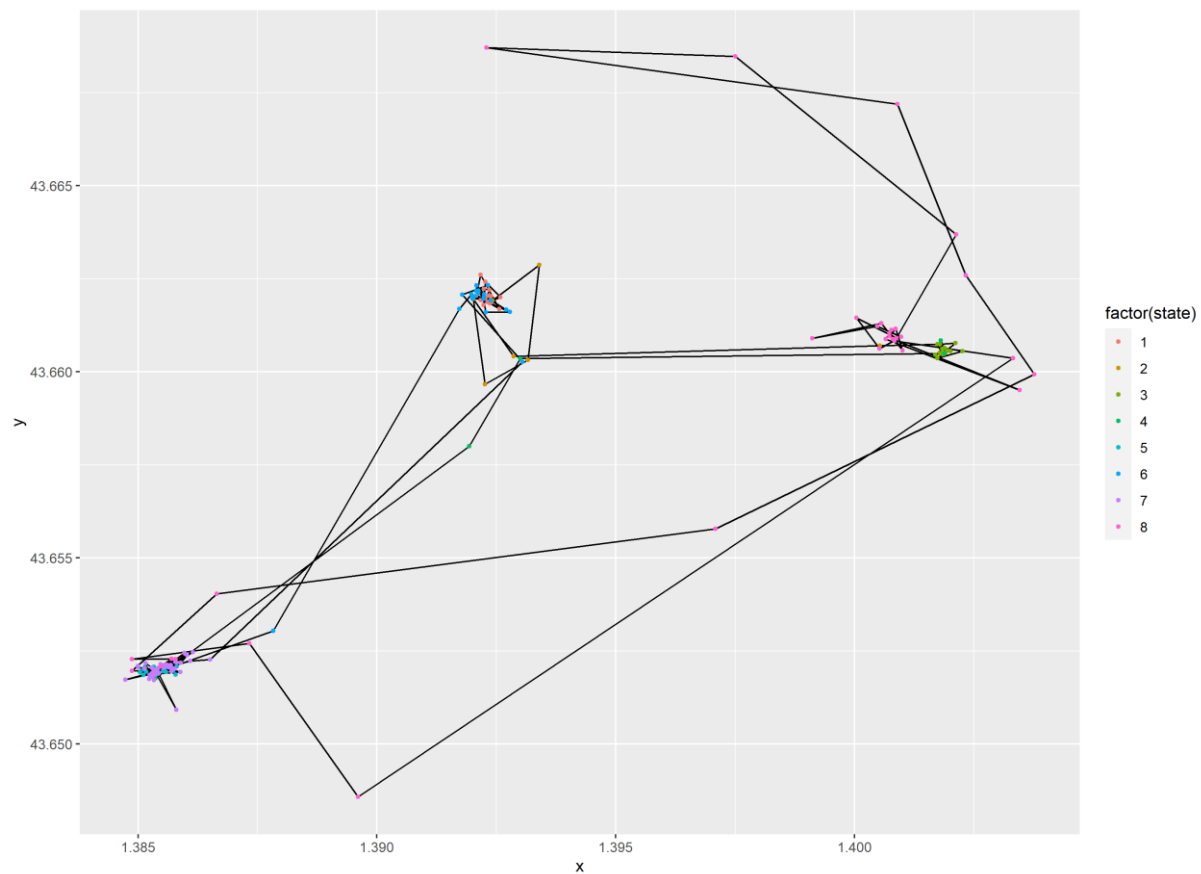


Figure 8 : Résultat de la segmentation comportementale : Carte représentant les points GPS et la couleur du segment auquel ils appartiennent. Individu : 40397. Semaine de suivi : 50.

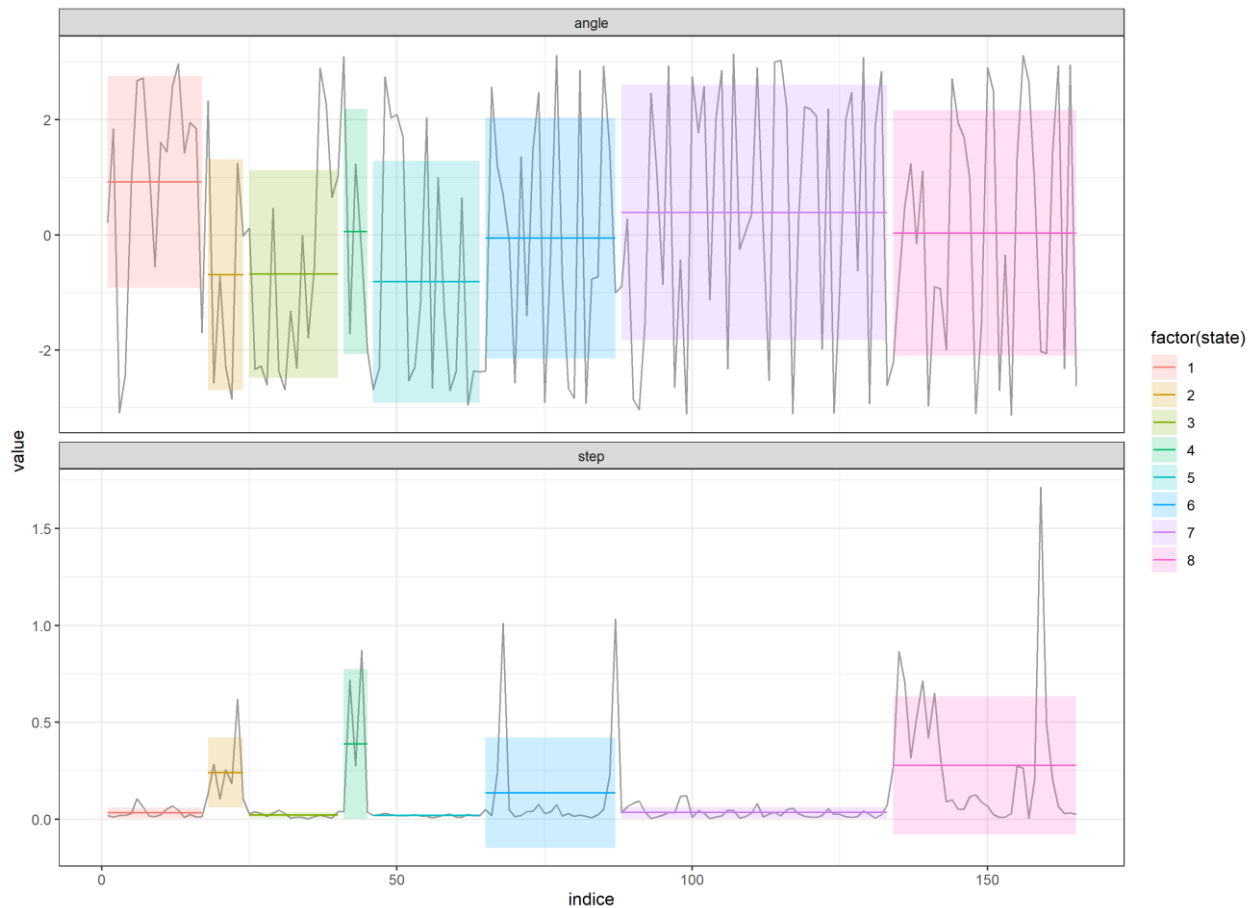


Figure 9 : Résultat de la segmentation comportementale. En haut : angle du pas. En bas : longueur du pas. Chaque couleur représente un segment de la série de données. Individu : 40397. Semaine de suivi : 50.

ii. Chevreuils

La segmentation comportementale du chevreuil effectuée sur des données acquises toutes les heures ne permet pas de visualiser des segments correspondants à des périodes de remise. En revanche, en segmentant les données acquises toutes les 10 minutes, il est possible de repérer quelques segments correspondants à des remises.

Avec cette méthode de sélection des remises, on identifie 171 segments correspondants à des remises, ce qui représente 11% des données GPS disponibles.

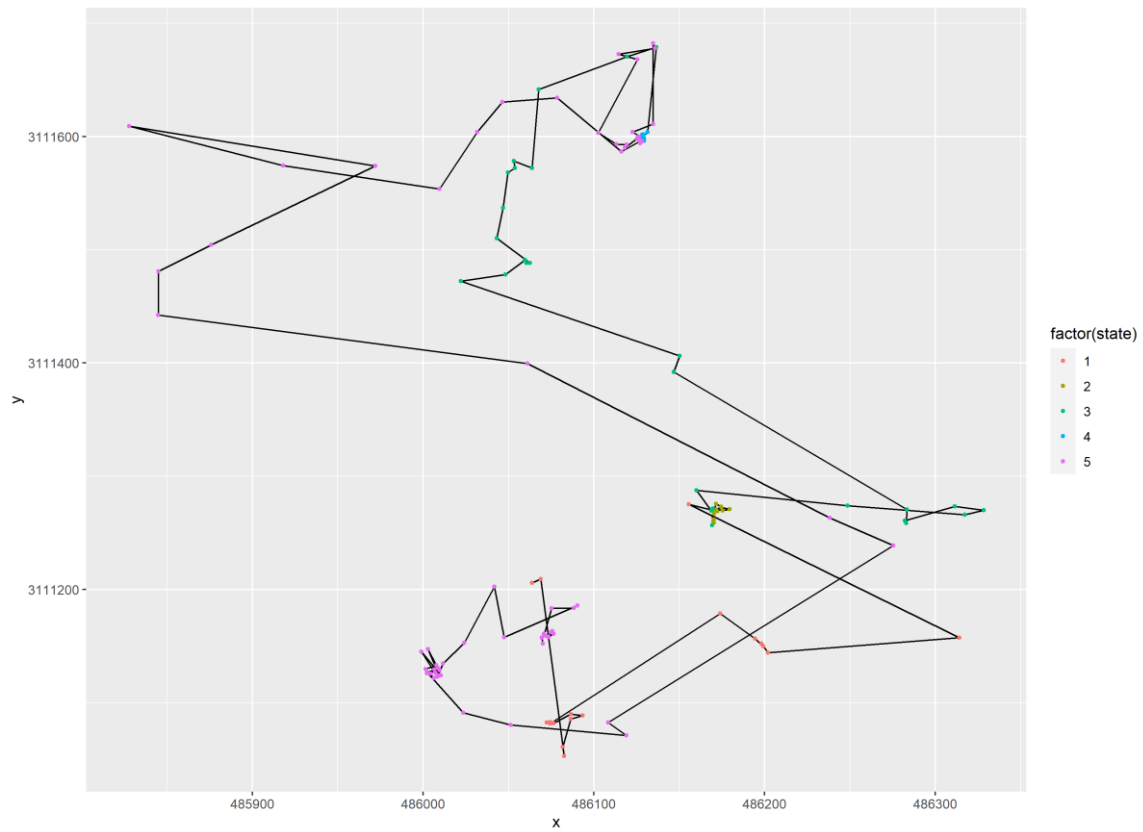


Figure 10 : Résultat de la segmentation comportementale : Carte représentant les points GPS et la couleur du segment auquel ils appartiennent. Individu : Individu : 953_16. Série de suivi de 10 min : n°7.

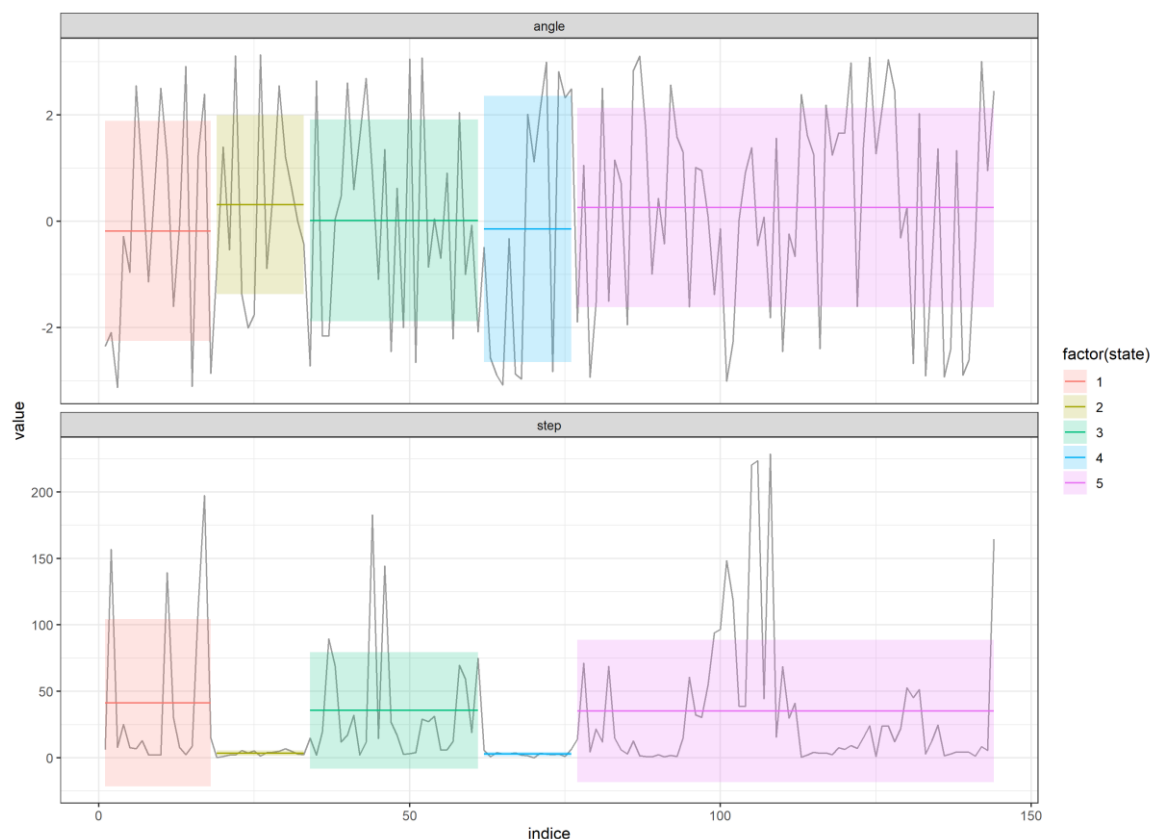


Figure 11 : Résultat de la segmentation comportementale. En haut : angle du pas. En bas : longueur du pas. Chaque couleur représente un segment de la série de données. Individu : 40397. Semaine de suivi : 50.

4. Comparaison des méthodes utilisées pour identifier remises

Parmi les 3 méthodes utilisées, il serait intéressant d'identifier la meilleure pour identifier et isoler points GPS situés dans des remises.

Il a été admis que le sanglier ne se remisait qu'exclusivement dans des boisements. Ainsi, la meilleure méthode d'identification des remises sera celle ayant le plus grand pourcentage de points-remises dans un boisement par rapport à tous les points-remises identifiés dans la méthode (Tableau 2).

Tableau 2 : Pourcentage de points-remises dans un boisement par rapport à tous les points-remises, en fonction de la méthode d'identification des remises appliquée.

<i>Méthode d'identification des remises</i>	Distance entre 2 localisations successives < 200m	5h-17h	Segmentation segclust2d
<i>Pourcentage de points-remises dans un boisement par rapport à tous les points-remises</i>	82%	77%	70%

La méthode basée sur la distance successive entre 2 localisations successives s'avère donc être la méthode la plus efficace pour identifier les remises des sangliers. Ainsi, pour identifier

les caractéristiques des remises des sangliers, nous partirons des remises identifiées et extraites des données GPS avec cette méthode.

Pour le chevreuil, il n'est pas possible d'identifier la meilleure méthode d'identification des remises puisque nous ne possédons pas d'information sur les remises des chevreuils. Ainsi, les caractéristiques des remises seront étudiées sur les points-remises issues des deux méthodes et l'on essaiera d'identifier les similitudes existantes.

B. Identification des caractéristiques des remises

L'objectif a été d'identifier les caractéristiques des points-remises isolés dans la partie précédente afin d'affiner la définition des remises dans le modèle Via Fauna et ainsi améliorer les points de départ et d'arrivée du modèle.

Par soucis de simplification, « points-remises » sera utilisé pour désigner les points identifiés comme se trouvant dans des remises, grâce aux méthodes utilisés en III.A. De la même manière « points-remises-boisements » sera utilisé pour désigner les points identifiés comme étant dans des remises et se trouvant dans des boisements.

1) Sangliers

i. Principe de la méthode

Nous l'avons vu, les sangliers ne se remettent exclusivement que dans des boisements. Ainsi, trois caractéristiques des boisements de remise ont été étudiés pour le sanglier :

- Leur taille
- Leur compacité
- Leur distance à la route et au bâti

La compacité est le rapport entre le périmètre et l'aire d'une surface : plus le résultat est proche de 0 plus le polygone est proche d'un cercle (boisement) et inversement plus il est proche de la valeur 1 et plus le polygone est proche d'une ligne (haie/ripisylves).

Pour étudier la taille et la compacité des boisements de remise, on réalise sur QGIS l'intersection entre les points de remises sangliers (identifiés avec la méthode basée sur la distance entre deux localisations successives) et les boisements avec la caractéristique à étudier fixée (taille ou compacité). On calcule ensuite le pourcentage suivant :

$$\frac{\text{Nombre de points remises dans un boisement aux caractéristiques fixées (taille ou compacité)}}{\text{Nombre de points remises dans un boisement}}$$

On gardera les critères aboutissant au plus grand pourcentage de points.

Pour étudier les distances des boisements de remise à la route et au bâti, la plus petite distance à ces éléments a été calculée pour chaque point GPS sur R, à partir des couches routes et bâti de l'IGN. Il est important de préciser que pour calculer les distances des points-remises à la route, un critère a été appliqué pour ne calculer les distances que pour les routes

d'importance et étant perçues comme des obstacles par les animaux : autoroutes, quasi-autoroutes, routes à 1 chaussée, routes à 2 chaussées et bretelles. Les chemins, routes empierrées, sentiers, pistes cyclables et escaliers ont ainsi été exclus des calculs de distance.

ii. Résultats et interprétations

Taille du boisement de remise

98% des points-remises-boisements se trouvent dans un boisement d'une surface supérieure ou égale à 2 hectares. Ces boisements ne représentaient que 0,6% des boisements de la zone d'étude.

Compacité du boisement de remise

98% des points-remises-boisements se trouvent dans un boisement d'une compacité inférieure ou égale à 0,1, donc plutôt compact. Ces boisements ne représentaient qu'1,3% des boisements de la zone d'étude.

Distance à la route et au bâti du boisement de remise

Tableau 3 : Valeurs du quantile à 2% pour les distances à la route et au bâti

Distance à ...	Q2%
Route	34 m
Bâti	9 m

Le quantile à 2% représente la plus petite distance tel que au moins 2% des valeurs sont inférieures ou égales à cette distance. En d'autres termes, en utilisant les résultats trouvés et recensés dans le Tableau 3, au moins 2% des distances à la route des remises-sangliers sont inférieures ou égales 34 m. On a donc 98% des remises des sangliers qui se trouvent à une distance strictement supérieure à 34 mètres de la route.

Sur le Tableau 3, on constate que les remises des sangliers semblent être plus proches des bâtiments qu'elles ne le sont de la route.

Ainsi la remise définie dans le modèle Via Fauna semble bien correspondre aux caractéristiques écologiques des remises des sangliers (boisement de plus de 2 hectares). Nous pouvons jouer sur les caractéristiques de la compacité et la distance à la route et au bâti pour affiner la définition des remises dans le modèle.

2) Chevreuils

i. Principe de la méthode

La définition des remises des chevreuils repose sur la même méthode que le travail effectué précédemment sur les sangliers. L'occupation du sol des remises a été étudiée en complément.

ii. Résultats et interprétations

a. Occupation du sol des remises

Tableau 4 : Pourcentage de points-remises se trouvant sur une occupation du sol (OS) donnée, en fonction de la méthode d'identification des remises choisie

OS / Méthode d'identification des points-remises utilisée	Distance entre 2 localisations successives <20m	Segmentation segclust2d
Formations herbacées indifférenciées en contexte agricole	64,5%	66,9%
Peuplements forestiers et arbustifs hors contexte urbain	35,3%	33,0%
TOTAL	99,8 %	99,9 %

Quelle que soit la méthode d'identification des remises utilisée, plus de 99% des remises des chevreuils se trouvent soit sur une occupation du sol de type agricole soit dans un boisement. Il semble également que les remises des chevreuils se trouvent en majorité sur des « formations herbacées indifférenciées en contexte agricole ». Cette dénomination est le fruit du regroupement, effectué par Anaïs Sentenac lors de la construction du modèle Via Fauna, de plusieurs occupations et usages du sol de l'OCSGE. Elle regroupe les classes suivantes de l'OCSGE : « Formations herbacées – Agriculture », « Formations herbacées – Activité d'extraction », « Formations herbacées – Pêche et aquaculture » et « Formations herbacées – Sans usage ».

Comme les remises des chevreuils semblent se situer sur deux occupations du sol distinctes, le travail précédent de caractérisation des remises a été effectué pour chacune des deux occupations du sol. Les distances à la route et au bâti de toutes les remises du chevreuil (c'est-à-dire toutes occupations du sol confondues) ont aussi été calculées et figurent en Annexe 3.

b. Caractéristiques des remises sur OS agricole

Tableau 5 : Valeurs des quantiles à 2% et 98% pour les distances des remises-chevreuils sur OS agri à la route, au bâti et au boisement, en fonction de la méthode d'identification des remises utilisée

Distance à ...	Méthode d'identification des points-remise utilisée	Q2%	Q98%
ROUTE	Distance entre 2 localisations successives <20m	15m	X
	Segmentation segclust2d	11 m	X

BATI	Distance entre 2 localisations successives <20m	50 m	X
	Segmentation segclust2d	60 m	X
BOIS	Distance entre 2 localisations successives <20m	0,9	131
	Segmentation segclust2d	0,8	131

On constate que les remises sur OS agricole sont toujours relativement proches d'un boisement : 98% des points-remises sont à moins de 131 mètres des boisements. De plus, les remises sur OS agricole semblent être toujours plus éloignées des bâtis qu'elles ne le sont de la route.

c. Caractéristiques des remises sur OS bois

Taille du boisement de remise

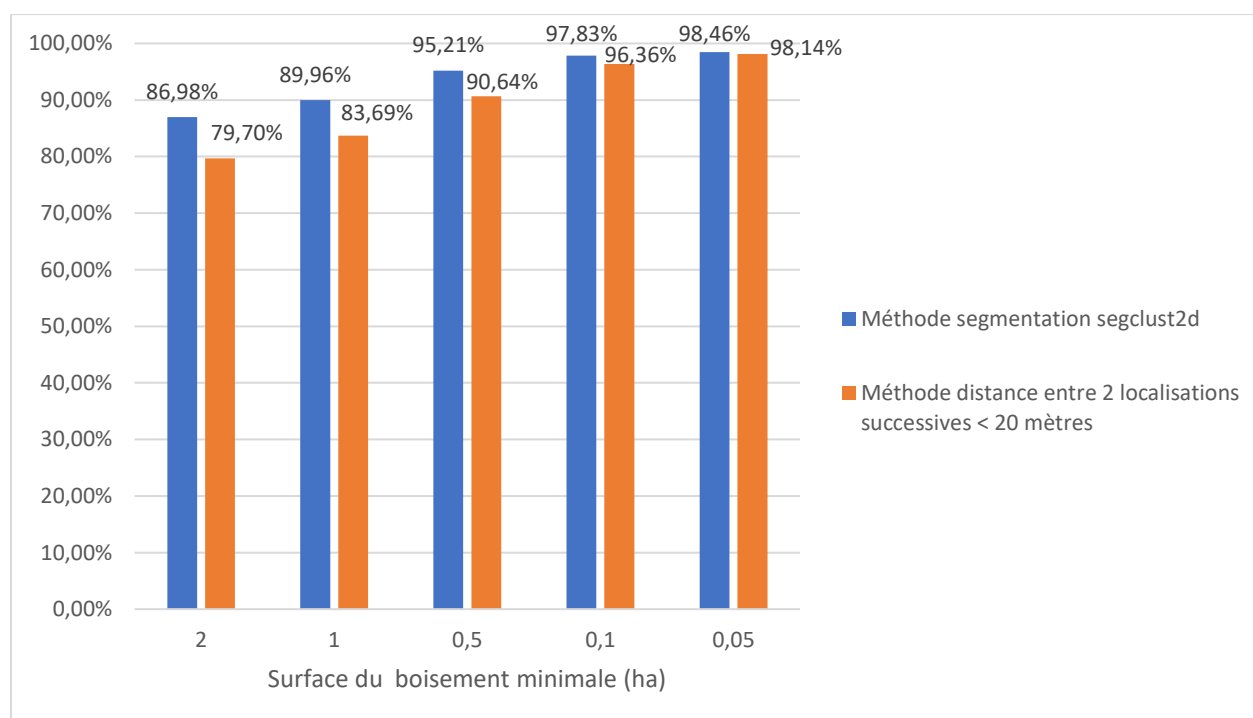


Figure 12 : Pourcentage de points-remises-boisements respectant le critère de taille appliqué

Compacité du boisement de remise

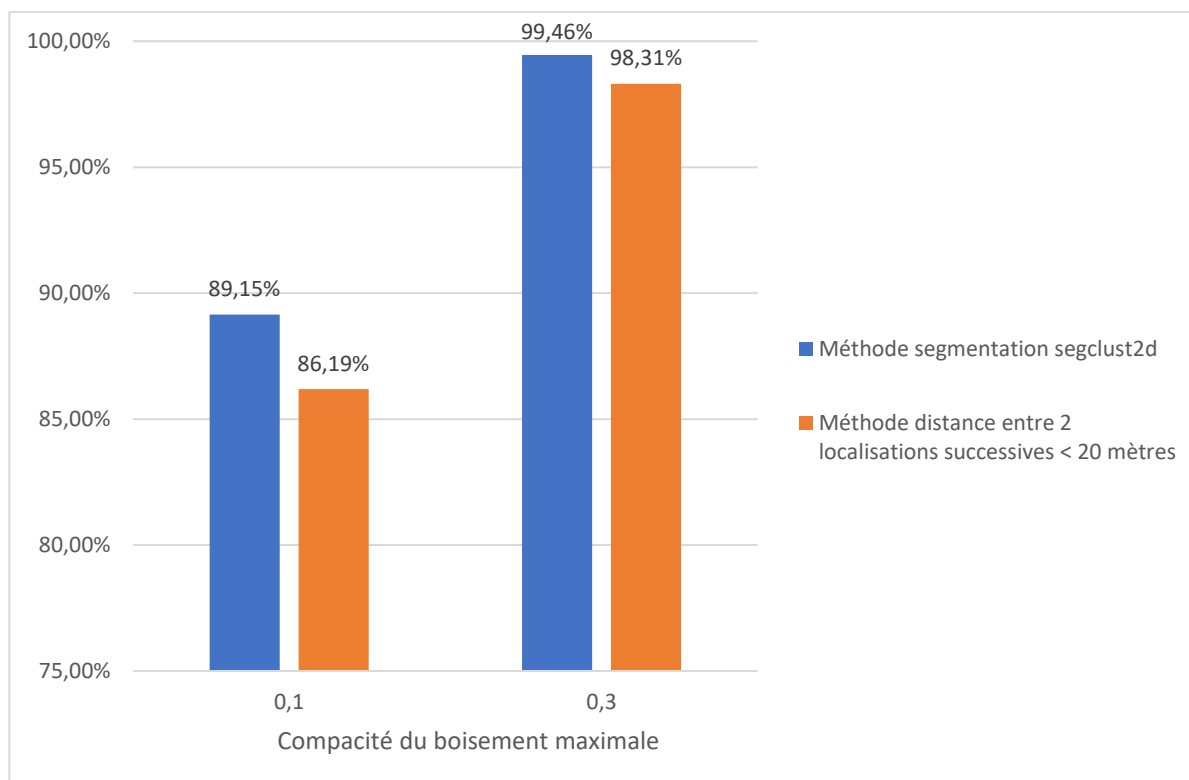


Figure 13 : Pourcentage de points-remises-boisements respectant le critère de compacité appliqué

Distance à la route et au bâti du boisement de remise

Tableau 6 : Valeurs du quantile à 2% pour les distances des remises-chevreuils-boisements à la route et au bâti, en fonction de la méthode d'identification des remises utilisée

Distance à ...	Méthode d'identification des points-remise utilisée	Q2%
ROUTE	Distance entre 2 localisations successives <20m	11 m
	Segmentation segclust2d	8 m
BATI	Distance entre 2 localisations successives <20m	35 m
	Segmentation segclust2d	31 m

Comme pour les remises de chevreuils sur OS agricole on constate que les remises sur OS boisement sont majoritairement plus éloignées du bâti qu'elles ne le sont de la route. Les remises sur OS agricole sont plus proches des routes que les remises-boisement.

3) Conclusion sur les caractéristiques des remises du sanglier et du chevreuil

Tableau 7 : Résumé des caractéristiques des remises du chevreuil et du sanglier.

NB : pour l'espèce chevreuil, chaque variable a été obtenue grâce au calcul de la moyenne de la variable dans les 2 méthodes d'identification des remises (distance entre 2 localisations successives et segmentation segclust2d).

	CHEVREUIL		SANGLIER
	Remise OS agricole chevreuil (65%)	Remise boisement chevreuil (35%)	Remise boisement sanglier
Taille du boisement	X	>= 0.05 ha	>= 2 ha
Compacité du boisement	X	<= 0,3	<= 0,1
Distance à la route	> 13 m	> 10 m	> 34 m
Distance au bâti	> 55 m	> 33 m	> 9 m
Distance au boisement	< 131 m	X	X

D'une part, cette étude a montré que l'espèce fictive du modèle Via Fauna correspond assez bien aux caractéristiques écologiques des remises du sanglier, en tout cas pour cette zone d'étude précise. En effet, pour cette structure paysagère donnée, les remises des sangliers se trouvent en majorité dans des boisements d'une taille supérieure ou égale à 2 hectares, comme il avait été admis initialement lors de la construction du modèle Via Fauna. Il serait intéressant de conduire de nouveau cette étude dans une zone avec une structure paysagère tout autre (par exemple une zone beaucoup plus rurale) pour voir si la définition des remises dans le modèle Via Fauna correspond toujours aux caractéristiques écologiques des sangliers en termes de remise.

D'autre part, l'étude a montré que pour cette structure paysagère donnée, les caractéristiques écologiques de l'espèce fictive du modèle Via Fauna ne correspondent pas à celles du chevreuil en termes de remises. En effet, pour une telle structure paysagère, le chevreuil se remisera en majorité dans un boisement plus petit et moins compact que ce qu'il a été admis dans le modèle Via Fauna. De la même manière que pour le sanglier, il serait intéressant de conduire de nouveau l'étude de caractérisation des remises des chevreuils dans une zone avec une structure paysagère tout autre (une zone plus urbanisée par exemple).

On observe que les chevreuils, pendant leur période de remise, occupent des boisements plus petits et moins compacts que les sangliers. De plus, quelle que soit l'occupation du sol considérée (boisement ou OS agricole), les remises du chevreuil sont plus proches des routes que celles des sangliers. Au contraire, les sangliers occupent des boisements plus proches du bâti par rapport au chevreuil. Cependant les résultats de cette étude et les interprétations qui en découlent peuvent largement être biaisés par le fait qu'elles aient été menées sur deux zones très différentes pour les deux espèces. En effet, les sangliers se trouvent dans une zone très urbanisée, au cœur de l'agglomération toulousaine, alors que les chevreuils occupent quant à eux un territoire beaucoup plus rural, loin de la ville et des infrastructures de transport majeures. La promiscuité des sangliers au bâti par rapport aux chevreuils s'explique donc très probablement par la densité de bâti beaucoup plus élevée dans la zone d'étude des sangliers. De la même façon, les routes des deux zones d'étude sont très différentes, notamment en termes

de fréquentation. Il est très probable que les remises des sangliers soient en moyenne plus éloignées des routes de la zone d'étude, plus fréquentées et représentant un réel obstacle pour les animaux, par rapport à la zone d'étude des chevreuils, qui eux se remettent dans des espaces relativement proches des routes de faible fréquentation et donc peu dangereuses.

Il serait donc très risqué de vouloir dresser une comparaison entre les remises des sangliers et celles des chevreuils puisqu'elles se situent dans deux zones aux structures paysagères très différentes. Pour s'affranchir de ce biais et pouvoir dresser une comparaison inter-espèce rigoureuse, il serait nécessaire de conduire l'étude de caractérisation des remises dans une zone d'étude commune aux chevreuils et sangliers ou du moins dans deux zones d'étude avec la même structure paysagère.

C. Discussion et perspectives d'amélioration

Cette étude fournit des caractéristiques supplémentaires concernant les remises des chevreuils et des sangliers, qui pourront être intégrées dans la définition de l'habitat dans le modèle Via Fauna. Ainsi, cette nouvelle définition des remises pourra être utilisé comme outil d'aide à la décision lors de projets d'aménagements du territoire. Il est néanmoins essentiel d'insister sur le fait que les caractéristiques des remises mises en évidence dans cette étude ne sont applicables que pour la zone d'étude dans laquelle le travail a été effectué.

Dans le Tableau 8 figurent quelques pistes qui auraient pu être explorées pour améliorer la définition des remises dans le modèle Via Fauna et qui pourront l'être si ce travail venait à être repris.

Tableau 8 : Pistes à explorer pour améliorer la définition des remises dans le modèle Via Fauna.

Pistes à explorer	Description	Aide bibliographique
<p>Intégrer la notion de réponse fonctionnelle dans la modélisation</p>	<p>Pour que le modèle Via Fauna puisse être utilisé sur tout type de territoire, il est essentiel d'intégrer la notion de réponse fonctionnelle dans la modélisation : l'animal sélectionne son habitat différemment en fonction de la structure paysagère.</p>	<p>Morellet, N., Van Moorter, B., Cargnelutti, B. <i>et al.</i> (2011). Landscape composition influences roe deer habitat selection at both home range and landscape scales. <i>Landscape Ecol</i> 26, 999-1010. https://doi.org/10.1007/s10980-011-9624-0</p> <p>Dupke, C., Peters, A., Morellet, N. & Heurich, M. (2021). Holling meets habitat selection – functional response of large herbivores revisited. 10.21203/rs.3.rs-208563/v1.</p> <p>Amburgey SM, Miller DAW, Rochester CJ, et al. The influence of species life history and distribution characteristics on species responses to habitat</p>

		fragmentation in an urban landscape. <i>J Anim Ecol.</i> 2021;90: 685–697. https://doi.org/10.1111/1365-2656.13403
Identifier les points GPS localisés dans des remises en fonction de la saison de l'année	En effet le paysage change en fonction de la saison considérée, notamment en milieu agricole où les cultures ne sont pas les mêmes selon le mois de l'année. Les animaux sont alors susceptibles de sélectionner leur habitat de manière différente en fonction de la saison. On revient ici à la notion de réponse fonctionnelle	
Utiliser les données d'activité pour identifier les points GPS localisés dans des remises	Il existe un package R qui segmente les données GPS en se basant sur les données d'activité (cf aide bibliographique). Cette méthode aurait permis d'être plus précis dans l'identification des points GPS se trouvant dans des remises. On aurait ainsi pu s'affranchir de la méthode basée sur la distance entre 2 localisations successives où le choix de la distance de 20 mètres pour les chevreuils fût au final très arbitraire.	Michelot, T., Langrock, R., & Patterson, T. A. (2016). moveHMM: An R package for the statistical modelling of animal movement data using hidden Markov models. <i>Methods in Ecology and Evolution</i> . https://doi.org/10.1111/2041-210X.12578 McClintock, B. T., & Michelot, T. (2018). momentuHMM: R package for generalized hidden Markov models of animal movement. <i>Methods in Ecology and Evolution</i> , 9(6), 1518–1530. https://doi.org/10.1111/2041-210X.12995
Identifier les points GPS localisés dans des remises en fonction de la période de chasse	On aurait ainsi pu observer l'influence de la chasse sur le lieu de remise des animaux.	

IV- Amélioration des notes de friction attribuées à chaque occupation du sol

Nous l'avons vu, le second paramètre clé du modèle Via Fauna sont les frictions attribuées à chaque occupation du sol. Dans le modèle, ces dernières sont les mêmes qu'il s'agisse de l'espèce chevreuil ou de l'espèce sanglier. Aux vues des données GPS, il semble pourtant que ces 2 espèces utilisent l'occupation du sol de manière différente pour leurs déplacements. Par exemple, le chevreuil semble beaucoup plus affranchi du milieu boisé que ce soit pour des déplacements ou dans des remises, par rapport au sanglier.

Ainsi, l'objectif de cette partie est d'identifier pour chacune des 2 espèces considérées, les notes de frictions attribuées à chaque occupation du sol (OS) permettant de traduire le plus fidèlement possible l'utilisation du sol lors de leurs déplacements réels. Pour cela, une analyse de sensibilité des frictions a été conduite.

A) Principe de l'analyse de sensibilité

1. Les différents cas de friction étudiés

Pour étudier la sensibilité des notes de frictions, il convient d'envisager différents cas de frictions. Dans chaque cas, le principe est de modifier une par une la valeur de la friction attribuée à l'OS considérée, les autres restant fixées à la valeur de friction initialement admise lors de la construction du modèle Via Fauna. On voit alors, dans chaque cas de friction, l'effet que cela produit sur la sortie du modèle, à savoir les chemins de moindre coût (CMC).

Pour ne pas avoir à modifier une à une les frictions des 28 classes d'occupation du sol existantes, une nouvelle nomenclature, avec seulement 13 classes d'occupation du sol, a été créée (Annexe 4) et l'analyse de sensibilité ne se basera que sur les 3 occupations du sol majoritairement utilisées lors du déplacement des Ongulées : les boisements, les haies et les formations herbacées en milieu ouvert.

Tableau 9 : Cas de friction étudiés pour l'analyse de sensibilité des frictions.

	CAS VIA FAUNA	CAS 1 Haies défavorisées	CAS 2 MO défavorisé	CAS 3 MO favorisé
Haies	1	2	1	1
Formations herbacées en milieu ouvert	2	2	3	1
Boisements (hors-urbain)	1	2	1	1

2. Lancement du modèle Via Fauna pour chaque cas de friction

Afin de bien observer l'effet de la variation du paramètre friction sur la sortie du modèle, il est essentiel de fixer le second paramètre du modèle Via Fauna : les points de départ et d'arrivée des CMC. Pour cela, pour chaque cas de friction, les CMC seront générés entre les points GPS d'arrivée et de départ des trajets réels des animaux.

3. Confrontation des trajectoires des CMC avec les trajectoires GPS réels

L'objectif est ensuite de confronter les CMC générés dans chaque cas de friction et pour chaque individu aux trajets GPS réels de ces derniers. Comparer visuellement sur une carte si le CMC généré se rapproche plus ou moins du trajet GPS réel n'est pas la bonne stratégie à adopter et serait de toute façon peu rigoureux. En fait, il est essentiel de s'intéresser aux caractéristiques des trajets, représentées par des variables quantitatives, pour pouvoir les comparer au mieux.

Ainsi, pour chaque point du CMC et pour chaque point du trajet de dispersion les variables suivantes sont calculées :

- Distance à la route la plus proche
- Distance au bâti le plus proche
- Si le point se trouve sur l'occupation du sol « Formations herbacées en milieu ouvert » : distance au boisement le plus proche
- Occupation du sol sur laquelle se trouve le point

La moyenne de chacune de ces variables pour l'ensemble du trajet GPS ou l'ensemble du CMC est ensuite calculée. Ainsi, pour chaque trajet, théorique comme réel, les variables suivantes sont disponibles et vont permettre de comparer les 2 trajets entre eux dans chacun des cas de friction étudiés :

- Distance moyenne du trajet à la route
- Distance moyenne du trajet au bâti
- Distance moyenne des points du trajet sur une OS « Formations herbacées en milieu ouvert » au boisement
- Pourcentage de points du trajet sur OS « Formations herbacées en milieu ouvert »
- Pourcentage de points du trajet sur OS bois
- Pourcentage de points du trajet sur OS haie

Pour ces 6 moyennes, la différence entre leurs valeurs sur le trajet GPS et sur le CMC a été calculée. Enfin, dans chaque cas de friction, chacune de ces 6 différences ont ensuite été sommées. Cette somme constitue la note du cas de friction : plus le cas envisagé possède une note basse, plus le CMC généré par le modèle Via Fauna dans ce cas de friction possède les mêmes caractéristiques que le trajet GPS réel et donc reflète bien la réalité. L'annexe 5 explique la façon dont les différentes variables ont été calculées grâce au script R.

Pour mener l'analyse de sensibilité des notes de frictions dans le cas de l'espèce chevreuil, les trajets de dispersion de 9 disperseurs ont été sélectionnés sur la zone d'étude figurant Figure 14.

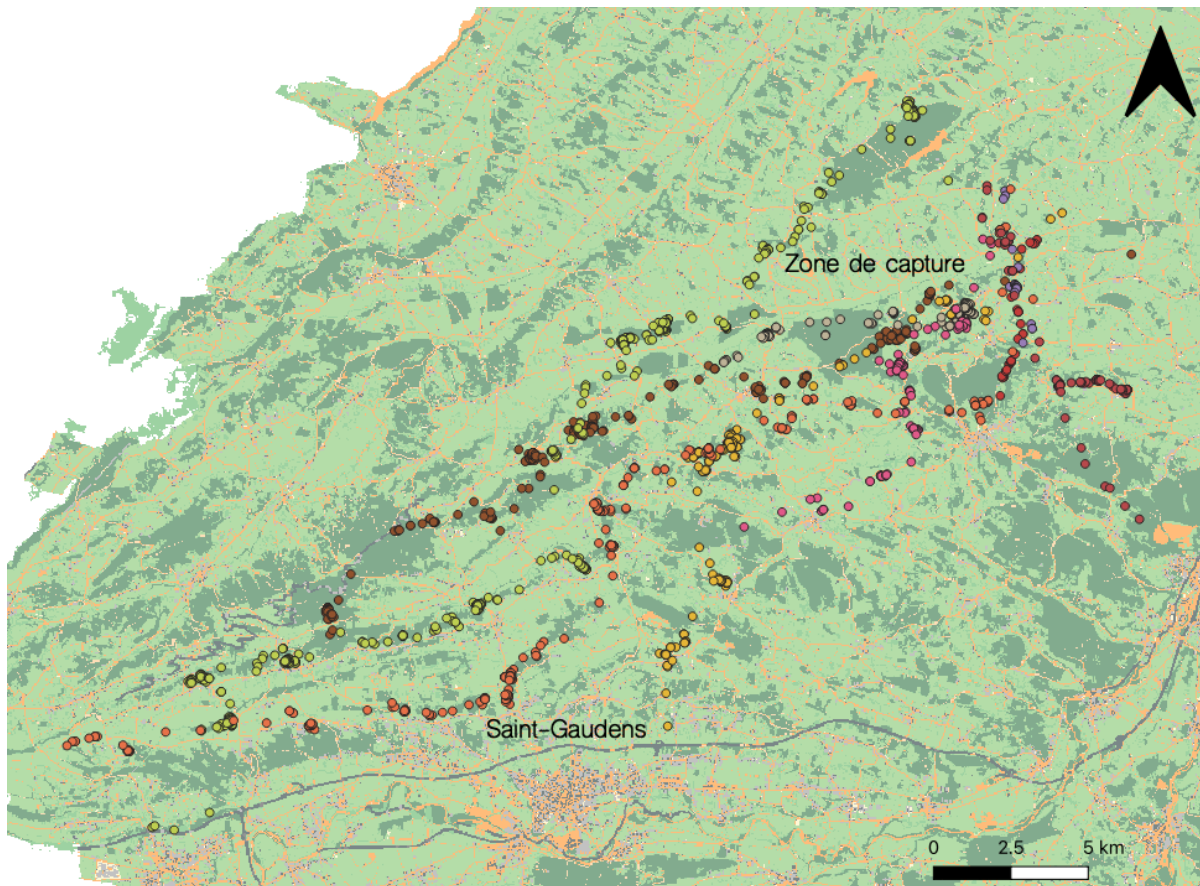


Figure 14 : Zone d'étude des chevreuils de l'analyse de sensibilité. Zone capture : Fabas, Eoux, Saint-André, Peyrissas, etc.

Pour les sangliers, 9 trajets ont été sélectionnés sur deux zones d'étude différentes : une au nord-est de Toulouse et regroupant 7 trajets de l'individu équipé du collier 40401 (Figure 15), une autre au sud-est de Toulouse et présentant 2 trajets GPS de l'individu équipé du collier 40400 (Figure 16).

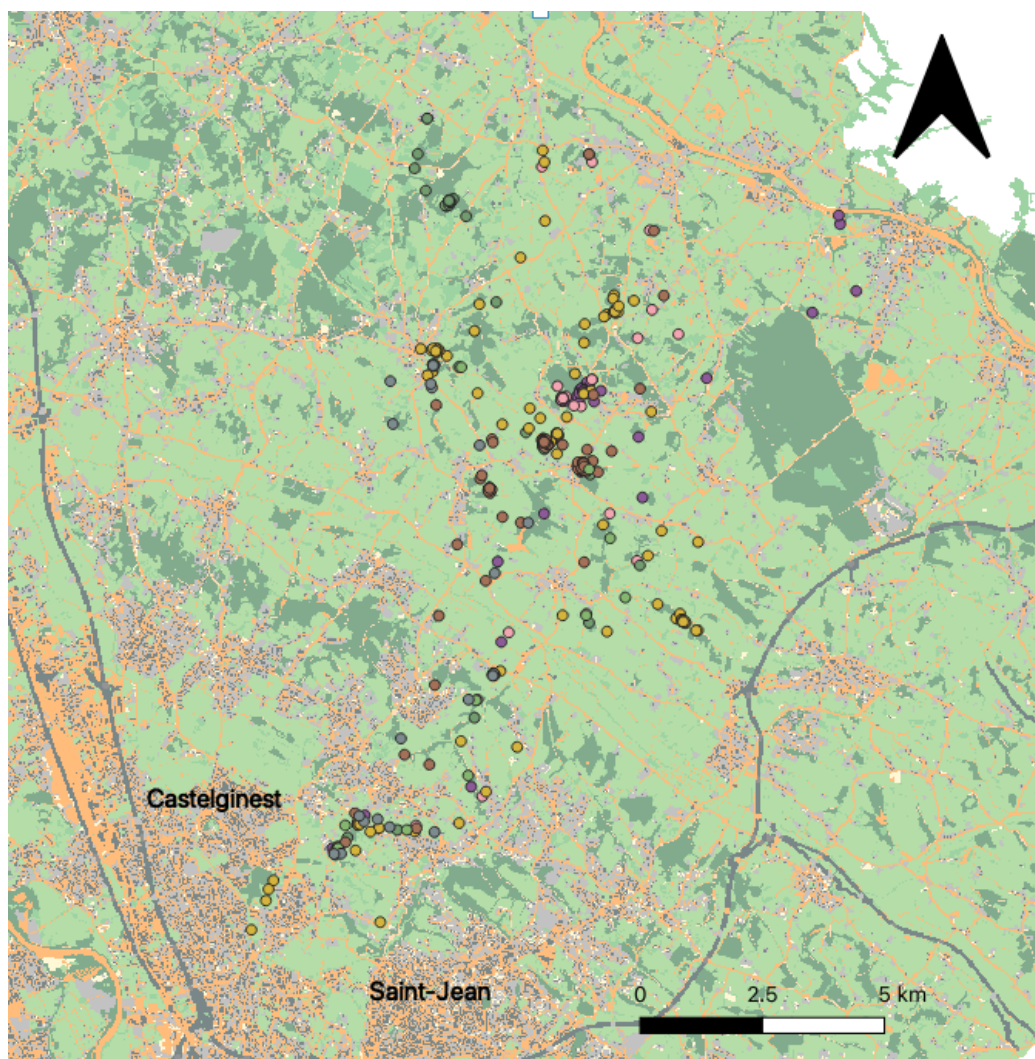


Figure 15 : Zone d'étude nord des sangliers de l'analyse de sensibilité. Individu 40401

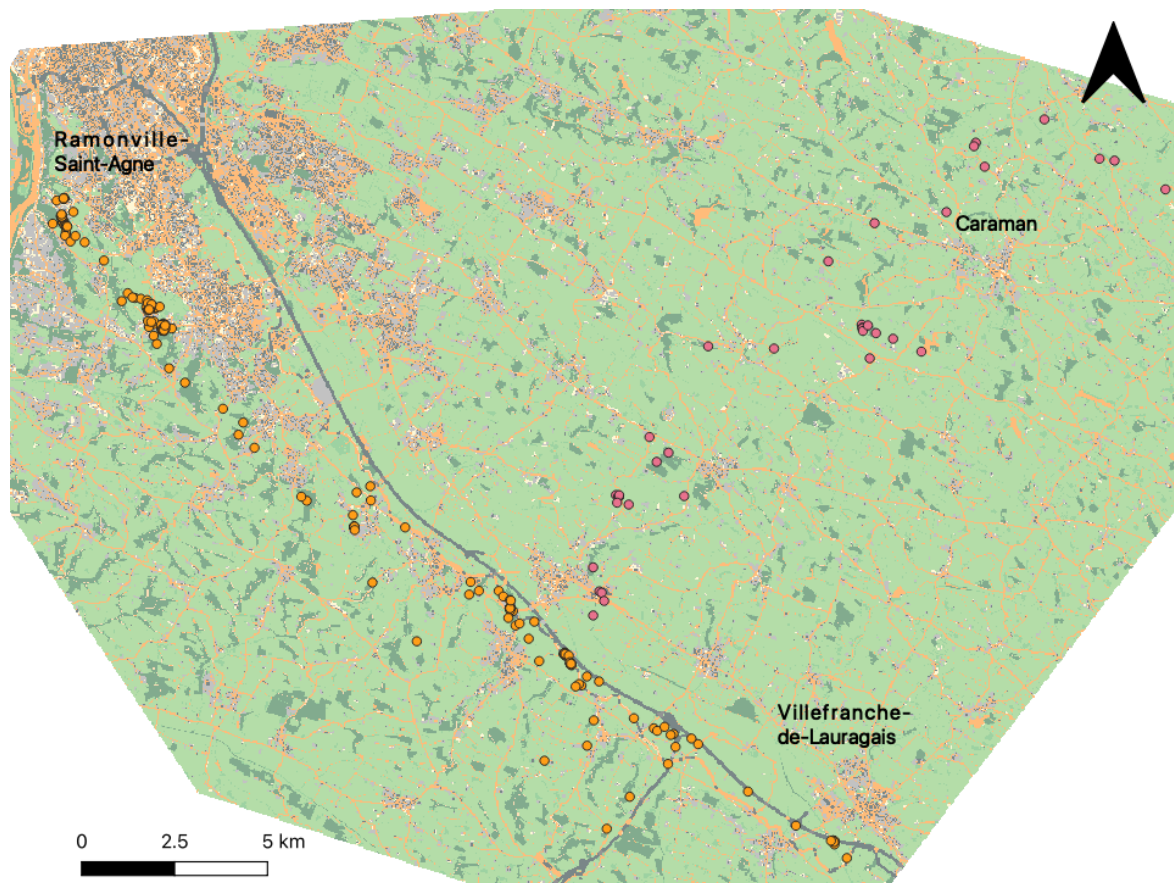


Figure 16 : Zone d'étude sud des sangliers de l'analyse de sensibilité. Individu 40400

B) Résultats et interprétations

Sur tous les graphiques de résultats présentés, la variable liée à l'élément du paysage « haie » ne sera pas représentée. En effet, pour tous les cas de friction étudiés, le pourcentage de points du trajet (CMC ou GPS) sur l'OS haie était toujours nul. Ceci est lié au fait que la carte d'occupation du sol identifie très mal les haies. A titre l'exemple, la carte d'occupation du sol générée pour la zone d'étude sud des sangliers (Figure 16) disait ne comporter aucune haie. Il conviendrait donc de revoir l'identification des haies lors de la construction de la carte d'occupation du sol pour pouvoir ensuite étudier l'utilisation réelle de ces éléments par les animaux lors de leurs déplacements.

1. Comparaison sangliers/chevreuils

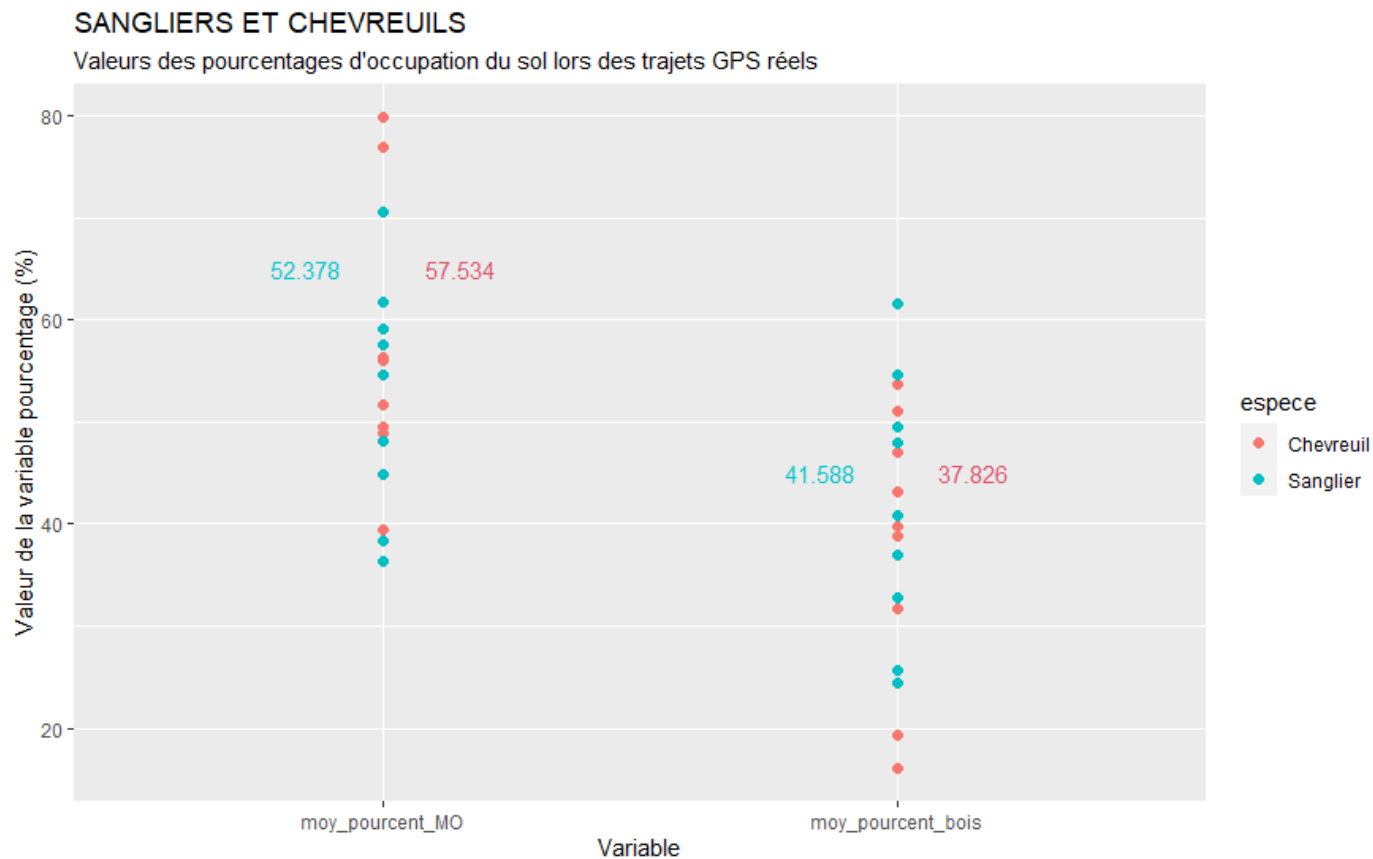


Figure 17 : Valeurs des pourcentages d'occupation du sol lors des trajets GPS des 9 individus chevreuils et des 9 trajets des 2 individus sangliers.

Tout d'abord, observons dans quelles proportions sont utilisées les deux occupations du sol étudiées (milieu ouvert et boisement) lors du trajet GPS du sanglier et du chevreuil.

Bien que les structures paysagères dans lesquelles évoluent les deux espèces soient très différentes, nous observons que globalement le chevreuil et le sanglier utilisent les OS boisement et milieu ouvert dans des proportions semblables lors de leurs déplacements GPS réels. Cette similitude dans l'utilisation de l'OS chez les deux espèces témoigne d'un certain « choix » fait par ces dernières, dans des zones à la structure paysagère très différente.

Ce constat nous permettra de minimiser l'incertitude liée aux zones d'étude différentes lorsque nous observerons et interpréterons la différence GPS/CMC pour les deux variables d'occupation du sol. Qu'en est-il quant aux variables de distances aux différents éléments du paysage ?

SANGLIERS ET CHEVREUILS

Valeurs des distances à différents éléments (route, bâti, bois) lors des trajets GPS réels

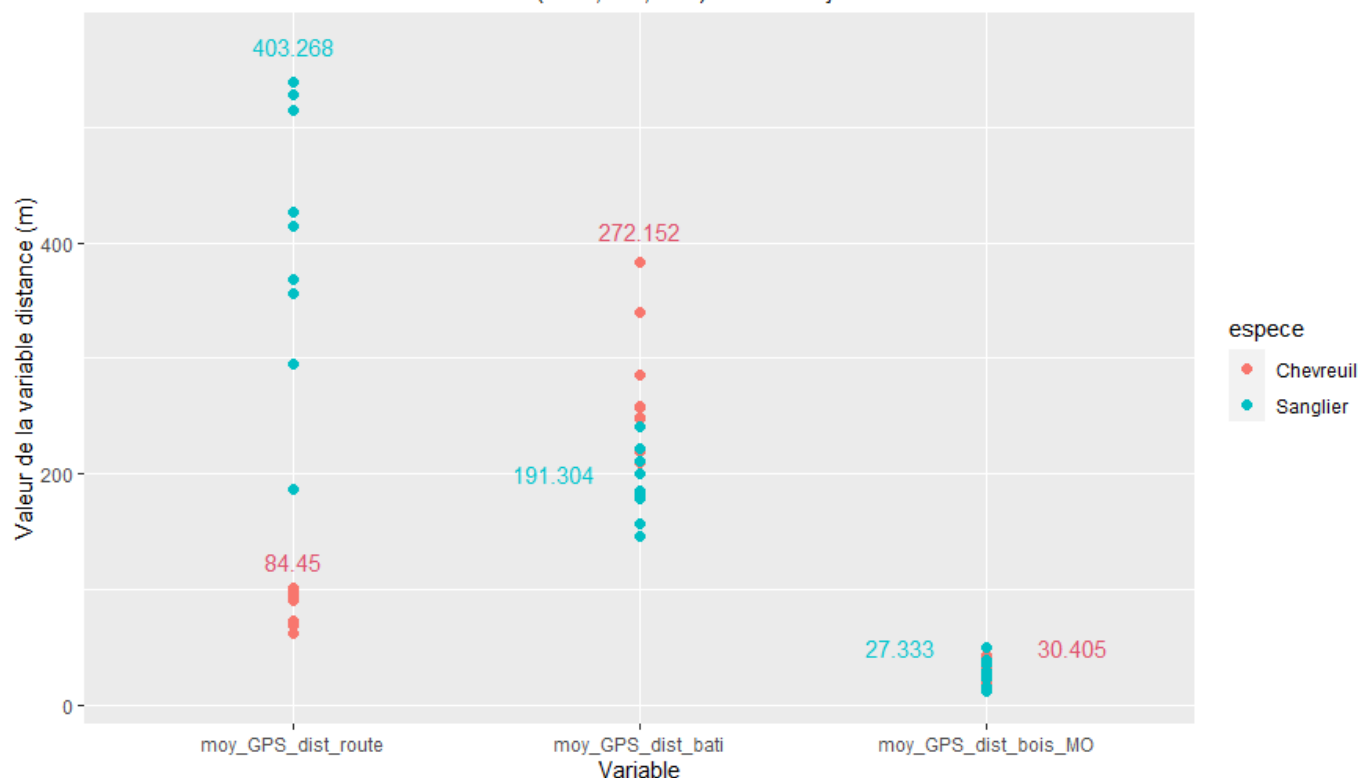


Figure 18 : Valeurs des moyennes de distance aux éléments paysagers (route, bâti et bois) lors des trajets GPS des 9 individus chevreuils et des 9 trajets des 2 individus sangliers.

Nous observons que les variables de distance à la route et au bâti ont des valeurs très différentes pour l'espèce sanglier et pour l'espèce chevreuil. Une fois de plus, ceci s'explique par le fait que les deux espèces évoluent dans des zones à la structure paysagère différente. Par exemple, on peut supposer que la grande différence de distance moyenne à la route entre le sanglier et le chevreuil (différence de plus de 300 mètres) est dû au fait que le sanglier s'éloigne plus des routes, très fréquentées en zone péri-urbaine et donc représentant un obstacle pour lui, alors que le chevreuil ne considère pas les routes de sa zone d'étude comme un obstacle majeur puisqu'elles sont très peu fréquentées. Ce constat nous montre qu'une comparaison sanglier/chevreuil est très délicate lorsque nous considérerons les variables de distance à la route et au bâti.

En revanche, nous constatons que la distance moyenne au boisement lorsque l'animal est sur une OS de milieu ouvert est globalement la même pour le sanglier et le chevreuil. Il faudrait recourir à une vérification de ces propos, mais ceci peut montrer que le chevreuil comme le sanglier semblent avoir tendance à longer les boisements lorsqu'ils se trouvent en milieu ouvert. Facilement repérables en milieu ouvert, ils possèdent ainsi un « refuge » ou « cachette » rapidement utilisable si quelque chose venait à les perturber et les effrayer (présence humaine notamment).

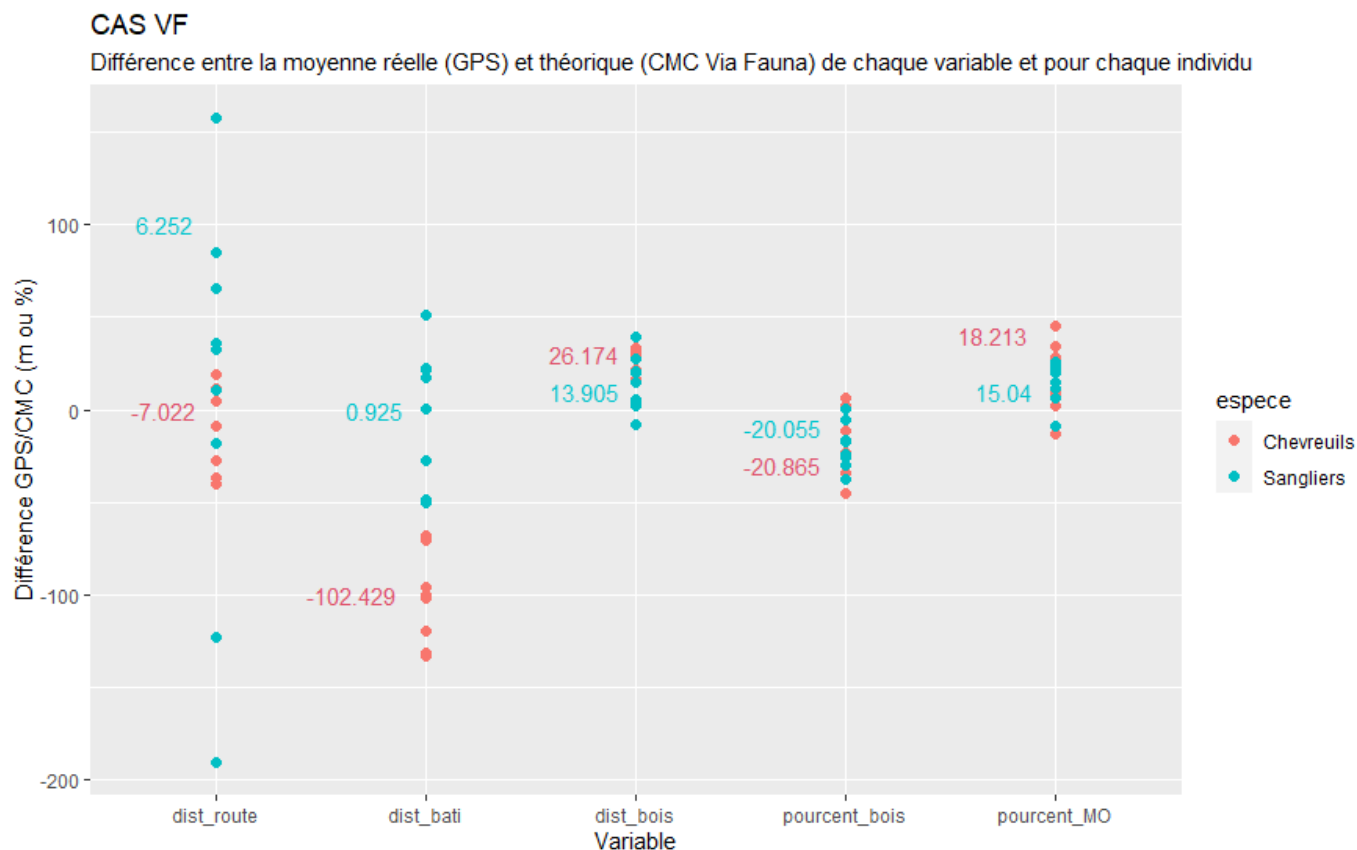


Figure 19 : Différence entre la moyenne GPS réelle (trajet GPS) et théorique (CMC Via Fauna) de chaque variable (distance à un élément du paysage ou pourcentage d'occupation du sol) pour les espèces chevreuils et sangliers.

NB : Chaque point représente un disperseur pour les chevreuils ou un trajet de l'individu 40400 ou de l'individu 404001 pour les sangliers.

Il convient maintenant d'observer la différence GPS/CMC dans le cas Via Fauna, c'est-à-dire dans un cas où les notes de frictions sont telles qu'elles ont été définies lors de la construction du modèle.

Que ce soit pour le sanglier ou pour le chevreuil, la différence GPS/CMC pour la variable pourcent_bois est négative. Cela signifie que le pourcentage de points sur une OS boisée est supérieur lors du trajet CMC par rapport au trajet GPS. Ainsi pour ces deux structures paysagères données, le modèle Via Fauna semble trop favorisé l'OS « boisements (hors-urbain) pour modéliser les OS les plus favorables aux déplacements des animaux.

A l'inverse, la différence GPS/CMC pour la variable pourcent_MO est toujours positive, que l'on considère l'espèce sanglier ou chevreuil. Cela signifie cette fois-ci que le pourcentage de points sur une OS en milieu ouvert est inférieur lors du trajet CMC par rapport à ce qu'il se passe sur le terrain, lors du trajet GPS. Ainsi pour ces deux structures paysagères données, le modèle Via Fauna semble ne pas assez favoriser l'OS « Formations herbacées en milieu ouvert ».

Nous constatons également que les différences GPS/CMC pour les variables de distance sont globalement assez faibles et distribuées autour de 0 pour le sanglier. Il semble donc que le CMC tracé par le modèle Via Fauna se rapproche assez bien du trajet GPS réel du sanglier lorsque l'on s'intéresse aux distances moyennes du trajet aux différents éléments paysagers.

2. Sangliers

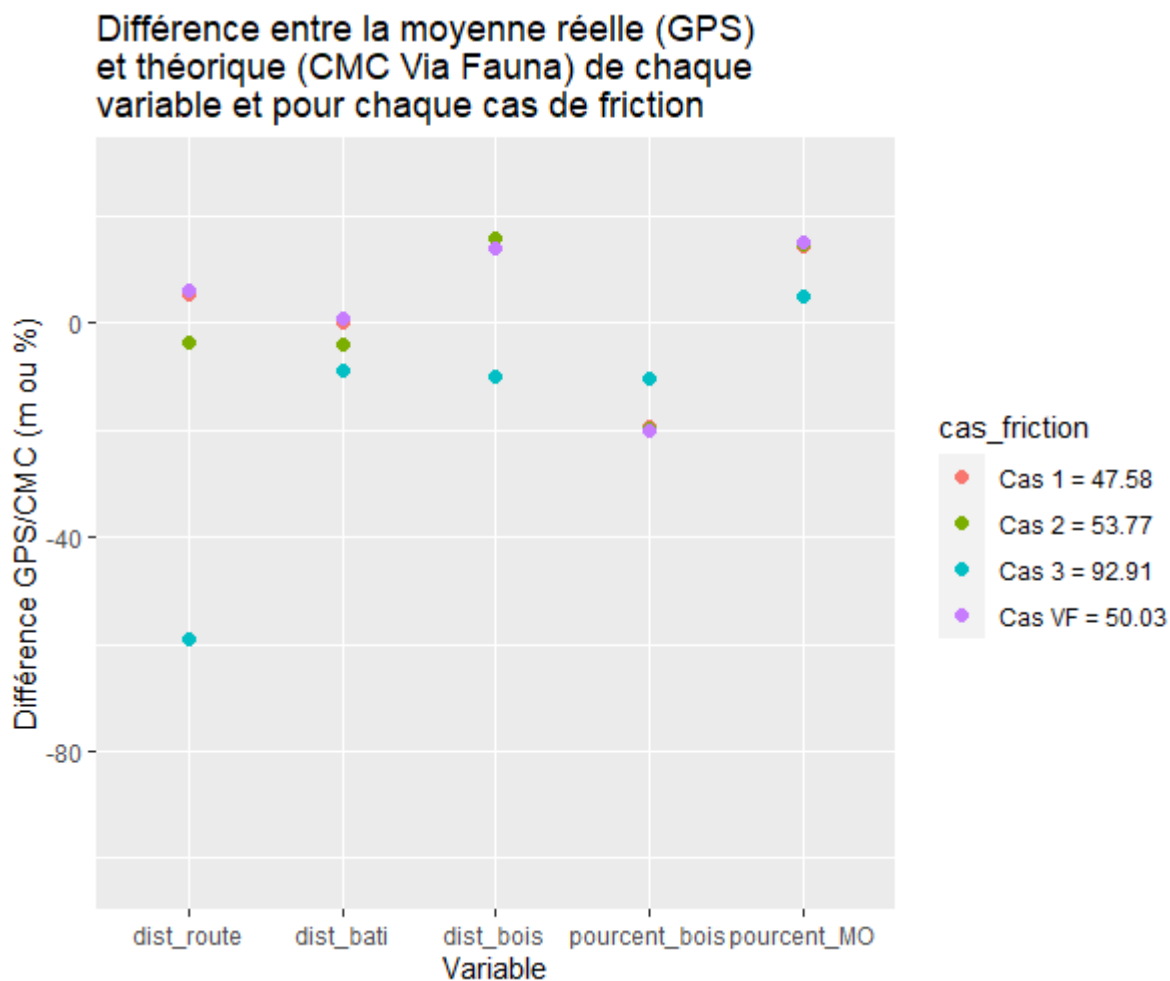


Figure 20 : Différence entre la moyenne GPS réelle (trajet GPS) et théorique (CMC Via Fauna) de chaque variable (distance à un élément du paysage ou pourcentage d'occupation du sol) pour chaque cas de friction.

NB : le nombre figurant à droite des cas de friction dans la légende représente la note de chaque cas, calculée comme la somme des différences GPS/CMC en valeur absolue pour les 5 variables d'étude.

Le cas 1, cas de friction où les haies ont été défavorisées, s'avère être le cas avec la note la plus faible. Or nous avons vu que la carte d'occupation du sol identifiait très mal les haies. Nous ne pouvons donc affirmer avec certitude que le cas 1 est le cas de friction permettant de se rapprocher au mieux des caractéristiques du trajet GPS réel.

Globalement, les notes attribuées à chaque cas de friction sont relativement faibles et homogènes si l'on ignore la différence GPS/CMC de distance à la route qui fait augmenter considérablement la note du cas 3. Ceci montre que de manière générale le modèle Via Fauna tel qu'il est construit modélise bien les meilleures occupations du sol permettant le déplacement du sanglier, en tout cas pour la structure paysagère considérée.

3. Chevreuils

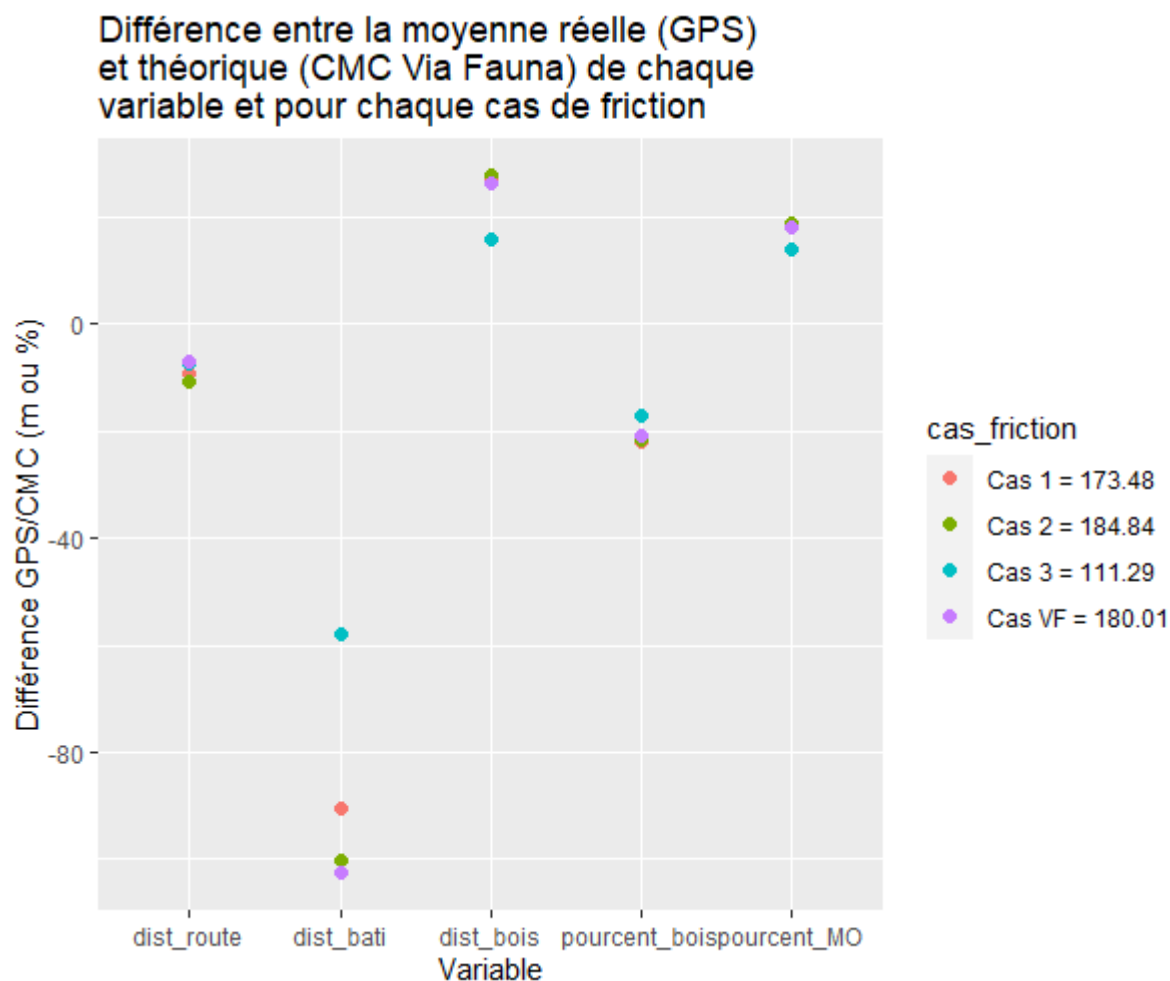


Figure 21 : Différence entre la moyenne GPS réelle (trajet GPS) et théorique (CMC Via Fauna) de chaque variable (distance à un élément du paysage ou pourcentage d'occupation du sol) pour chaque cas de friction.

NB : le nombre figurant à droite des cas de friction dans la légende représente la note de chaque cas, calculée comme la somme des différences GPS/CMC en valeur absolue pour les 5 variables d'étude.

Pour le chevreuil, le cas 3 possède de loin, la meilleure note. Ce constat est confirmé par la Figure 19, sur laquelle nous avons vu que le cas Via Fauna avait tendance à trop défavoriser les déplacements des chevreuils en milieu ouvert par rapport à la réalité.

On constate également que les notes de chaque cas de friction sont globalement bien plus élevées qu'elles ne le sont pour le sanglier, ce qui prouve que le modèle Via Fauna est globalement construit pour modéliser les meilleures occupations du sol permettant le déplacement du sanglier et non pas du chevreuil. Il convient donc d'essayer d'aller plus loin dans l'amélioration du modèle Via Fauna pour l'espèce chevreuil.

C) Construction de tampons autour des éléments paysagers

1. Principe

Nous l'avons vu, quel que soit le cas de friction considéré, les notes du chevreuil sont encore très élevées par rapport à celles du sanglier. De plus, il n'était pas forcément judicieux de comparer les variables de distance aux différents éléments du paysage entre CMC et trajet GPS puisque le modèle Via Fauna ne « regarde » pas le paysage alentour lorsqu'il trace les CMC, il utilise seulement les occupations du sol les plus favorables en longeant les pixels. Dans la réalité, lors du trajet GPS réel, l'animal a intégré son environnement, ce qu'il y avait autour de lui, et fait de réels choix lors de son déplacement.

Pour essayer d'intégrer cet aspect dans la modélisation, nous avons envisagé de nouveaux cas de frictions, cette fois-ci basés sur la construction de tampons autour de deux éléments paysagers : un tampon attractif autour des boisements et répulsif autour des routes. Pour cela, il faut en fait construire une nouvelle carte d'occupation du sol qui intégrera les deux tampons. Un guide méthodologique en Annexe 6 explique comment les cartes d'occupations du sol avec ces zones tampons ont été obtenues.

Tableau 10 : Cas de friction étudiés.

	CAS TAMP 1 Tampons autour boisements ET autour routes	CAS TAMP 2 Tampon autour boisements = friction boisements VF	CAS TAMP 3 Gradient de friction boisements/tampon boisements/MO	CAS TAMP 4 Tampon autour routes
Friction du tampon attractif à 30 mètres autour des boisements	2 (+ Friction MO = 4)	1 (+ Friction MO = 2)	2 (+ Friction MO = 4)	Rien
Friction du tampon répulsif à 90 mètres des routes	10 (Frictions routes = 10/100/10000)	Rien	Rien	10 (Frictions routes = 10/100/10000)

Dans le cas Tamp 1, les frictions des OS sont les mêmes que celles admises lors de la construction modèle Via Fauna, mis à part pour l'OS « Formations herbacées en milieu ouvert » (nommée MO dans le Tableau 10) à laquelle on attribue une friction de 4 et non pas de 2 comme dans le modèle Via Fauna. Le tampon de 30 mètres autour des boisements possédera quant à lui une friction de 2. Cependant, cela ne signifie pas qu'un bâtiment à 30 mètres d'un boisement possédera une friction de 2. En effet, l'ordre de fusion des couches lors de la construction de la carte d'OS est très important et permet d'éviter ces écueils (Annexe 6).

2. Résultats et interprétations

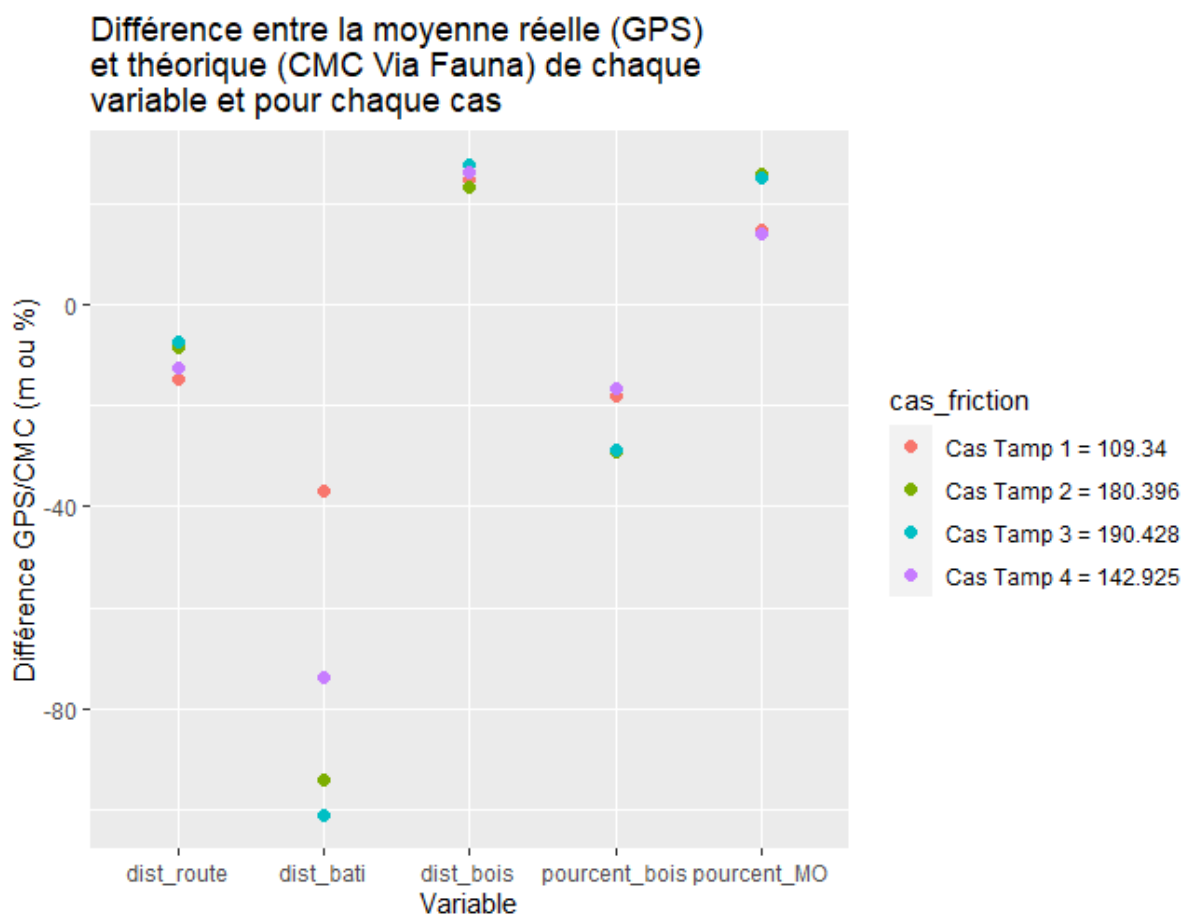


Figure 22 : Différence entre la moyenne GPS réelle (trajet GPS) et théorique (CMC Via Fauna) de chaque variable (distance à un élément du paysage ou pourcentage d'occupation du sol) pour chaque cas de tampon étudié.

NB : le nombre figurant à droite des cas de friction dans la légende représente la note de chaque cas, calculée comme la somme des différences GPS/CMC en valeur absolue pour les 5 variables d'étude.

La meilleure note revient au cas Tamp 1, soit le cas où l'on applique à la fois un tampon attractif autour des boisements et un tampon répulsif autour des routes. Ce cas permet d'obtenir une note de 109 alors qu'elle était au minimum de 111 avec le cas 3 étudié précédemment.

Rappelons que le modèle Via Fauna vise à identifier les meilleures occupations du sol permettant le déplacement de la faune sauvage et se base pour cela sur les occupations du sol et non les distances aux différents éléments paysagers. Ainsi, si l'on ne se concentre que sur les variables d'occupation du sol, c'est dans le cas Tamp 4 que la différence GPS/CMC est la plus faible, soit le cas où seul le tampon répulsif autour des routes est appliqué. Il semble donc qu'appliquer un tampon répulsif autour des routes soit ce qui contribue largement à améliorer la note, comme dans le cas Tamp 1.

D. Conclusion sur l'amélioration de la définition des notes de friction attribuées à chaque occupation du sol

Nous avons vu que pour 2 structures paysagères totalement différentes, l'utilisation du sol lors du déplacement des sangliers et des chevreuils était globalement la même. Ceci témoigne d'un réel choix fait par ces deux animaux en termes d'occupation du sol. Que ce soit

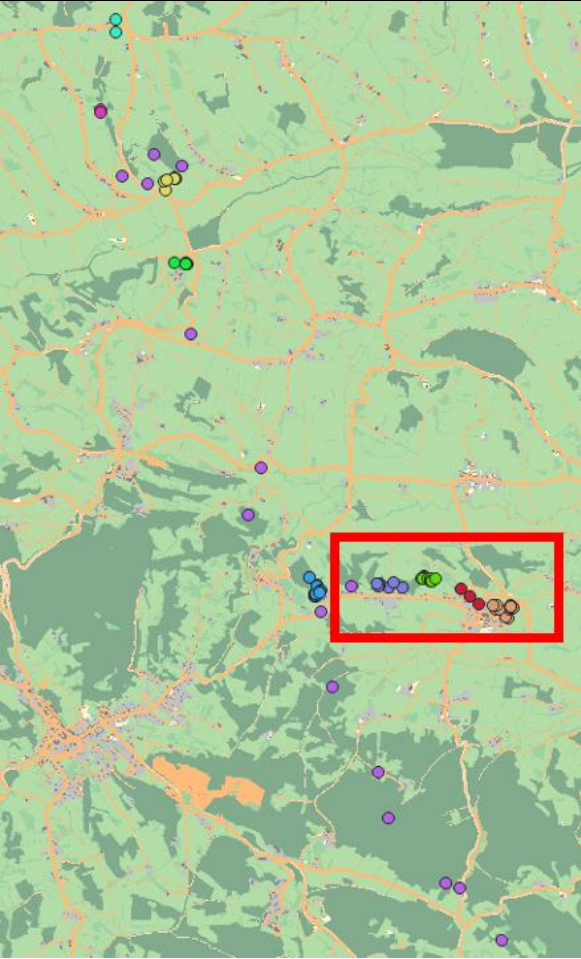
pour le chevreuil ou pour le sanglier, il semble que le modèle Via Fauna ne favorise pas assez le milieu ouvert pour modéliser les meilleures occupations du sol permettant le déplacement de ces espèces.

Les notes relativement faibles calculées pour le sanglier nous ont montré que de manière générale le modèle Via Fauna tel qu'il est construit modélise bien les meilleures occupations du sol permettant le déplacement du sanglier, en tout cas pour la structure paysagère considérée. Les notes calculées pour l'espèce chevreuil sont quant à elles bien plus élevées que dans le cas du sanglier. On comprend ici qu'il est absolument nécessaire de traiter différemment les deux espèces dans le modèle Via Fauna, que ce soit en termes de frictions ou en termes d'habitats/points de départ-arrivée comme nous l'avons vu dans la partie précédente.

Pour le chevreuil, il a été montré qu'abaisser la note de friction du milieu ouvert pour favoriser cette OS dans le modèle Via Fauna, ou encore créer un tampon attractif autour des boisements et répulsif autour des routes, était une bonne manière d'améliorer la modélisation, en tout cas pour cette structure paysagère donnée.

E. Discussion et perspectives d'amélioration

Pistes à explorer	Description	Aide bibliographique
<p>Intégrer la notion de réponse fonctionnelle dans la modélisation</p>	<p>Pour que le modèle Via Fauna puisse être utilisé sur tout type de territoire, il est essentiel d'intégrer la notion de réponse fonctionnelle dans la modélisation : l'animal sélectionne les occupations du sol différemment lors de son trajet en fonction de la structure paysagère.</p>	<p>Vimercati G, Kruger N, Secondi J. Land cover, individual's age and spatial sorting shape landscape resistance in the invasive frog <i>Xenopus laevis</i>. <i>J Anim Ecol.</i> 2021;00:1–14. https://doi.org/10.1111/1365-2656.13445</p> <p>Amburgey SM, Miller DAW, Rochester CJ, et al. The influence of species life history and distribution characteristics on species responses to habitat fragmentation in an urban landscape. <i>J Anim Ecol.</i> 2021;90: 685–697. https://doi.org/10.1111/1365-2656.13403</p> <p>Van Moorter, B., Kivimäki, I., Panzacchi, M. and Saerens, M. (2021), Defining and quantifying effective connectivity of landscapes for species' movements. <i>Ecography.</i> https://doi.org/10.1111/ecog.05351</p> <p>Girardet, X., Conruyt-Rogéon, G. & Foltête, JC. Does regional landscape connectivity influence the location of roe deer roadkill hotspots?. <i>Eur J Wildl Res</i> 61, 731–742 (2015). https://doi.org/10.1007/s10344-015-0950-4</p> <p>Panzacchi, M., Van Moorter, B., Strand, O., Saerens, M., Kivimäki, I., St. Clair, C.C., Herfindal, I. and Boitani, L. (2016), Predicting</p>

		<p>the <i>continuum</i> between corridors and barriers to animal movements using Step Selection Functions and Randomized Shortest Paths. <i>J Anim Ecol</i>, 85: 32-42. https://doi.org/10.1111/1365-2656.12386</p> <p>Laliberté, Jérôme & St-Laurent, Martin-Hugues. (2020). Validation of functional connectivity modeling: The Achilles' heel of landscape connectivity mapping. <i>Landscape and Urban Planning</i>. 202. 103878. 10.1016/j.landurbplan.2020.103878.</p>
<p>Supprimer les points GPS compris dans des phases « d'exploration » lors du calcul de statistiques sur le trajet GPS.</p>	<p>Comme nous pouvons le voir sur la Figure 23 ci-contre, les trajets de dispersion peuvent être ponctués de phases où l'animal « explore » son environnement avant de poursuivre sa dispersion à proprement parler. Ces phases d'exploration peuvent biaiser les statistiques calculées sur le trajet GPS. Par exemple ici, la distance au bâti calculée pour le trajet GPS du chevreuil sera plus faible que ce qu'elle n'est lors du trajet de dispersion hors phases d'exploration. Il conviendrait donc ne sélectionner que les points GPS inclus dans le trajet de la dispersion sans ces phases exploratoires afin de caractériser au mieux les trajets GPS réels.</p>	 <p>Figure 23 : Trajet GPS réel du chevreuil F1410_18. En rouge : phase « d'exploration ».</p>
<p>Ne calculer que les distances aux</p>	<p>Lors de l'analyse de sensibilité les distances à tous types de routes ont été calculées. Ceci peut expliquer la</p>	

<p>routes d'importante</p>	<p>distance relativement proche des chevreuils à la route. En effet, la zone d'étude des chevreuils comporte de nombreux chemins et sentiers de forêts, qui ont été pris en compte lors du calcul des distances aux routes alors qu'ils ne constituent pas un obstacle pour ces animaux lors de leurs trajets.</p> <p>Il conviendrait d'ajouter un critère dans la sélection des routes pour le calcul de distance afin de ne pas calculer la distance aux petites routes ne constituant pas un obstacle pour les animaux (les petits sentiers de forêts, les chemins de terre, etc.).</p>	
<p>Faire une carte de friction pour le jour et une autre pour la nuit.</p>	<p>En effet, sur un territoire donné, les animaux ne se déplaceront pas de la même manière la nuit et le jour. Les routes sont par exemple beaucoup moins fréquentées la nuit que le jour, et sont donc moins perçues comme un obstacle pour les animaux.</p> <p>Il pourrait être intéressant d'élaborer deux cartes d'occupation du sol distinctes, avec des frictions différentes : une pour le jour et une pour la nuit. Par exemple, les routes se</p>	

	verront attribuer des frictions plus faibles sur la carte d'OS nuit que sur la carte d'OS jour.	
--	---	--

V- Bibliographie

Börger, L. & Fryxell, J. (2012). Quantifying individual differences in dispersal using net square displacement. *Dispersal and Spatial Evolutionary Ecology* (eds J. Clobert, M. Baguette, T. Benton & J. Bullock). Oxford University Press, Oxford.

Bunnefeld, N., Börger, L., Van Mooter, B., Rolandsen, C.M., Dettki, H., Solberg, E.J. & Ericsson, G. (2011). A model-driven approach to quantify migration patterns : individual, regional and yearly differences. *Journal of Animal Ecology*, 80, 466-476.

Cargnelutti B, Coulon A, Hewison AJM, Goulard M, Angibault J-M, Morellet N (2007) Testing Global Positioning System performance for wildlife monitoring using mobile collars and known reference points. *Journal of Wildlife Management* 71, 1380-1387.

Direction régionale de l'Environnement, de l'Aménagement et du Logement Occitanie, Etat des lieux de l'artificialisation en Occitanie en 2020, (2020).

Ducros, D., Morellet, N., Patin, R., Atmeh, K., Debeffe, L., Cargnelutti, B., Chaval, Y., Lourtet, B., Coulon, A. & Hewison, M. (2020). Beyond dispersal versus philopatry? Alternative behavioural tactics of juvenile roe deer in a heterogeneous landscape. *Oikos*. 129: 81–92.

Gérard, J.F., Teillaud, P., Spitz, F., Mauget, R., Campan, R. (1991). Le sanglier. *Revue d'écologie*. 6 : 11-67.

Linnell, J., Wahlstrom, K. & Gaillard, J.M. (1998). From birth to independence. In *The European roe deer: the biology of success* (eds R. Andersen, P. Duncan & J.D.C. Linnell), pp. 257-283. Scandinavian University Press, Oslo.

Mauget, R. (1980). Régulations écologiques, comportementales et physiologiques (fonction de reproduction), de l'adaptation du Sanglier, *Sus scrofa* L., au milieu. Thèse d'état, Université François Rabelais, Tours.

Pettorelli, N., Dray, S., Gaillard, J.M., Chessel, D., Duncan, P., Illius, A., Guillon, N., Klein, F. & Van Laere, G. (2003) Spatial variation in springtime food resources influences the winter body mass of roe deer fawns. *Oecologia*, 137, 363-369.

Sentenac Anaïs. Elaboration d'une méthodologie d'identification des ruptures de continuités écologiques terrestres vis-à-vis des infrastructures linéaires de transport existantes en Midi-Pyrénées.

Spitz, F., Janeau, G. & Valet, G. (1984). - Éléments de démographie du Sanglier dans la région de Grésigne. Acta oecol., Oecol. applic., 8 : 333-341.

Stenseth, N.C. & Lidicker, W.Z. (1992) Animal Dispersal: Small Mammals as a Model Chapman and Hall, London

VI- Annexes

Annexe 1 *Script R – Identification des données GPS situées dans des remises : méthode distance entre 2 localisations successives*

Ce document vise à présenter une des deux méthodes utilisées pour isoler les points GPS correspondant à des remises. La méthode décrite ici se base sur la distance entre 2 localisations GPS successives dans le temps. Pour rappel, une remise est défini comme tout lieu où un individu (sanglier ou chevreuil) n'effectue que de petits mouvements et où l'activité est moindre.

Principe de la méthode

En partant de la matrice de distance générée plus tôt lors du stage, la différence de distance (toujours par rapport au point de référence) entre deux données GPS successives dans le temps a été calculée pour chaque animal suivi. Un compteur a ensuite compté, pour chaque heure de suivi, le nombre de fois où l'évènement A se produisait.

A= « la distance entre 2 données GPS successives est supérieure ou égale à x mètres. »

Les données GPS pour lesquelles le compteur valait 0 ont ensuite été isolées, c'est-à-dire les données pour lesquelles la distance entre 2 points successifs est inférieure à x mètres. Elles correspondent aux remises ainsi définies dans cette méthode.

Grâce à cette méthode, on parvient donc d'une part à isoler les points GPS localisés dans des remises, et d'autre part à identifier la période de remise (si elle existe) de l'espèce étudiée en générant le diagramme en bâton du compteur de l'évènement A en fonction du temps.

Script Sanglier

```
#Importation des packages
library(ggplot2)
library(lubridate)
library(chron)

#Lecture de La matrice de distance; dans ce fichier la distance entre chaque point GPS sanglier et Le capitole (point de ref) est calculée
data<-read.csv("/Users/macbook/Desktop/ici/matrice_distance_tous_points_sangliers.csv")
```



```
Nhour=levels(as.factor(as.character(indtot$heure_arrondi)))
for (l in 1:length(Nhour)){ #Boucle-heure
  c=0
  #Création d'un sous-ensemble de données pour chaque heure de suivi
  ind_heure=subset(indtot,indtot$heure_arrondi==Nhour[l])
  for (k in 1:nrow(ind_heure)){
    #Si la différence de distance entre 2 positions GPS successives est
    supérieure à x mètres,
    #Le compteur augmente de 1
    if (ind_heure$diff[k]>=x) { #Valeur de x à changer en conséquence
      c=c+1}
    #Tableau de résultats donnant pour chaque heure de suivi le résultat
    du compteur
    tb$heure[l]=Nhour[l]
    tb$compteur[l]=c} }
  tb$heure=as.numeric(tb$heure)
  tb$CollarID=NInd[i]
  tb=tb[order(tb$heure),]
  #Collage du tableau de résultats-individu au tableau de résultats final
  L'occurrence de l'événement A pour chaque heure de la journée et pour chaq
  ue individu précédemment traité dans la boucle
  tb_res=rbind(tb_res,tb)
  #Représentation graphique
  ggplot(tb,aes(x=heure,y=compteur)) + geom_bar(stat = "identity") + scale
  _x_continuous(breaks=seq(0,23,1)) + title (main = "Diagramme en bâtons don
  nant pour chaque heure de suivi le résultat du compteur, c'est à dire l'oc
  currence de l'événement : la distance entre deux positions successives est
  supérieure ou égale à 100 mètres")
  #Sauvegarde du résultat
  setwd("/Users/macbook/Desktop/ici")
  ggsave(filename=paste(NInd[i],"_histo_matrice_distance.png"), width = 11
  , height = 8)
  print(NInd[i])}

#Suppression de la première ligne de NA du tableau de résultats
tb_res=tb_res[-1,]

#Création d'un compteur binaire, qui prend la valeur 1 si l'événement A se
produit au moins une fois et 0 sinon
tb_res$compteurbi=tb_res$compteur
tb_res$compteurbi[tb_res$compteur>=1]=1
#Attribution d'un identifiant unique pour chaque heure de chaque id_suivi
tb_res$ID_heure=paste(tb_res$CollarID,tb_res$heure,sep="_")
data$ID_heure=paste(data$CollarID,data$heure_arrondi,sep="_")
tb_res=tb_res[,c("ID_heure","compteurbi")]
#Intégration de la variable compteurbi à chaque lignes du jeu de données
ata
data=merge(data,tb_res,by.x = "ID_heure",by.y = "ID_heure")

#Les remises correspondent à toutes les données où le compteur binaire vau
t 0
```



```
remises=data[(data$compteurbi==0),]  
  
#Simplification du jeu de données recensant toutes les remises identifiées  
grâce à la méthode  
remises=remises[,c("CollarID", "UTC_Date", "UTC_Time_ok", "heure_arrondi", "x",  
"y")]  
  
#Sauvegarde des données  
write.csv(remises, "/Users/macbook/Desktop/ici/remises_dist_infxm_sangliers  
.csv")
```

Script Chevreuil

Pour les chevreuils, la méthode diffère du fait que l'on utilise les données de suivi GPS acquises toutes les 10 minutes. Ainsi il est nécessaire d'assigner une variable unique à chaque animal, pour chaque série de suivi GPS de 10min.

```
#Importation des packages  
library(ggplot2)  
library(lubridate)  
library(chron)  
  
#Lecture de la matrice de distance; dans ce fichier la distance entre chaq  
ue point GPS sanglier et Le capitole (point de ref) est calculée  
data<-read.csv("/Users/macbook/Desktop/ici/matrice_distance_tous_points_ch  
evreuil_10min.csv")  
  
#Conversion de Loc_date en format POSIXct (date/heure)  
data$loc_date<-as.POSIXct(strptime(as.factor(data$loc_date), format="%Y-%m-  
%d %H:%M:%S"), format="%Y-%m-%d %H:%M:%S")  
  
#Ordonnance du jeu de données en fonction du collier de l'animal (cap_bagu  
e), puis de manière chronologique (Loc_date)  
data=data[order(data$cap_bague,data$loc_date),]  
  
#Création d'une variable NInd donnant la liste de tous les individus du je  
u de données  
NInd=levels(as.factor(as.character(data$cap_bague)))  
  
#Suppression des NA du jeu de données  
data<-na.omit(data)  
  
#Création d'une nouvelle colonne n_suivi, initialement remplie de NA  
data$n_suivi=NA  
  
#Boucle permettant d'assigner un numéro de suivi unique à chaque animal, p  
our chaque série de suivi de 10min  
for (l in 1:length(NInd)) { #Boucle-individu  
  #Création d'un sous-ensemble de données pour chaque individu suivi  
  ind=subset(data,data$cap_bague==NInd[l])  
  #Création d'une nouvelle variable indice, initialement remplie de NA  
  ind$indice=NA  
  #La première ligne du jeu de données prend 1 comme valeur d'indice
```

```
ind$indice[1]=1
#Boucle attribuant à chaque ligne du jeu de données La valeur d'indice 0
ou 1
for (j in 2:nrow(ind)) {
  #Calcul de la différence de temps, en heures, entre deux données de su
ivi successives
  diff=as.numeric(difftime(ind$loc_date[j],ind$loc_date[j-1], units = "h
ours"))
  #Si la différence de temps entre 2 données de suivi successives est su
périeure à 20 heures, alors l'indice
#attribué à la ligne concernée est 1
  if (diff>20) {
    ind$indice[j]=1
    #Sinon, l'indice attribué est 0
  }else{
    ind$indice[j]=0}}
i=0
#Boucle attribuant un numéro de suivi à chaque série de suivi GPS
for (j in 1:nrow(ind)) {
  #Si l'indice vaut 0, alors le numéro de suivi est i
  if (ind$indice[j]==0){
    ind$n_suivi[j]=i
    #Sinon (si l'indice vaut 1), alors le numéro de suivi augmente sa va
leur de 1
  }else{
    i=i+1
    ind$n_suivi[j]=i}}
#Transfert de la variable n_suivi au jeu de données data
data$n_suivi[data$cap_bague==NInd[1]]=ind$n_suivi}

#Création d'une nouvelle variable id_suivi, issue du collage de cap_bague
et n_suivi.
#Une valeur unique est donc assignée à chaque animal, pour chaque série de
suivi GPS 10 minutes
data$id_suivi=paste(data$cap_bague,data$n_suivi,sep="-")

#Création d'une nouvelle variable donnant la liste de tous les id_suivi du
jeu de données
Nid_suivi=levels(as.factor(as.character(data$id_suivi)))

#Création d'une variable donnant l'heure arrondi (minutes et secondes élim
inées)
data$heure_arrondi<-hour(data$loc_date)

#Simplification du jeu de données
data$dist_matrice<-data$Matrice._1
data<-data[,c("cap_bague","id_suivi","loc_date","cpos_date","cpos_heure","
heure_arrondi","pos_x","pos_y","dist_matrice")]

#Boucle calculant pour chaque individu suivi, la différence de distance en
tre 2 positions GPS successives
for (i in 1:length(Nid_suivi)) {
```



```
setwd("/Users/macbook/Desktop/ici")
ggsave(filename=paste(Nid_suivi[i], "_histo_matrice_distance.png"), width
= 11, height = 8)
print(Nid_suivi[i])}

#Suppression de la première ligne de NA du tableau de résultats
tb_res=tb_res[-1,]

#Création d'un compteur binaire, qui prend la valeur 1 si l'événement A se
produit au moins une fois et 0 sinon
tb_res$compteurbi=tb_res$compteur
tb_res$compteurbi[tb_res$compteur>=1]=1

#Attribution d'un identifiant unique pour chaque heure de chaque id_suivi
tb_res$ID_heure=paste(tb_res$id_suivi,tb_res$heure,sep="_")
data$ID_heure=paste(data$id_suivi,data$heure_arrondi,sep="_")
tb_res=tb_res[,c("ID_heure", "compteurbi")]

#Intégration de la variable compteurbi à chaque lignes du jeu de données d
ata
data=merge(data,tb_res,by.x = "ID_heure",by.y = "ID_heure")

#Les remises correspondent à toutes les données où le compteur binaire vau
t 0
remises=data[(data$compteurbi==0),]
#Simplification du jeu de données
remises=remises[,c("cap_bague", "loc_date", "cpos_date", "cpos_heure", "pos_x"
, "pos_y")]

#Sauvegarde des données
write.csv(remises, "/Users/macbook/Desktop/ici/remises_dist_inf20m_chevreui
ls.csv")
```

Annexe 2 *Script R – Identification des données GPS situées dans des remises : méthode segmentation comportementale segclust2d*

Ce document vise à présenter une des deux méthodes utilisées pour isoler les points GPS correspondant à des remises. La méthode utilisée ici se base sur un package R nommé segclust2d et élaboré par R. Patin. (Patin R et al. 2019.).

Principe de la méthode

Cette méthode utilisée est capable de segmenter une série chronologique bivariée et de regrouper des segments similaires (correspondant à des phases stationnaires) dans une classe commune (correspondant à un état donné). Selon la finalité recherchée, la méthode peut identifier deux types d'éléments (spatial ou comportemental) : - Des phases stationnaires

correspondant à des domaines vitaux temporaires lorsqu'elle est basée sur des séries chronologiques de localisation bivariées. - Des modes de déplacement lorsqu'elle est basée sur des séries chronologiques bivariées de mesures telles que la vitesse et la tortuosité. C'est ce type de segmentation, que l'on nommera segmentation comportementale, qui a été utilisée pour identifier les remises des sangliers et chevreuils.

La segmentation comportementale se base sur le fait qu'en général, les animaux en recherche de nourriture alternent entre un mode de recherche intensive (zone concentrée en ressources), caractérisé par une tortuosité élevée et une faible vitesse, et un mode de recherche extensive (transit), caractérisé par une tortuosité faible et une vitesse élevée. Ainsi, dans la méthode segclust2d, les longueurs de pas et les angles de rotation ont été calculés à partir des données GPS, puis un traitement de ces 2 métriques a permis de segmenter les données en fonction du comportement de l'animal.

Paramètres d'entrée dans la méthode segclust2d : * Nombre maximal de segments : K * Longueur minimale d'un segment (exprimée en nombre de localisations) : Lmin Par défaut, Lmin ne peut être inférieur à 5 pour garantir une estimation suffisamment fiable des paramètres.

SCRIPT CHEVREUIL

Pour l'espèce chevreuil, une segmentation pour chaque série de données acquises toutes les 10 minutes pendant 24 heures a été réalisée.

Préparation des données pour la segmentation

```
#Chargement des librairies nécessaires à l'exécution du script
library(segclust2d)
library(ggplot2)
library(moveHMM)
library(lubridate)
library(chron)
library(RPostgreSQL)
library(DBI)

#Accès à la base de données CEFS
dbcefs<- dbConnect(PostgreSQL(), host="pggeodb.nancy.inra.fr", dbname="db_
cefs", user="smillet", password="millets")

#Sélection des disperseurs qui possèdent un suivi GPS toutes les 10 minute
s
data <- dbGetQuery(dbcefs,"SELECT cap_bague, cpos_date, cpos_heure, pos_x,
pos_y,
                                pos_distance_route, pos_distance_bois, pos_
distance_bati, pos_distance_haie
                                FROM public.v_aniposi_gpsgsm
                                WHERE cpos_prog10minutes= TRUE AND (cap_bag
ue='190_05' OR cap_bague='232_06'
                                OR cap_bague='280_06' OR cap_bague='338_07'
OR cap_bague='342_07' OR cap_bague='344_07'
```

```
OR cap_bague='386_08' OR cap_bague='350_07' OR cap_bague='352_07'
OR cap_bague='434_08' OR cap_bague='406_08' OR cap_bague='420_08'
' OR cap_bague='494_09' OR cap_bague='446_08' OR cap_bague='F542_09'
OR cap_bague='520_09' OR cap_bague='500_09' OR cap_bague='510_09'
' OR cap_bague='F582_11' OR cap_bague='F532_09' OR cap_bague='536_10'
' OR cap_bague='F586_11' OR cap_bague='634_11' OR cap_bague='F554_11'
OR cap_bague='664_11' OR cap_bague='658_11' OR cap_bague='660_11'
OR cap_bague='710_12' OR cap_bague='666_11' OR cap_bague='698_12'
' OR cap_bague='760_12' OR cap_bague='736_12' OR cap_bague='F732_12'
3' OR cap_bague='782_13' OR cap_bague='F1064_12' OR cap_bague='772_1
OR cap_bague='F1094_13' OR cap_bague='784_13' OR cap_bague='822_13'
4' OR cap_bague='F861_14' OR cap_bague='F1019_13' OR cap_bague='853_1
OR cap_bague='927_16' OR cap_bague='870_14' OR cap_bague='905_15'
OR cap_bague='952_16' OR cap_bague='938_16' OR cap_bague='946_16'
7' OR cap_bague='953_16' OR cap_bague='F1243_16' OR cap_bague='971_1
' OR cap_bague='3003_18' OR cap_bague='976_17' OR cap_bague='1000_17
8' OR cap_bague='F1374_18' OR cap_bague='3009_18' OR cap_bague='3027_1
_18' OR cap_bague='3061_19' OR cap_bague='F1410_18' OR cap_bague='F1347
OR cap_bague='3072_19')")
```

#Conversion de La date et de L'heure au format character

```
data$cpos_date<-as.character(data$cpos_date)
data$cpos_heure<-as.character(data$cpos_heure)
```

#Création d'une nouvelle variable Loc_date, issue du collage de la date et de L'heure des relevés GPS

```
data$loc_date<-paste(data$cpos_date,data$cpos_heure)
```

#Conversion de La date de suivi en format POSIXct

```
data$cpos_date<-as.POSIXct(strptime(as.factor(data$cpos_date), format="%Y-
%m-%d"), format="%Y-%m-%d")
```

#Conversion de L'heure de suivi en format times

```
data$cpos_heure<-chron(times. = data$cpos_heure)
```



```
#Conversion de Loc_date en format POSIXct (date/heure)
data$loc_date<-as.POSIXct(strptime(as.factor(data$loc_date),format="%Y-%m-%d %H:%M:%S"),format="%Y-%m-%d %H:%M:%S")

#Ordonnance du jeu de données en fonction du collier de l'animal (cap_bague), puis de manière chronologique (Loc_date)
data=data[order(data$cap_bague,data$loc_date),]

#Création d'une variable NInd donnant la liste de tous les individus du jeu de données
NInd=levels(as.factor(as.character(data$cap_bague)))

#Suppression des NA du jeu de données
data<-na.omit(data)

#Préparation des données : calcul de la longueur de pas et de l'angle
processeddata <- prepData(data,type="UTM", coordNames=c("pos_x","pos_y"))

#Préparation des données : suppression des lignes de NA
processeddata=na.omit(processeddata)

#Ajout d'une colonne n_suivi remplie de NA
processeddata$n_suivi=NA

# Cette boucle permet d'assigner un numéro de suivi unique à chaque animal et pour chaque phase
#de suivi de 10 minutes.
for (l in 1:length(NInd)) { #Boucle-individu
  #Création d'un sous-ensemble de données pour chaque individu suivi
  ind=subset(processeddata,processeddata$cap_bague==NInd[l])
  ind=ind[,c("step","angle","loc_date","cpos_date", "cpos_heure", "x","y")]
}
#Suppression des NA dans le jeu de données ind
ind=na.omit(ind)
#Création d'une nouvelle variable, initialement remplie de NA
ind$indice=NA
#La première ligne du jeu de données prend 1 comme valeur d'indice
ind$indice[1]=1
#Boucle attribuant à chaque ligne du jeu de données la valeur d'indice 0 ou 1
for (j in 2:nrow(ind)) {
  #Calcul de la différence de temps, en heures, entre deux données successives
  diff=as.numeric(difftime(ind$loc_date[j],ind$loc_date[j-1], units = "hours"))
  #Si la différence de temps entre 2 données successives est supérieure à 20 heures, alors l'indice
  #attribué à la ligne concernée est 1
  if (diff>20) {
    ind$indice[j]=1
  }
  #Sinon, l'indice attribué est 0
}
```

```
}else{
  ind$indice[j]=0}}
#Création d'une nouvelle variable n_suivi
ind$n_suivi=NA
i=0
#Boucle attribuant un numéro de suivi à chaque série de suivi GPS
for (j in 1:nrow(ind)) {
  #Si L'indice vaut 0, alors Le numéro de suivi est i
  if (ind$indice[j]==0){
    ind$n_suivi[j]=i
    #Sinon (si L'indice vaut 1), alors Le numéro de suivi augmente sa valeur de 1
  }else{
    i=i+1}}
for (j in 1:nrow(ind)){
  if (ind$indice[j]==1){
    ind$n_suivi[j]=ind$n_suivi[j+1]}}
#Transfert de la variable n_suivi au jeu de données processeddata
processeddata$n_suivi[processeddata$cap_bague==NInd[1]]=ind$n_suivi
}

#Création d'une nouvelle variable id_suivi, issue du collage de cap_bague et n_suivi.
#Une valeur unique est donc assignée à chaque animal, pour chaque série de suivi GPS 10 minutes
processeddata$id_suivi=paste(processeddata$cap_bague,processeddata$n_suivi,sep="-")

#Création d'une nouvelle variable donnant la Liste de tous les id_suivi du jeu de données
Nid_suivi=levels(as.factor(as.character(processeddata$id_suivi)))
```

Segmentation comportementale

```
#Création d'une matrice de NA qui se remplira avec les résultats de la segmentation
res_seg=matrix(NA,1,22,dimnames=list(NA,c('id_suivi','x','y','step','angle','loc_date','cpos_date','cpos_heure','x_ind','subsample_ind','indice','state','prop','mu.step','mu.angle','sd.step','sd.angle','state_ordered','pos_distance_route','pos_distance_bois','pos_distance_bati','pos_distance_haie'))))

#Boucle de segmentation tournant pour chaque individu et pour chaque numéro de suivi unique
for (i in 1:length(Nid_suivi)) { #Boucle-individu
  #Création d'un sous-ensemble de données pour chaque individu suivi
  indtot=subset(processeddata,processeddata$id_suivi==Nid_suivi[i])
  #Création d'un nouveau jeu de données pour chaque individu suivi et suppression des lignes NA
  ind=indtot[,c("step","angle","loc_date","cpos_date","cpos_heure","x","y","pos_distance_route",
```

```
      "pos_distance_bois", "pos_distance_bati", "pos_distance_haie")]; ind=na.omit(ind)
  if (nrow(ind)>5) {
    #Segmentation
    seg <- segmentation(ind, lmin = 5, Kmax = 20, seg.var = c("step","angle"),scale.variable = FALSE)
    #Création d'un graphique donnant la segmentation en fonction de x et en fonction de y
    plot.segmentation(seg)
    #Sauvegarde du graphique obtenu
    setwd("C:/Users/smillet/Stage/Def habitat/3- segclust2d/Chevreuils/Cpt")
    ggsave(filename=paste(Nid_suivi[i],"_segclust2d.png"), width = 11, height = 8) # ggplot2
    #Résultat statistique de la segmentation
    temp_res_seg=augment(seg)
    #Ajout d'une colonne précisant pour chaque relevé GPS le nom de l'individu concerné par la segmentation
    temp_res_seg$id_suivi=Nid_suivi[i]
    #Ajout d'une colonne loc_date précisant la date et l'heure de chaque relevé GPS dans le résultat de la segmentation
    temp_res_seg$loc_date=ind$loc_date ; temp_res_seg$loc_date=as.character(temp_res_seg$loc_date)
    #Ajout d'une colonne date précisant la date de chaque relevé GPS dans le résultat de la segmentation
    temp_res_seg$cpos_date=ind$cpos_date ; temp_res_seg$cpos_date=as.character(temp_res_seg$cpos_date)
    #Ajout d'une colonne heure précisant l'heure de chaque relevé GPS dans le résultat de la segmentation
    temp_res_seg$cpos_heure=ind$cpos_heure ; temp_res_seg$cpos_heure=as.character(temp_res_seg$cpos_heure)
    temp_res_seg$pos_distance_route=ind$pos_distance_route
    temp_res_seg$pos_distance_bois=ind$pos_distance_bois
    temp_res_seg$pos_distance_bati=ind$pos_distance_bati
    temp_res_seg$pos_distance_haie=ind$pos_distance_haie
    #Ajout des résultats de la segmentation au tableau res_seg (initialement remplie de NA)
    res_seg=rbind(res_seg,temp_res_seg)
    #Création d'une carte donnant le résultat de la segmentation
    segmap(seg, coord.names = c("x","y"))
    #Sauvegarde de la carte
    ggsave(filename=paste(Nid_suivi[i],"_plot_segment.png"), width = 11, height = 8) # ggplot2
    #A chaque tour de la boucle, affichage du n_suivi qui vient d'être segmenté
    print(Nid_suivi[i])
  }

#Suppression de la première ligne de NA dans le tableau de résultats de segmentation
res_seg=res_seg[-1,]

res_seg_sauv=res_seg
```

```
#Conversion de Loc_date au format POSIXct
res_seg$loc_date=as.POSIXct(strptime(res_seg$loc_date,"%Y-%m-%d %H:%M:%S"),tz="GMT")
#Conversion de date au format POSIXct
res_seg$cpos_date=as.POSIXct(strptime(res_seg$cpos_date,"%Y-%m-%d"),tz="GMT")
#Conversion de heure au format times
res_seg$cpos_heure=chron(times. = res_seg$cpos_heure)

#Sauvegarde du tableau de résultats
setwd("C:/Users/smillet/Stage/Def habitat/3- segclust2d/Chevreuils/Cpt")
save(res_seg,file="res_segcpt_chevreuils_10min.RData")

#Sauvegarde du tableau de résultats au format csv
write.csv(res_seg,"C:/Users/smillet/Stage/Def habitat/3- segclust2d/Chevreuils/Cpt/res_segcpt_chevreuils_10min.csv")
```

Sélection des remises

A partir de la segmentation comportementale, chaque segment correspondant potentiellement à une remise a été identifié. Pour chaque segment de remise potentielle, les valeurs de médiane et variance de la longueur du pas ont été relevés puis la moyenne de chacune de ces 2 variables a été calculée. Deux valeurs seuils de remise ont ainsi pu être identifiées : * Valeur seuil de la médiane de la longueur d'un pas : $\mu.step=20$ * Valeur seuil de la variance de la longueur d'un pas : $sd.step=17$

Au delà de ces valeurs seuils, le segment ne peut être considéré comme une remise.

```
#Affichage du tableau de résultats de La segmentation comportementale
res_seg=res_seg[,c("id_suivi",'state','mu.step','sd.step',"loc_date","cpos_date","cpos_heure","x","y","prop","pos_distance_route","pos_distance_bois","pos_distance_haie","pos_distance_bati")]

#Extraction des remises à partir du tableau de résultats de La segmentation comportementale, en appliquant les valeurs seuils
remises=res_seg[(res_seg$mu.step<=20 & res_seg$sd.step<=17),]

#Ordonnance du jeu de données des remises en fonction du collier de l'animal, puis de la date
remises=remises[order(remises$id_suivi,remises$loc_date),]

#Sauvegarde du tableau de résultats des remises
setwd("C:/Users/smillet/Stage/Def habitat/3- segclust2d/Chevreuils/Cpt")
save(remises,file="remises_chevreuils_10min.RData")

#Sauvegarde du tableau de résultats des remises au format csv
write.csv(remises,"C:/Users/smillet/Stage/Def habitat/3- segclust2d/Chevreuils/Cpt/remises_chevreuils_10min.csv")
```

SCRIPT SANGLIER

Pour l'espèce sanglier, la segmentation comportementale a été effectuée pour chaque semaine de suivi et pour chaque individu, avec des données GPS d'une acquisition toutes les heures.

Préparation des données pour la segmentation

```
#Chargement des Librairies nécessaires à l'exécution du script
library(segclust2d)
library(ggplot2)
library(moveHMM)
library(lubridate)
library(chron)

#Accès au dossier où se trouve Les données GPS à utiliser
setwd('C:/Users/smillet/Stage/Données GPS/Materiel GPS_FDC31/Materiel GPS_
SMILLET/Données Vectronic (2019-2020)/shape GPS/shape_données_regulières')

#Lecture des fichiers de données GPS
data40398<-read.csv('collar40398_regul.csv')
data40400_1<-read.csv('collar40400_1_regul.csv')
data40400_2<-read.csv('collar40400_2_regul.csv')
data40401<-read.csv('collar40401_regul.csv')
data40394<-read.csv('collar40394_regul.csv')
data40397<-read.csv('collar40397_regul.csv')

#Regroupement des données GPS issues des différents individus en un jeu de
données unique
data<-rbind(data40398,data40400_1,data40400_2,data40401,data40397,data4039
4)

#Création d'une nouvelle variable Loc_date, issue du collage de la date et
de l'heure des relevés GPS
data$loc_date<-paste(data$UTC_Date,data$UTC_Time)

#Conversion de la date de suivi en format POSIXct
data$date<-data$UTC_Date
data$date<-as.POSIXct(strptime(as.factor(data$date), format="%d/%m/%Y"), fo
rmat="%d-%m-%Y")

#Conversion de l'heure de suivi en format POSIXct
data$heure<-data$UTC_Time
data$heure<-chron(times. = data$heure)

#Conversion de Loc_date en format POSIXct (date/heure)
data$loc_date<-as.POSIXct(strptime(as.factor(data$loc_date), format="%d/%m/
%Y %H:%M:%S"), format="%d-%m-%Y %H:%M:%S")

#Ordonnance du jeu de données en fonction du collier de l'animal (CollarID
), puis de manière chronologique (Loc_date)
data=data[order(data$CollarID,data$loc_date),]
```

```
#Création d'une variable NInd donnant La Liste de tous Les individus du jeu de données data
NInd=levels(as.factor(as.character(data$CollarID)))

#Renommage des variables "Latitude" et "Longitude" respectivement en "x" et "y"
data$x=data$Longitude
data$y=data$Latitude..

#Création d'une nouvelle variable week qui donne pour chaque relevé GPS Le numéro de La semaine
data$week=week(as.POSIXct(strptime(as.factor(data$UTC_Date),format="%d/%m/%Y"),format="%d-%m-%Y"))

#Simplification du jeu de données (-->diminution du temps de traitement)
data<-na.omit(data[,c("CollarID","x","y","loc_date", "date", "heure", "week")])
```

Segmentation comportementale

```
#Création d'une matrice de NA qui se remplira avec Les résultats de La segmentation
res_seg=matrix(NA,1,19,dimnames=list(NA,c('x','y','step','angle','loc_date','CollarID','week','date','heure','x_ind','subsample_ind','indice','state','prop','mu.step','mu.angle','sd.step','sd.angle','state_ordered'))))

#Préparation des données : calcul de La Longueur de pas et de L'angle
processeddata <- prepData(data,coordNames=c("x","y"))
#Préparation des données : suppression des lignes de NA
processeddata=na.omit(processeddata)

#Boucle de segmentation tournant pour chaque individu et pour chaque semaine
for (i in 1:length(NInd)) { #Boucle-individu
  #Création d'un sous-ensemble de données pour chaque individu suivi
  indtot=subset(processeddata,processeddata$CollarID==NInd[i])
  #Création d'un nouveau jeu de données pour chaque individu suivi et suppression des lignes NA
  ind=indtot[,c("step","angle","loc_date","week","date", "heure", "x","y")] ; ind=na.omit(ind)
  #Création d'une nouvelle variable Nweek_ind donnant pour chaque individu La liste de tous Les n° de semaine de suivi
  Nweek_ind=levels(as.factor(as.character(ind$week)))
  for (j in 1:length(Nweek_ind)) { #Boucle-semaine
    #Création d'un sous-ensemble de données pour chaque semaine de suivi
    ind_week=subset(ind,ind$week==Nweek_ind[j])
    #Création d'un nouveau jeu de données pour chaque semaine de suivi
    ind_cpt_week=ind_week[,c("step","angle","loc_date", "date", "heure", "x","y")]
    #Lancement de La segmentation si et seulement si il y a plus de 5 relevés GPS dans La semaine de suivi
```



```

if (nrow(ind_cpt_week)>5) {
  #Segmentation
  seg <- segmentation(ind_cpt_week, lmin = 5, Kmax = 20, seg.var = c("
step","angle"),scale.variable = FALSE)
  #Création d'un graphique donnant la segmentation en fonction de x et
en fonction de y
  plot.segmentation(seg)
  #Sauvegarde du graphique obtenu
  setwd("C:/Users/smillet/Stage/Def habitat/3- segclust2d/Sangliers/Cp
t")
  ggsave(filename=paste(NInd[i],"_week",Nweek_ind[j],"_segclust2d.png"
), width = 11, height = 8) # ggplot2
  #Résultat statistique de la segmentation
  temp_res_seg=augment(seg)
  #Ajout d'une colonne précisant pour chaque relevé GPS le nom de l'in
dividu concerné par la segmentation
  temp_res_seg$CollarID=indtot$CollarID[i]
  #Ajout d'une colonne loc_date précisant la date et l'heure de chaque
relevé GPS dans le résultat de la segmentation
  temp_res_seg$loc_date=ind_cpt_week$loc_date ; temp_res_seg$loc_date=
as.character(temp_res_seg$loc_date)
  #Ajout d'une colonne date précisant la date de chaque relevé GPS dan
s le résultat de la segmentation
  temp_res_seg$date=ind_cpt_week$date ; temp_res_seg$date=as.character
(temp_res_seg$date)
  #Ajout d'une colonne heure précisant l'heure de chaque relevé GPS da
ns le résultat de la segmentation
  temp_res_seg$heure=ind_cpt_week$heure ; temp_res_seg$heure=as.charac
ter(temp_res_seg$heure)
  #Ajout d'une colonne week précisant la semaine de suivi pour chaque
relevé GPS dans le résultat de la segmentation
  temp_res_seg$week=ind_week$week[j]
  #Ajout des résultats de la segmentation au tableau res_seg (initiale
ment remplie de NA)
  res_seg=rbind(res_seg,temp_res_seg)
  #Création d'une carte donnant le résultat de la segmentation
  segmap(seg, coord.names = c("x","y"))
  #Sauvegarde de la carte
  ggsave(filename=paste(NInd[i],"_week",Nweek_ind[j],"_plot_segment.pn
g"), width = 11, height = 8) # ggplot2
  #A chaque tour de la boucle-individu, affichage de l'individu suivi
segmenté
  print(Nweek_ind[j])}
  #A chaque tour de la boucle, affichage de la semaine de suivi segmenté
e
  print(NInd[i])
}}

#Suppression de la première ligne de NA dans le tableau de résultats de se
gmentation
res_seg=res_seg[-1,]

```

```
res_seg_sauv=res_seg

#Conversion de Loc_date au format POSIXct
res_seg$loc_date=as.POSIXct(strptime(res_seg$loc_date,"%Y-%m-%d %H:%M:%S"),tz="GMT")
#Conversion de date au format POSIXct
res_seg$date=as.POSIXct(strptime(res_seg$date,"%Y-%m-%d"),tz="GMT")
#Conversion de heure au format times
res_seg$heure=chron(times. = res_seg$heure)

#Sauvegarde du tableau de résultats
setwd("C:/Users/smillet/Stage/Def habitat/3- segclust2d/Sangliers/Cpt")
save(res_seg,file="res_segcpt_sangliers_week.RData")

#Sauvegarde du tableau de résultats au format csv
write.csv(res_seg,"C:/Users/smillet/Stage/Def habitat/3- segclust2d/Sangliers/Cpt/res_segcpt_sangliers_week.csv")
```

Sélection des remises

A partir de la segmentation comportementale, chaque segment correspondant potentiellement à une remise a été identifié. Pour chaque segment de remise potentielle, les valeurs de médiane et variance de la longueur du pas ont été relevés puis la moyenne de chacune de ces 2 variables a été calculée. Deux valeurs seuils de remise ont ainsi pu être identifiées : * Valeur seuil de la médiane de la longueur d'un pas : $\mu.step=0.05$ * Valeur seuil de la variance de la longueur d'un pas : $sd.step=0.042$

Au-delà de ces valeurs seuils, le segment ne peut être considéré comme une remise.

```
#Affichage du tableau de résultats de La segmentation comportementale
res_seg=res_seg[,c("CollarID", 'state', 'mu.step', 'sd.step', "loc_date", "week", "date", "heure", "x", "y", "prop")]

#Extraction des remises à partir du tableau de résultats de La segmentation comportementale, en appliquant les valeurs seuils
remises_plus=res_seg[(res_seg$mu.step<=0.05 & res_seg$sd.step<=0.042),]
#Ordonnance du jeu de données des remises en fonction du collier de l'animal, puis de la date
remises=remises[order(remises$CollarID,remises$loc_date),]

#Sauvegarde du tableau de résultats des remises
setwd("C:/Users/smillet/Stage/Def habitat/3- segclust2d/Sangliers/Cpt")
save(remises,file="remises_sangliers_week.RData")

#Sauvegarde du tableau de résultats des remises au format csv
write.csv(remises,"C:/Users/smillet/Stage/Def habitat/3- segclust2d/Sangliers/Cpt/remises_sangliers_week.csv")
```

Annexe 3 *Valeurs du quantile à 2% pour les distances des remises-chevreuils (remises dans boisements et remises sur OS agricole) à la route, au bâti et au boisement, en fonction de la méthode d'identification des remises utilisée*

Distance à l'élément paysager ...	Méthode d'identification des points-remises	Q2%
Route	Distance entre 2 localisations successives <20m	14 m
	Segmentation segclust2d	11 m
Bâti	Distance entre 2 localisations successives <20m	46 m
	Segmentation segclust2d	42 m

Annexe 4 *Nouvelle nomenclature des classes d'occupation du sol élaborée lors du stage*

Nom de l'OS dans la nouvelle nomenclature élaborée pendant le stage	Nom de l'OS dans la nomenclature élaborée par Anaïs
Bâtiments	Bâtiments
Zones urbanisées ou perturbées	- Zones urbanisées ou perturbées - Parcelles en travaux
Boisements (hors-urbain)	- Peuplements forestiers et arbustifs hors contexte urbain - Peuplements forestiers et arbustifs aux abords ou dans les emprises des ILT
Boisements urbains	Peuplements forestiers et arbustifs en contexte urbain
Sols nus ou enneigés	Sols nus ou enneigés
Formations herbacées en contexte urbain	Formations herbacées en contexte urbain
Formations herbacées en milieu ouvert	- Formations herbacées indifférenciées en contexte agricole - Formations herbacées aux abords ou dans les emprises des ILT - Cultures annuelles - Cultures pluriannuelles - Prairies permanentes - Prairies temporaires

	- Bandes enherbées des cours d'eau
Surfaces en eau	- Surfaces en eau - Milieux aquatiques temporaires - Milieux aquatiques permanents - Canaux à berges naturelles - Points d'eau
ILT mineures	- Voies ferrées - Routes - Départementales - Autres routes
ILT moyennes	Nationales
ILT majeures	- Canaux à berges artificielles - Autoroutes et quasi-autoroutes
Réseaux de haies	Réseaux de haies

Annexe 5 *Script R - Analyse sensibilité des frictions attribuées à chaque occupation du sol dans le modèle Via Fauna*

Le modèle Via Fauna présente deux paramètres clés : l'habitat (définissant les points de départ et d'arrivée du modèle) et les frictions attribuées à chaque occupation du sol.

Ce document vise à présenter une méthode d'analyse de la sensibilité des frictions attribuées à chaque occupation du sol dans le modèle Via Fauna. Cette analyse a permis d'améliorer les frictions utilisées dans le modèle Via Fauna, en fonction de l'espèce considérée notamment.

Dans ce document, on se base sur une zone d'étude sur laquelle il y a le trajet de dispersion de 9 disperseurs.

CALCUL PREALABLE DES VARIABLES POUR CHAQUE POINTS GPS DES TRAJETS DE DISPERSION

Il est d'abord nécessaire de calculer pour chaque point GPS du trajet de dispersion réel les valeurs de différentes variables, qui pourront ensuite être comparées à celles des trajets de dispersion théoriques prédits par le modèle Via Fauna.

Etapes préliminaires

Avant de lancer le script ci-dessous, les opérations suivantes auront été effectuées au préalable :

1. Identifier les trajets de dispersion se trouvant dans la zone d'étude (à partir de la liste des disperseurs de Laura Benoit)

2. Extraire les données de suivi GPS de la base de données CEFS pour les disperseurs précédemment identifiés
3. Pour chaque disperseur, appliquer les dates de dispersion définies par Laura Benoit pour ne garder que les données GPS correspondant au trajet de dispersion de l'individu.
4. Pour le trajet de dispersion de chaque disperseur, identifier les "clusters" de points, un cluster étant défini comme un ensemble de points distants au maximum de 150 mètres entre eux. Pour cela, sur QGIS : Boîte à outils de traitements/Analyse vectorielle/Partionnement DBSCAN/Taille minimum du cluster : 2 points, Distance maximale entre les points dans le cluster : 150 mètres

A l'issue de ces traitements, on possède pour chaque disperseur, un trajet de dispersion pour lequel chaque point GPS est représenté par un numéro de cluster, ou non si le point n'avait pas de voisin à 150 mètres.

Principe de la méthode

Pour chaque point GPS du trajet de dispersion réel les valeurs de différentes variables sont calculées :

- Distance du point GPS à la route la plus proche (dans un rayon de 5km)
- Distance du point GPS au bâti le plus proche (dans un rayon de 5km)
- Distance du point GPS au boisement le plus proche (dans un rayon de 5km)
- Occupation du sol présente sous le point GPS

Script

```
#Chargement des packages nécessaires
```

```
library(DBI)
```

```
library(RPostgreSQL)
```

```
#Chargement des trajets de dispersion de la zone d'étude
```

```
data510_09<-read.csv("F:/Confrontation trajectoires modèle/CHEVREUILS/Indiv confrontation/510_09/trajet_disp_510_09.csv")
```

```
data971_17<-read.csv("F:/Confrontation trajectoires modèle/CHEVREUILS/Indiv confrontation/971_17/trajet_disp_971_17.csv")
```

```
data3027_18<-read.csv("F:/Confrontation trajectoires modèle/CHEVREUILS/Indiv confrontation/3027_18/trajet_disp_3027_18.csv")
```

```
data3072_19<-read.csv("F:/Confrontation trajectoires modèle/CHEVREUILS/Indiv confrontation/3072_19/trajet_disp_3072_19.csv")
```

```
dataF582_11<-read.csv("F:/Confrontation trajectoires modèle/CHEVREUILS/Indiv confrontation/F582_11/trajet_disp_F582_11.csv")
```

```
dataF732_12<-read.csv("F:/Confrontation trajectoires modèle/CHEVREUILS/Indiv confrontation/F732_12/trajet_disp_F732_12.csv")
```

```
dataF861_14<-read.csv("/Volumes/NO NAME/Confrontation trajectoires modèle/CHEVREUIL/Indiv confrontation/F861_14/trajet_disp_F861_14.csv")
```

```
dataF1410_18<-read.csv("F:/Confrontation trajectoires modèle/CHEVREUILS/Indiv confrontation/F1410_18/trajet_disp_F1410_18.csv")
```

```
#Fusion de tous les trajets de dispersion dans un seul jeu de données
data<-rbind(data510_09,data938_16,data971_17,data3027_18,data3072_19,dataF
582_11,dataF732_12,dataF861_14,dataF1410_18)

#Liste du cap_bague de tous les disperseurs de la zone d'étude
NIndiv=levels(as.factor(as.character(data$cap_bague)))

#Tableau de résultats regroupant tous les disperseurs de la zone d'étude
data_tousind=matrix(NA,1,13,dimnames=list(NA,c('cap_bague','cpos_id','cpos
_date','cpos_heure','pos_x','pos_y',
                                                    'id_cluster', 'os','dist_rou
te','importance','nature',
                                                    'dist_bati','dist_bois'))))

#Connexion à la base de données CEFS
dbcefs<- dbConnect(PostgreSQL(), host="pggeodb.nancy.inra.fr", dbname="db_
cefs", user="....", password="....")

#Boucle affectant à chaque données GPS de chaque individu sa distance à la
route, au bâti, au bois et l'occupation du sol sous le point traité
for (j in 1:length(NIndiv)){ #Boucle-individu

  #Création d'une variable ind pour chaque disperseur traité dans la boucl
e
  ind=subset(data,data$cap_bague==NIndiv[j])
  ind=ind[,c("cpos_id","cap_bague","pos_x","pos_y", "cpos_date", "cpos_heu
re","id_cluster")]

  #Chargement de la variable ind dans une table temporaire de la base de d
onnées CEFS
  dbWriteTable(dbcefs, c("temporaire","millet_gps_25022021"), ind,row.name
s=FALSE)

  #Ajout d'une colonne geom (conversion de pos_x et pos_y de Lambert III à
Lambert 93) à la table
  dbSendQuery(dbcefs,"ALTER TABLE temporaire.millet_gps_25022021 ADD COLUM
N geom geometry")
  dbSendQuery(dbcefs,"UPDATE temporaire.millet_gps_25022021 SET geom = a.g
ege
                        FROM (SELECT cpos_id, ST_transform(ST_SetSRID(ST_MakePoi
nt(pos_x, pos_y),27573),2154) gege
                        FROM temporaire.millet_gps_25022021) a where millet_gps_
25022021.cpos_id = a.cpos_id")

  #Ajout d'une colonne OS à la table et attribution d'une OS à chaque poin
t GPS
  dbSendQuery(dbcefs,"ALTER TABLE temporaire.millet_gps_25022021 ADD COLUM
N OS text")
  dbSendQuery(dbcefs, "UPDATE temporaire.millet_gps_25022021 SET OS = env_
data.sm_os_casvf_ign.cat_sophie
                        FROM env_data.sm_os_casvf_ign WHERE ST_Intersects(temporai
```



```
re.millet_gps_25022021.geom,
      env_data.sm_os_casvf_ign.geom)")

  #Calcul de la distance à la route, au bâti et au bois le plus proche, dans un rayon de 5km
  data_ind<-dbGetQuery(dbcefs,"SELECT cap_bague, cpos_id, cpos_date, cpos_heure, pos_x, pos_y,
      id_cluster, OS, closest_road.dist_route, closest_road.
importance, closest_road.nature,
      closest_bati.dist_bati, closest_bois.dist_bois
FROM temporaire.millet_gps_25022021 as gps

      CROSS JOIN LATERAL
      (SELECT road.importance, road.nature,
      ST_Distance(road.geom, gps.geom) as dist_route
FROM env_data.topo_17_route_emprise_disp as road
WHERE ST_DWithin(road.geom, gps.geom, 5000)
ORDER BY ST_Distance(road.geom, gps.geom)
LIMIT 1
      ) AS closest_road

      CROSS JOIN LATERAL
      (SELECT
      ST_Distance(bati.geom, gps.geom) as dist_bati
FROM env_data.topo_17_bati_emprise_disp as bati
WHERE ST_DWithin(bati.geom, gps.geom, 5000)
ORDER BY ST_Distance(bati.geom, gps.geom)
LIMIT 1
      ) AS closest_bati

      CROSS JOIN LATERAL
      (SELECT
      ST_Distance(bois.geom, gps.geom) as dist_bois
FROM env_data.topo_17_veget_emprise_disp as bois
WHERE ST_DWithin(bois.geom, gps.geom, 5000)
ORDER BY ST_Distance(bois.geom, gps.geom)
LIMIT 1
      ) AS closest_bois

      ")

  #Ajout du tableau de données du disperseur traité dans la boucle au tableau de résultats de tous les disperseurs de la zone d'étude
  data_tousind=rbind(data_tousind,data_ind)

  #Suppression de la table temporaire de la base de données (pour pouvoir traiter le disperseur suivant)
  dbSendQuery(dbcefs,"DROP TABLE temporaire.millet_gps_25022021")

  #Affichage du disperseur traité pour suivre la progression dans l'exécution du script
  print(NIndiv[j])
```

```
} #Fin de La boucle-individu

#Suppression de La première ligne de NA du tableau de résultats
data_tousind=data_tousind[-1,]

#Sauvegarde du tableau de résultats
save(data_tousind,file="data_tousind.RData")
```

NB : Dans ce script, les distances à tout type de routes sont calculées : les petits sentiers de forêts, les chemins de terre, etc. Pour l'analyse de sensibilité conduite ici, il serait sûrement plus judicieux d'ajouter un critère dans la sélection des routes pour le calcul de distance. Par exemple, ajouter : *WHERE (nature='Autoroute' OR nature='Bretelle' OR nature='Quasi-autoroute' OR nature='Route à 1 chaussée' OR nature='Route à 2 chaussées')*

Ici ce critère de sélection des routes n'a pas été appliqué car la zone étudiée est très rural et ne possède que très peu d'ILT. Ainsi, la zone ne possède que rarement des routes d'une certaine importance (de type non-chemin) à moins de rayon 5km. Certains points GPS se seraient alors retrouvés sans calcul de distance à la route. On aurait pu bien sûr pû augmenter le rayon de recherche des routes à 10km ou 20km mais cela aurait augmenté considérablement le temps de calcul.

STATISTIQUES SUR LES TRAJETS DE DISPERSION REELS - CHEMINS GPS

Principe de la méthode

Maintenant que l'occupation du sol, les distances à la route, au bâti et au bois de chaque point GPS du trajet réel de dispersion ont été calculées, 100 échantillons de trajets GPS vont être générés afin de gommer l'effet des amas de points dans un cluster. En effet, les points d'un même cluster possèdent un certain poids dans le calcul de la moyenne des différentes variables sur l'ensemble du trajet GPS et peuvent ainsi contribuer à biaiser l'analyse de sensibilité.

Ce biais peut être supprimé en générant pour chaque disperseur, 100 échantillons de trajet GPS, où chaque échantillon correspond à un ensemble de points sélectionnés aléatoirement et de manière unique dans chaque cluster (cluster représenté par son id_cluster grâce à l'opération de Partitionnement DBSCAN effectuée précédemment sur QGis). Grâce à cette méthode, il n'y aura jamais 2 points appartenant au même cluster dans un échantillon de trajet GPS et le biais lié aux amas de point sera donc supprimé.

Dans cette méthode, pour chaque échantillon, les moyennes des variables sont calculées. Les moyennes des variables du trajet de dispersion réel d'un individu sont ainsi obtenues en calculant la moyenne des moyennes des 100 échantillons.

Script

```
#Tableau de résultats final
result_final_GPS_tous_indiv=matrix(NA,1,8,dimnames=list(NA,c('cap_bague',
'nb_pts_trajet_GPS','moy_GPS_dist_route',
'moy_GPS_dist_bati', 'moy_GPS_dist_bois_MO','moy_pourcent_MO','moy_pourcent_bois','moy_
```

```
pourcent_haie'))))

for (j in 1:length(NIndiv)) { #Boucle-individu
  #Création d'une variable ind pour chaque disperseur traité dans la boucle
  ind=subset(data_tousind,data_tousind$cap_bague==NIndiv[j])
  ind=ind[,c("cpos_id","cap_bague", "cpos_date", "cpos_heure", "pos_x","pos_y",
  "id_cluster", "os","dist_route","importance","nature",
  "dist_bati","dist_bois")]

  #Liste de tous les id_cluster de l'individu traité dans la boucle
  Ncluster=levels(as.factor(as.character(ind$id_cluster)))

  #Tableau de résultats de tous les points du trajet GPS pour chacun des 100 échantillons
  tous_ech_ind=matrix(NA,1,20,dimnames=list(NA,c('cpos_id','cap_bague','cpos_date','cpos_heure','pos_x',
  'pos_y','n_ech','id_cluster','os','dist_route','importance','nature',
  'moy_dist_route','dist_bati','moy_dist_bati','dist_bois',
  'moy_dist_bois_MO','pourcent_MO',
  'pourcent_bois','pourcent_haie'))))

  #Tableau de résultats avec les moyennes des variables du trajet GPS pour chacun des 100 échantillons
  resume_tous_ech_ind=matrix(NA,1,7,dimnames=list(NA,c('n_ech','moy_dist_route','moy_dist_bati',
  'moy_dist_bois_MO',
  'pourcent_bois',
  'pourcent_MO','pourcent_haie'))))

  #Boucle générant 100 échantillons du trajet GPS de l'individu traité dans la boucle
  for (i in 1:100){
    ech=ind[is.na(ind$id_cluster),] #Sélection des points GPS non-inclus dans un cluster
    ech$n_ech=i #Attribution d'un numéro d'échantillon

    #Boucle pour sélectionner aléatoirement un point dans chaque cluster de points
    for (k in 1:length(Ncluster)){
      cluster=subset(ind,ind$id_cluster==Ncluster[k])
      pt_cluster=cluster[sample(1:nrow(cluster),1,replace=FALSE),]
      pt_cluster$n_ech=i
      ech=rbind(ech,pt_cluster)
    }
  }
}
```

```
ech$moy_dist_route=mean(ech$dist_route) #Calcul de La distance moyenne
à la route pour L'échantillon de trajet GPS considéré
ech$moy_dist_bati=mean(ech$dist_bati) #Calcul de La distance moyenne a
u bâti pour L'échantillon de trajet GPS considéré

#Création de compteurs comptant Le nombre de points GPS sur chaque occ
upation du sol
compteur_MO=0
compteur_bois=0
compteur_haie=0
for (l in 1:nrow(ech)) {
  if ((ech$os[l]=='Formations herbacees en milieu ouvert') & (!is.na(e
ch$os[l]))) {
    compteur_MO=compteur_MO+1}
  if ((ech$os[l]=='Boisements (hors-urbain)') & (!is.na(ech$os[l]))) {
    compteur_bois=compteur_bois+1}
  if ((ech$os[l]=='Reseaux de haies') & (!is.na(ech$os[l]))) {
    compteur_haie=compteur_haie+1}
}

#Calcul de La distance moyenne au bois des points sur une OS MO pour L
'échantillon de trajet GPS considéré
ech_MO=ech[(ech$os=='Formations herbacees en milieu ouvert') & (!is.na
(ech$os)),]
ech$moy_dist_bois_MO=mean(ech_MO$dist_bois)

#Calcul du pourcentage de points sur chaque OS pour L'échantillon de t
rajat GPS considéré
ech$pourcent_MO=(compteur_MO/nrow(ech))*100
ech$pourcent_bois=(compteur_bois/nrow(ech))*100
ech$pourcent_haie=(compteur_haie/nrow(ech))*100

ech=ech[,c('cpos_id', 'cap_bague', 'cpos_date', 'cpos_heure', 'pos_x',
'pos_y', 'n_ech', 'id_cluster', 'os', 'dist_route', 'importance', 'nature'
,
'moy_dist_route', 'dist_bati', 'moy_dist_bati', 'dist_bois',
'moy_dist_bois_MO', 'pourcent_MO',
'pourcent_bois', 'pourcent_haie')]

tous_ech_ind=rbind(tous_ech_ind,ech)

resume_ech<-data.frame(n_ech=i, moy_dist_route=ech$moy_dist_route,
moy_dist_bati=ech$moy_dist_bati,moy_dist_bois_M
O=ech$moy_dist_bois_MO,
pourcent_bois=ech$pourcent_bois,pourcent_MO=ech
$pourcent_MO,pourcent_haie=ech$pourcent_haie)

resume_tous_ech_ind=rbind(resume_tous_ech_ind, resume_ech)

print(i) #Affichage du numéro de L'échantillon pour suivre la progress
ion dans L'exécution du script
```

```
}  
  
#Suppression de La première Ligne de NA  
tous_ech_ind=tous_ech_ind[-1,]  
resume_tous_ech_ind=resume_tous_ech_ind[-1,]  
  
result_finalGPS=data.frame(cap_bague=NIndiv[j], nb_pts_trajet_GPS=nrow(e  
ch),  
                                moy_GPS_dist_route=mean(resume_tous_ech_ind$m  
oy_dist_route),  
                                moy_GPS_dist_bati=mean(resume_tous_ech_ind$m  
oy_dist_bati),  
                                moy_GPS_dist_bois_MO=mean(resume_tous_ech_ind  
$moy_dist_bois_MO),  
                                moy_pourcent_MO=mean(resume_tous_ech_ind$po  
urcent_MO),  
                                moy_pourcent_bois=mean(resume_tous_ech_ind$po  
urcent_bois),  
                                moy_pourcent_haie=mean(resume_tous_ech_ind$po  
urcent_haie))  
  
result_final_GPS_tous_indiv=rbind(result_final_GPS_tous_indiv,result_fin  
alGPS)  
  
print(NIndiv[j]) #Affichage de L'individu traité pour suivre la progress  
ion dans L'exécution du script  
}  
  
#Suppression de La première Ligne de NA  
result_final_GPS_tous_indiv=result_final_GPS_tous_indiv[-1,]  
  
#Sauvegarde du résultat  
setwd("G:/Confrontation trajectoires modèle/CHEVREUILS/ANALYSE SENSIBILITE  
/TRAJET GPS REELS")  
save(result_final_GPS_tous_indiv,file="result_final_GPS_tous_indiv.RData")
```

STATISTIQUES SUR LES TRAJETS DE DISPERSION THEORIQUES - CHEMINS DE MOINDRE COUT DU MODELE VIA FAUNA

Etapas préliminaires

Pour chaque cas de friction considéré, et pour chaque disperseur de la zone d'étude :

1. Isoler en 2 shapes distincts les points de départ et d'arrivée du trajet de dispersion réel de chaque disperseur de la zone d'étude.
2. Lancer le modèle Via Fauna entre ces 2 points et enregistrer le chemin de moindre coût (CMC) généré.
- 3 Sur QGis : Boite à outils de traitements/Géométrie vectorielle/Points le long d'une géométrie/Distance : 1m, décalage de départ et de fin : 0m

4 Sur Qgis : Boite à outils de traitements/Table vecteur/Ajouter les champs X/Y à la couche/SCR= 27573-Lambert III/Préfixe de champ : pos_

5 Enregistrer au format csv le trajet CMC avec points tous les 1m.

Principe de la méthode

Cette méthode diffère de ce qui a été fait précédemment lors de l'élaboration des échantillons de trajet de dispersion. Ici, on part du CMC tracé par le modèle Via Fauna, avec un point tous les mètres. Afin de comparer le CMC au trajet réel GPS, on sélectionne aléatoirement dans le CMC le même nombre de point que le trajet GPS : c'est un échantillon. On répète l'opération 100 fois, on obtient donc 100 échantillons de trajets CMC.

De la même manière que précédemment, on calcule la moyenne des variables pour chaque échantillon, puis on fait la moyenne des 100 échantillons pour obtenir la moyenne de l'individu traité.

```
library(DBI)
library(RPostgreSQL)

#Connexion à La base de données CEFS
dbcefs<- dbConnect(PostgreSQL(), host="host_adress", dbname="db_name", use
r="xxxx", password="xxxx")

#Listant donnant Les CMC pour chaque disperseur de La zone d'étude, dans L
e Cas de friction du modèle Via Fauna
list_CasVF<-list.files(path="G:/Confrontation trajectoires modèle/CHEVREUI
LS/ANALYSE SENSIBILITE/CMC_tous_indiv_tous_cas", pattern = "CasVF", full.n
ame=TRUE)

for (m in 1:length(list_CasVF)){ #Boucle-individu

  #Lecture du fichier CMC correspondant à L'individu traité dans La boucle
  data_ind<-read.csv(list_CasVF[m])

  #Sélection des variables précédemment calculées sur Le trajet GPS de L'i
ndividu traité dans La boucle
  GPS=subset(result_final_GPS_tous_indiv,result_final_GPS_tous_indiv$cap_b
ague==data_ind$cap_bague[1])

#Tableau de résultats final
tous_ech_ind=matrix(NA,1,16,dimnames=list(NA,c('id_ligne','pos_x', 'pos_y'
,'n_ech','os','dist_route',
'importance','nature',
'moy_dist_route','dist_bati
','moy_dist_bati', 'dist_bois',
'moy_dist_bois_MO','pourcen
t_MO',
'pourcent_bois','pourcent_h
aie'))))
```



```
#Tableau de résultats avec Les moyennes pour chaque échantillon d'un individu
resume_tous_ech_ind=matrix(NA,1,7,dimnames=list(NA,c('n_ech','moy_dist_route','moy_dist_bati',
                                                    'moy_dist_bois_MO','pourcent_bois',
                                                    'pourcent_MO','pourcentage_haie'))))

#Boucle générant 100 échantillons du trajet CMC de l'individu traité dans la boucle
for (i in 1:100){
  #Sélection aléatoirement dans Le CMC du même nombre de point que le trajet et GPS de l'individu correspondant
  ech=data_ind[sample(1:nrow(data_ind), GPS$nb_pts_trajet_GPS,replace=FALSE),]
  ech$id_ligne<-seq(1,nrow(ech),1)
  ech$n_ech=i
  ech=ech[,c('id_ligne','n_ech','pos_x','pos_y')]

  #Chargement de la variable ech dans une table temporaire de la base de données CEFS
  dbWriteTable(dbcefs, c("temporaire","millet_gps_25022021"), ech,row.names=FALSE)

  #Ajout d'une colonne geom (conversion de pos_x et pos_y de Lambert III à Lambert 93) à la table
  dbSendQuery(dbcefs,"ALTER TABLE temporaire.millet_gps_25022021 ADD COLUMN geom geometry")
  dbSendQuery(dbcefs,"UPDATE temporaire.millet_gps_25022021 SET geom = a.gege
                    FROM (SELECT id_ligne, ST_transform(ST_SetSRID(ST_MakePoint(pos_x, pos_y),27573),2154)  gege
                    FROM temporaire.millet_gps_25022021) a where millet_gps_25022021.id_ligne = a.id_ligne")

  #Ajout d'une colonne OS à la table et attribution d'une OS à chaque point GPS
  dbSendQuery(dbcefs,"ALTER TABLE temporaire.millet_gps_25022021 ADD COLUMN OS text")
  dbSendQuery(dbcefs, "UPDATE temporaire.millet_gps_25022021 SET OS = env_data.sm_os.cat_sophie
                    FROM env_data.sm_os WHERE ST_Intersects(temporaire.millet_gps_25022021.geom,
                    env_data.sm_os.geom)")

  #Calcul de la distance à la route, au bâti et au bois le plus proche, dans un rayon de 5km
  ech<-dbGetQuery(dbcefs,"SELECT id_ligne, n_ech, pos_x, pos_y, geom, OS,
                    closest_road.dist_route, closest_road.importance, clos
```

```
est_road.nature,
    closest_bati.dist_bati, closest_bois.dist_bois
FROM temporaire.millet_gps_25022021 as gps

CROSS JOIN LATERAL
(SELECT road.importance, road.nature,
ST_Distance(road.geom, gps.geom) as dist_route
FROM env_data.topo_17_route_emprise_disp as road
WHERE ST_DWithin(road.geom, gps.geom, 5000)
ORDER BY ST_Distance(road.geom, gps.geom)
LIMIT 1
) AS closest_road

CROSS JOIN LATERAL
(SELECT
ST_Distance(bati.geom, gps.geom) as dist_bati
FROM env_data.topo_17_bati_emprise_disp as bati
WHERE ST_DWithin(bati.geom, gps.geom, 5000)
ORDER BY ST_Distance(bati.geom, gps.geom)
LIMIT 1
) AS closest_bati

CROSS JOIN LATERAL
(SELECT
ST_Distance(bois.geom, gps.geom) as dist_bois
FROM env_data.topo_17_veget_emprise_disp as bois
WHERE ST_DWithin(bois.geom, gps.geom, 5000)
ORDER BY ST_Distance(bois.geom, gps.geom)
LIMIT 1
) AS closest_bois

")

ech$moy_dist_route=mean(ech$dist_route) #Calcul de la distance moyenne à
La route pour L'échantillon de trajet CMC considéré
ech$moy_dist_bati=mean(ech$dist_bati)#Calcul de la distance moyenne au b
âti pour L'échantillon de trajet CMC considéré

#Création de compteurs comptant Le nombre de points sur chaque occupatio
n du sol
compteur_MO=0
compteur_bois=0
compteur_haie=0
for (l in 1:nrow(ech)) {
  if ((ech$os[l]=='Formations herbacees en milieu ouvert') & (!is.na(ech
$os[l]))) {
    compteur_MO=compteur_MO+1}
  if ((ech$os[l]=='Boisements (hors-urbain)') & (!is.na(ech$os[l]))) {
    compteur_bois=compteur_bois+1}
  if ((ech$os[l]=='Reseaux de haies') & (!is.na(ech$os[l]))) {
    compteur_haie=compteur_haie+1}
}
```

```
#Calcul de La distance moyenne au bois des points sur une OS MO pour L'échantillon de trajet CMC considéré
ech_MO=ech[(ech$os=='Formations herbacees en milieu ouvert') & (!is.na(ech$os)),]
ech$moy_dist_bois_MO=mean(ech_MO$dist_bois)

#Calcul du pourcentage de points sur chaque OS pour L'échantillon de trajet GPS considéré
ech$pourcent_MO=(compteur_MO/nrow(ech))*100
ech$pourcent_bois=(compteur_bois/nrow(ech))*100
ech$pourcent_haie=(compteur_MO/nrow(ech))*100

ech=ech[,c('id_ligne','pos_x','pos_y','n_ech','os','dist_route','importance','nature',
           'moy_dist_route','dist_bati','moy_dist_bati','dist_bois',
           'moy_dist_bois_MO','pourcent_MO',
           'pourcent_bois','pourcent_haie')]

tous_ech_ind=rbind(tous_ech_ind,ech)

resume_ech<-data.frame(n_ech=i, moy_dist_route=mean(ech$dist_route),
                      moy_dist_bati=mean(ech$dist_bati),moy_dist_bois_MO=mean(ech_MO$dist_bois),
                      pourcent_bois=(compteur_bois/nrow(ech))*100,
                      pourcent_MO=(compteur_MO/nrow(ech))*100,
                      pourcent_haie=(compteur_haie/nrow(ech))*100)

resume_tous_ech_ind=rbind(resume_tous_ech_ind,resume_ech)

dbSendQuery(dbcefs,"DROP TABLE temporaire.millet_gps_25022021")

print(i) #Affichage du numéro de L'échantillon traité pour suivre la progression dans L'exécution du script
}

#Suppression de La première ligne de NA
resume_tous_ech_ind=resume_tous_ech_ind[-1,]


#Sauvegarde des données
setwd("G:/Confrontation trajectoires modèle/CHEVREUILS/ANALYSE SENSIBILITE/CAS TAMPON/CMC_tous_indiv_tous_cas/results")
save(resume_tous_ech_ind,file=paste("resume_tous_ech_CMC", substr(list_CasVF[m],115,130),".RData"))


print(list_CasVF[m]) #Affichage de L'individu traité pour suivre la progression dans L'exécution du script
}
```

Annexe 6 *Guide méthodologique pour réaliser une carte d'occupation du sol intégrant les tampons répulsifs autour des routes et attractifs autour des boisements*

1. Sur la couche shape de la carte d'occupation du sol, ne sélectionner que les occupations du sol ayant pour nom « Boisements (hors-urbain) » dans la nouvelle nomenclature élaborée lors du stage (cf Annexe 4).
2. Appliquer un tampon de 30 mètres autour de ces boisements (30 mètres correspond à la distance moyenne au bois des points sur OS MO pour les 9 individus de la zone d'étude).
3. Sur la couche shape des routes recensées par l'IGN, ne sélectionner que les routes de type « Autoroutes », « Nationales » et « Départementales ».
4. Appliquer un tampon de 84 mètres autour de de ces routes (84 mètres correspond distance moyenne au bois des points sur OS MO pour les 9 individus de la zone d'étude).
5. Fusionner les 2 couches tampons et toutes les couches « vm » dans l'ordre suivant (de haut en bas) pour réaliser une nouvelle carte d'OS qui intègre les tampons.

Type d'élément	Ordre de fusion des couches
Éléments naturels et agricoles	Zones enneigées /Sols nus
	Formation herbacées indifférenciées contexte agricole (cultures ou prairies)
	Cultures
	Prairies temporaires
	Prairies permanentes
Éléments de corridors	Peuplement forestiers et arbustifs hors contexte urbain (Probablement non clôturés)
	Haies
	Bandes enherbées
	Cours d'eau temporaires
	Cours d'eau permanents
	Points d'eau (FRC)
Éléments imperméabilisants	Surfaces en eau
	Peuplement forestiers et arbustifs en contexte urbain (Probablement clôturés)
	Peuplements forestiers et arbustifs aux abords ou dans les emprises des ILT
	Formations herbacées aux abords ou dans les emprises des ILT
	Parcelles anthropisées ou perturbées (en travaux)
	Formations herbacées en contexte urbain (Probablement clôturés/ jardins)
	Zones urbanisées ou perturbées (parkings, zones d'activités, carrières, décharges...)
	Bâtiments
	Voies Ferrées
	Routes + autres routes à 1 chaussée
	Départementales 1 chaussée
	Départementales 2 chaussées
	Nationales 1 chaussée
Nationales 2 chaussées	
Autoroutes + Bretelles + Quasi-autoroutes	
Couches des éléments imperméabilisants (grillages, murs...)	
Éléments perméabilisants	Couches des éléments perméabilisants (Buses, ponts...)

 Tampon 30m forêts

 Tampon 84m routes

6. Dans le fichier Excel attribuant une friction à chaque occupation du sol, ajouter 2 nouvelles lignes permettant d'attribuer une friction aux éléments des 2 couches tampons.



7. Réaliser une jointure entre la nouvelle carte d'occupation du sol et nomenclature attribuant une friction à chaque OS. Champ de jointure : code_cs_us.