

Développement d'une application pour l'exploration et l'analyse de données GPS en télémétrie animale

Auteur : Watthez, Quentin

Promoteur(s) : 11781; Lejeune, Philippe

Faculté : Gembloux Agro-Bio Tech (GxABT)

Diplôme : Master en bioingénieur : gestion des forêts et des espaces naturels, à finalité spécialisée

Année académique : 2021-2022

URI/URL : <http://hdl.handle.net/2268.2/15260>

Avertissement à l'attention des usagers :

Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative" (BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.

DÉVELOPPEMENT D'UNE APPLICATION POUR L'EXPLORATION ET L'ANALYSE DE DONNÉES GPS EN TÉLÉMÉTRIE ANIMALE

QUENTIN WATTHEZ

**TRAVAIL DE FIN D'ÉTUDES PRÉSENTÉ EN VUE DE L'OBTENTION DU DIPLÔME DE
MASTER BIOINGÉNIEUR EN GESTION DES FORêTS ET DES ESPACES NATURELS**

ANNÉE ACADEMIQUE 2021-2022

CO-PROMOTEUR(S): ALAIN LICOPPE, PHILIPPE LEJEUNE

Toute reproduction du présent document, par quelque procédé que ce soit, ne peut être réalisée qu'avec l'autorisation de l'auteur et de l'autorité académique de Gembloux Agro-Bio Tech.

Le présent document n'engage que son auteur.

DÉVELOPPEMENT D'UNE APPLICATION POUR L'EXPLORATION ET L'ANALYSE DE DONNÉES GPS EN TÉLÉMÉTRIE ANIMALE

QUENTIN WATTHEZ

**TRAVAIL DE FIN D'ÉTUDES PRÉSENTÉ EN VUE DE L'OBTENTION DU DIPLÔME DE
MASTER BIOINGÉNIEUR EN GESTION DES FORêTS ET DES ESPACES NATURELS**

ANNÉE ACADEMIQUE 2021-2022

CO-PROMOTEUR(S): ALAIN LICOPPE, PHILIPPE LEJEUNE

Remerciements

Tout d'abord, je tiens à remercier Alain Licoppe et Philippe Lejeune, mes co-promoteurs, pour avoir accueilli cette idée de projet avec enthousiasme et m'avoir donné la chance de le développer en toute autonomie. Leurs conseils avisés m'auront permis de progresser dans l'exploration de cette fascinante thématique de l'étude du déplacement des animaux.

Ensuite, je remercie les personnes ayant participé à l'enquête sur les cartes interactives pour leurs précieux retours. J'adresse des remerciements particuliers à Gérard Jadoul et Valentin Claes du bureau d'étude EcoFirst pour la mise à disposition de précieuses données GPS m'ayant permis valider l'outil. Mes remerciements vont également aux agents DNF du cantonnement de Nassogne et à mes nombreux collègues du DEMNA.

J'exprime aussi ma reconnaissance envers les membres du jury pour avoir accepté de donner leur temps et leur expertise afin d'évaluer la qualité de ce travail: Hélène Soyeurt, Marc Dufrêne et Arnaud Monty. Je tiens aussi à remercier Hugues Claessens pour sa disponibilité et ses conseils qui m'ont orienté tout au long de mon parcours à Gembloux Agro-Bio Tech.

Enfin, cette aventure n'aurait jamais pu aboutir sans le soutien inconditionnel de Pauline, ma femme, et de nos deux enfants, Florentin et Raphaël. Merci pour votre présence et votre patience au cours de ces trois années bien denses.

Résumé

Au cours des dernières décennies, le développement et la démocratisation de la technologie GPS ont multiplié les possibilités d'études et de suivis de la faune sauvage. Cette progression s'est accompagnée d'une taille croissante des jeux de données, complexifiant ainsi la tâche de traitement et d'analyse. Ce phénomène s'est traduit par le développement de solutions logicielles pour aider les scientifiques dans l'exploitation de ces données.

Ce travail de fin d'études consiste à développer un nouvel outil informatique permettant d'automatiser la production de cartes interactives et l'analyse de données de télémétrie animale. Ce projet poursuit un double objectif: d'une part, créer une solution pour communiquer efficacement les déplacements des animaux équipés par colliers GPS. D'autre part simplifier significativement l'analyse en intégrant de nombreux traitements sur le jeu de données initial.

Dénommée « TracksMining », l'application se veut innovante pour son intégration globale de l'ensemble des traitements en un processus d'exécution unique et performant. Implémenté avec Python, le logiciel peut être utilisé sans disposer de compétences spécifiques en programmation ou en traitement SIG. La première version de l'application présentée dans ce rapport a démontré toute son utilité dans un contexte professionnel et validé ses résultats et performances. Destinée à être améliorée au fur et à mesure de son utilisation au sein de la cellule Faune du DEMNA, l'application sera ultérieurement mise à disposition de la communauté scientifique et du grand public.

Abstract

In recent decades, the rise of GPS technology has multiplied the possibilities for studying and monitoring wildlife. This advent was followed by an increase in databases sizes, complicating the task of processing and analysis data. This phenomenon has led to the development of software solutions to help mining this massive amount of information.

This work aims to develop an innovative computer application with a dual objective. First, it deals with the creation of a new communication tool to display the moves of tagged animals through an interactive map. Secondly, it automates the analysis of animal telemetry data by integrating many processing steps on the initial dataset.

Called "TracksMining", this software stands out with a global integration of all stages in a fast and unique execution flow. Implemented with Python, the software innovates with its ease of use as no specific skills in programming or GIS are required. As presented in this report, the first version of this computer program has proved to be useful and efficient in a professional context and validated its performances. At a first instance, this software is intended to be used and improved within the DEMNA. However, TracksMining should be made later available for the scientific community and the general public.

Table des matières

REMERCIEMENTS	i
RÉSUMÉ/ABSTRACT.....	ii
TABLE DES MATIÈRES	iii
LISTE DES FIGURES	iv
LISTE DES TABLEAUX.....	v
LISTE DES ANNEXES	vi
ABRÉVIATIONS	vii
I. INTRODUCTION	1
1.1. CONTEXTE	1
1.2. PROBLÉMATIQUE INITIALE.....	2
II. ETAT DE L'ART	3
2.1. INTRODUCTION.....	3
2.2. L'AVÈNEMENT DE LA TECHNOLOGIE GPS EN TÉLÉMÉTRIE ANIMALE.....	3
2.2.1. <i>Une technologie aux nombreux avantages</i>	3
2.2.2. <i>Des applications variées</i>	3
2.2.3. <i>De nouveaux challenges</i>	5
2.2.4. <i>La multiplication des solutions logicielles</i>	6
2.3. EXPLOITATION ET ANALYSE DE DONNÉES : QUELS BESOINS DE LA COMMUNAUTÉ?.....	7
2.3.1. <i>Pré-traitement des données</i>	7
2.3.2. <i>Estimation du domaine vital</i>	8
2.3.3. <i>Visualisation de données</i>	12
2.3.4. <i>Analyses graphiques</i>	13
III. MÉTHODES	18
3.1. DÉFINITION DU CONCEPT GÉNÉRAL.....	18
3.2. PROTOTYPAGE.....	18
3.3. DÉFINITION DES CRITÈRES D'ÉVALUATION	19
3.4. DÉVELOPPEMENT DE L'APPLICATION.....	20
3.4.1. <i>Introduction</i>	20
3.4.2. <i>Choix du langage de programmation</i>	20
3.4.3. <i>Syntaxe</i>	20
3.4.4. <i>Utilisation de modules externes</i>	21
3.4.5. <i>Architecture logicielle</i>	22
3.4.6. <i>Logique générale</i>	22
3.4.7. <i>Classes de traitement</i>	23
3.4.8. <i>Fichiers produits</i>	23

IV. RÉSULTATS	24
4.1. FONCTIONNALITÉS DE L'APPLICATION	24
4.1.1. Accès à l'interface utilisateur	24
4.1.2. Chargement des données	24
4.1.3. Scan du fichier	25
4.1.4. Configuration des paramètres	26
4.1.5. Options et paramètres	26
4.1.6. Téléchargement des résultats	28
4.2. CARTE INTERACTIVE ET FONCTIONNALITÉS	29
4.2.1. Ouverture du document et vue générale	29
4.2.2. Navigation entre les couches	29
4.2.3. Description des couches de bases	30
4.3. ANALYSES GRAPHIQUES	35
4.3.1. Distance parcourue	35
4.3.2. Distance à un point de référence	35
4.3.3. Evolution de la superficie du domaine vital	35
4.3.4. Utilisation de l'habitat	36
V. DISCUSSION.....	37
5.1. EVALUATION DE L'ATTEINTE DES OBJECTIFS.....	37
5.1.1. Opérationnalité dans un contexte professionnel	37
5.1.2. Intégration des traitements élémentaires	37
5.1.3. Appréciation du logiciel par des utilisateurs potentiels	37
5.1.4. Validation des résultats	39
5.1.5. Validation des performances	40
5.1.6. Extrapolation à plusieurs modèles biologiques	42
5.1.7. Structuration cohérente et logique	43
5.1.8. Mise à disposition d'une interface utilisateur	43
5.2. QUELLE PLUS-VALUE POUR TRACKSMINING ?.....	44
5.2.1. Minimiser les dépendances	44
5.2.2. Mettre à disposition publique une solution complémentaire	45
5.2.3. Proposer une approche globale	45
5.2.4. Créer un logiciel accessible	45
5.2.5. Favoriser la communication	45
VI. CONCLUSIONS ET PERSPECTIVES.....	46
6.1. CONCLUSIONS	46
6.2. PERSPECTIVES.....	47
6.2.1. Permettre le suivi simultané de plusieurs individus	47
6.2.2. Adapter le programme aux données d'observations	47
6.2.3. Elargir les possibilités d'analyse	47
6.2.4. Améliorer la détection automatique des localisations erronées	48
6.2.5. Implémenter de nouveaux estimateurs de l'habitat	48
6.2.6. Augmenter les performances générales	48
VII. BIBLIOGRAPHIE.....	49

Liste des figures

FIGURE 1: PRÉSENTATION DE DONNÉES SOUS FORME D'UNE CARTE STATIQUE	1
FIGURE 2: DOMAINE VITAL DE PLUSIEURS INDIVIDUS PAR MCP	8
FIGURE 3: SUPERPOSITION DE DOMAINES VITAUX PAR CHP	9
FIGURE 4 : PRÉSENTATION D'UN KERNEL DE DENSITÉ AVEC LES ZONES NOYAUX EN ROUGE.....	9
FIGURE 5 : PRÉSENTATION DE L'UTILISATION DE L'HABITAT PAR GRID CELL	10
FIGURE 6 : PRÉSENTATION DE L'UTILISATION DE L'HABITAT PAR LoCoH	10
FIGURE 7: PRÉSENTATION DIRECTE DES GÉOLOCALISATIONS.....	12
FIGURE 8 : ISOPLÈTHES D'UN KERNEL DE DENSITÉ.....	12
FIGURE 9 : SUPERPOSITION DES ESTIMATIONS D'UN DOMAINE VITAL PAR MCP AVEC UN GRID CELL ET UN KERNEL DE DENSITÉ	12
FIGURE 10 : TRAJECTOIRES D'UN OISEAU MIGRATEUR ET DISPERSION DE RENARDS	13
FIGURE 11: VITESSE MOYENNE SELON LA PÉRIODE DU JOUR	14
FIGURE 12: PRÉSENTATION DE LA DISTANCE MENSUELLE TOTALE ET MOYENNE QUOTIDIENNE PARCOURUE	14
FIGURE 13: PRÉSENTATION DU NSD	14
FIGURE 14 : EVOLUTION DE LA TAILLE DU DOMAINE VITAL EN FONCTION DU NOMBRE DE GÉOLOCALISATIONS	15
FIGURE 15: EVOLUTION DU DOMAINE VITAL HEBDOMADAIRE D'UN MILAN ROYAL	15
FIGURE 16 : PROJECTION DES POINTS SUR UN SHAPEFILE REPRENANT LES DIFFÉRENTS TYPES D'HABITATS	16
FIGURE 17 : COMPARAISON DE LA PROPORTION D'HABITATS DANS LE DOMAINE VITAL	16
FIGURE 18 : COMPARAISON PAR HABITAT DES SUPERFICIES DISPONIBLES ET UTILISÉES DANS LE DOMAINE VITAL D'UN PÉCARI	17
FIGURE 19 : FRÉQUENCES CUMULÉES DES LOCALISATIONS EN FONCTION DE LA DISTANCE AU NID	17
FIGURE 20 : VUE D'ENSEMBLE DU PROCESSUS	23
FIGURE 21 : PAGE D'ACCUEIL ET EXPLORATEUR DE FICHIER.....	24
FIGURE 22 : FORMULAIRE DES PARAMÈTRES	26
FIGURE 23 : SIGNALEMENT VISUEL D'UN OUTLIER POTENTIEL.....	26
FIGURE 24 : DE GAUCHE À DROITE : OPENSTREETMAP, CARTODBPOSITRON, STAMENTERRAIN	27
FIGURE 25 : VUE DES PRINCIPALES OPTIONS DE PERSONNALISATION DU STYLE.....	28
FIGURE 26 : TÉLÉCHARGEMENT DU DOSSIER COMPRESSÉ.....	28
FIGURE 27 : VUE GÉNÉRALE DE LA CARTE INTERACTIVE.....	29
FIGURE 28 : PANNEAU DE NAVIGATION.....	29
FIGURE 29 : TOOLTIP AU SURVOL (À GAUCHE) ET POPUP AU CLIC DE SOURIS (À DROITE)	30
FIGURE 30 : LES TRAJECTOIRES À LARGE ÉCHELLE (GAUCHE) ET À PETITE ÉCHELLE (DROITE)	30
FIGURE 31 : TOOLTIP AU SURVOL (À GAUCHE) ET POPUP AU CLIC DE SOURIS (À DROITE)	30
FIGURE 32 : CENTROÏDE DES GÉOLOCALISATIONS.....	31
FIGURE 33 : ENVELOPPE TOTALE DES GÉOLOCALISATIONS	31
FIGURE 34 : CARTE DE CHALEUR	31
FIGURE 35 : MCP 95%.....	32
FIGURE 36 : KDE AVEC ISOPLÈTHES 95% (ORANGE) ET 50 (ROUGE)	32
FIGURE 37 : K-LoCoH AVEC ISOPLÈTHE DE 90%.....	32
FIGURE 38 : GRID CELL AVEC MAILLES DE 100M	32
FIGURE 39 : INTERACTION D'UN RATON-LAVEUR AVEC UN PARCELLAIRE FORESTIER	33
FIGURE 40 : HISTOGRAMME D'INTERACTION MENSUELLE DISPONIBLE EN POPUP POUR CHAQUE ENTITÉ	33
FIGURE 41 : ANIMATION TEMPORELLE, FAISANT DÉFILER LES LOCALISATIONS PAR GROUPES DE 24H	34
FIGURE 42 : OUTIL DE DESSIN, UTILISÉ POUR CRÉER UN MCP SUR LES LOCALISATIONS D'UNE MÊME JOURNÉE.....	34
FIGURE 43: DISTANCE MENSUELLE PARCOURUE AVEC DÉCLINAISON SELON LA PÉRIODE DU JOUR.....	35
FIGURE 45: COMPARAISON DE LA SUPERFICIE D'UN MCP ET D'UN KDE POUR DIFFÉRENTES VALEURS	35
FIGURE 46 : GRAPHIQUE DE DISTRIBUTION DES LOCALISATIONS AU SEIN DES HABITATS	36
FIGURE 47 : CONFRONTATION HABITAT DISPONIBLE / UTILISÉ AU SEIN DU DOMAINE VITAL	36
FIGURE 48 : KERNEL DE DENSITÉ PAR ADEHABITAT (GAUCHE) ET TRACKSMINING (DROITE).....	39
FIGURE 49 : K-LoCoH PAR ADEHABITAT (GAUCHE) ET TRACKSMINING (DROITE).....	40
FIGURE 50 : APERÇU DE LA NOUVELLE INTERFACE UTILISATEUR EN COURS DE DÉVELOPPEMENT	43
FIGURE 51 : R-LoCoH (r=15KM) APPLIQUÉ AUX OBSERVATIONS PONCTUELLES DE LA MARTRE DES PINS ENTRE 2001 ET 2021	47

Liste des tableaux

TABLE 1 : DÉFINITION DES CRITÈRES PERMETTANT D'ÉVALUER L'ATTEINTE DE L'OBJECTIF.....	19
TABLE 2 : RÉSUMÉ DES MODULES EXTERNES UTILISÉS PAR TRACKSMINING	22
TABLE 3 : RÉCAPITULATIF DES COLONNES ESSENTIELLES À UN FICHIER VALIDE	24
TABLE 4 : FORMATS DE DATES RECOMMANDÉS (EXEMPLE DU 18 AVRIL 2021 À 15H53).....	25
TABLE 5 : RÉCAPITULATIF DU CONTENU DES SOUS-DOSSIERS.....	28
TABLE 6 : PERFORMANCES ABSOLUES DE L'APPLICATION (SECONDES)	40
TABLE 7 : CARACTÉRISTIQUES DE L'ORDINATEUR UTILISÉ POUR LE CALCUL DES PERFORMANCES.....	41
TABLE 8 : COMPARAISON DE LA VITESSE D'EXÉCUTION MÉDIANE (SECONDES) POUR 3 ESTIMATEURS	41

Liste des annexes

ANNEXE 1 : LIEN VERS LE DÉPÔT PUBLIC GitHub CONTENANT LE CODE SOURCE ET LES VIDÉOS EXPLICATIVES.....	61
ANNEXE 2 : FORMULAIRE D'ÉVALUATION DE L'ENQUÊTE SUR LES CARTES INTERACTIVES	61

Abréviations

API : Application Programming Interface

BSD : Berkeley Software Distribution License

CHP : Characteristic Hull Polygon

CRS : Coordinate Reference System

DEMNA : Département de l'Etude du Milieu Naturel et Agricole

DNE : Direction de la Nature et de l'Eau

DNF : Département Nature et Forêts

GCM : Grid Cell Method

GPS: Global Positioning System

KDE: Kernel Density Estimator

LoCoH : Local Convex Hull.

MCP : Minimum Convex Polygon

MVC : Modèle-Vue-Contrôleur.

NNCH : Nearest Neighbour Convex Hull

NSD : Net Square Displacement

POO : Programmation Orientée Objet

PEP-8 : Python Extension Proposal

SIG : Système d'information géographique

SPW : Service Public de Wallonie

I. Introduction

1.1. Contexte

En février 2021, parallèlement à la poursuite de mes études de Master bioingénieur à Gembloux Agro-Bio Tech, j'ai eu l'opportunité d'intégrer la cellule « Faune sauvage » du Département de l'étude du milieu naturel et agricole (DEMNA). Cet organisme est un département du Service Public de Wallonie dont la mission générale consiste à récolter des données sur l'état de l'environnement en Wallonie, tant sur la faune et la flore que sur les habitats. Le DEMNA est sous-divisé en plusieurs directions dont fait partie la direction de la Nature et de l'Eau (DNE). Celle-ci regroupe elle-même plusieurs cellules, chacune spécialisée dans des missions spécifiques.

Au sein de cette direction, les missions de la cellule sont essentiellement axées sur le suivi des populations de mammifères en Wallonie. Initialement appelée « Laboratoire de Faune Sauvage et Cynégétique », la cellule était centrée sur le suivi du grand gibier mais ses activités de recherche se sont progressivement étendues à l'étude des mammifères protégés et invasifs. Actuellement, mon emploi se focalise essentiellement sur le suivi des mammifères faisant l'objet d'un statut de protection Natura 2000. Au moyen de différents types d'inventaires, il consiste principalement à mieux connaître l'état des populations sur notre territoire.

Grâce au travail réalisé depuis des années par les agents du DEMNA, les agents du DNF ainsi qu'une multitude d'observateurs et naturalistes bénévoles, nous disposons d'importantes bases de données biologiques liées à des inventaires de terrain ou à des suivis plus spécifiques. C'est dans l'optique de mettre en valeur ces données que j'ai entrepris d'initier le développement d'un nouvel outil pour faciliter leur exploration et leur interprétation.

Ce projet professionnel s'intégrait pleinement au contenu de ma formation. C'est dans ce cadre que j'ai eu la chance de pouvoir développer un travail de fin d'études permettant de combiner deux thématiques qui m'intéressent particulièrement : la télémétrie animale et la programmation. Ce document couvre l'ensemble du processus de développement du logiciel, depuis l'identification de la problématique initiale jusqu'au produit final. Les aspects fonctionnels de l'outil y seront particulièrement développés: avantages, limites et perspectives d'évolution.

1.2. Problématique initiale

Les missions de la cellule Faune comportent occasionnellement des suivis d'individus au moyen de techniques de télémétrie, principalement par l'usage de colliers GPS. Ces études ont généralement pour objectif de mieux cerner les comportements des espèces étudiées, tant pour répondre à des questions de biologie, que pour dégager des solutions de gestion appropriées.

Ces projets amènent régulièrement les membres de la cellule à travailler en étroite collaboration avec d'autres structures, en particulier avec les agents gestionnaires du Département de la Nature et des Forêts (DNF). La transparence des données de télémétrie est un facteur stimulant la recherche et le partenariat entre différents collaborateurs (Dwyer et al., 2015). Une bonne communication est donc primordiale pour garder la cohésion et la motivation autour de ces projets de suivis d'animaux par GPS. Cette régularité d'envoi de données, bien que nécessaire, peut cependant constituer un travail répétitif et chronophage.

D'un point de vue pratique, les données de télémétrie animale ont une composante temporelle qui peut les rendre difficiles à interpréter lorsque présentées sous format statique tel qu'une carte pdf (Figure 1). Ce manque de lisibilité augmente avec le nombre de géolocalisations et nécessite alors l'emploi de légendes complexes (gradients de couleurs, de tailles, etc.) ou la multiplication des supports via des cartes thématiques.

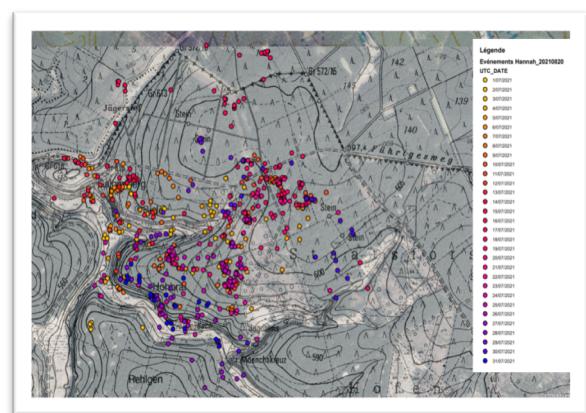


Figure 1: Représentation de données sous forme de carte statique

Enfin, ces suivis génèrent des quantités importantes de données qu'il convient de pouvoir exploiter pour en tirer un maximum d'informations. Ce travail d'analyse nécessite des logiciels adaptés et des compétences spécifiques (Dwyer et al., 2015 ; Urbano et al., 2015) et doit s'opérer selon une certaine standardisation pour que les résultats puissent se comparer entre études et individus (Laver & Kelly, 2008; Peris et al, 2020). Ces difficultés seront détaillées au chapitre 2.2.3.

C'est dans ce contexte qu'a alors émergé l'idée de créer un logiciel spécifique. Ce nouvel outil, baptisé « TracksMining » vise différents objectifs:

- Augmenter l'efficacité du travail en automatisant une tâche répétitive.
- Améliorer la qualité de la communication entre les différents partenaires en produisant des cartes interactives, permettant d'appréhender des données dynamiques et complexes.
- Combiner l'exploration à l'analyse de données en faisant abstraction de compétences en programmation et en traitement SIG grâce à une interface facile et intuitive.
- Standardiser les traitements, tout en permettant une transparence complète sur les paramètres et les algorithmes.

II. Etat de l'art

2.1. Introduction

Cette synthèse bibliographique se décline en 3 parties :

Premièrement, elle vise à mettre en évidence l'importance croissante de la technologie GPS dans le domaine de la télémétrie animale au cours des dernières années, ses multiples applications et les défis techniques qui accompagnent la gestion de ce type de données. Par la suite, quelques-unes des applications informatiques communément utilisées pour le traitement des données de télémétrie seront brièvement présentées.

Ensuite, une revue bibliographique identifie les techniques de pré-traitement, d'exploration visuelle, d'analyse et de représentation graphique des données de télémétrie animale dans la littérature scientifique. Cette partie a pour objectif de définir les fonctionnalités prioritaires et pertinentes à implémenter dans un logiciel.

Finalement, l'intérêt de la création d'un nouveau logiciel dans le cadre de ce travail de fin d'études sera argumenté.

2.2. L'avènement de la technologie GPS en télémétrie animale

2.2.1. Une technologie aux nombreux avantages

Au cours des dernières décennies, les travaux de recherche sur la faune sauvage ont largement profité de la démocratisation de la technologie satellitaire GPS (Marchand et al., 2015; Hofman et al., 2017). Ces récepteurs intégrés à des balises fixées sur les animaux permettent d'enregistrer avec précisions des localisations géographiques sur base d'une fréquence d'acquisition programmable. Les suivis peuvent désormais se réaliser en toutes saisons et en toutes conditions, sur des terrains ou à des périodes pour lesquels l'acquisition de données étaient auparavant très difficiles (Marchand et al., 2015).

L'intérêt majeur des systèmes GPS est la grande facilité de mise en œuvre accompagnée du développement de produits spécifiques à des animaux de plus en plus petits (Klein et al., 2006). Malgré des imprécisions qui subsistent dans certaines localisations, particulièrement en présence d'un couvert végétal dense, cette technologie permet de réaliser des suivis impossibles à réaliser avec des émetteurs VHF (Klein et al., 2006).

2.2.2. Des applications variées

De tels suivis permettent de récolter de nombreuses variables: utilisation et sélection des habitats par les animaux, caractéristiques de leurs mouvements (distances parcourues, vitesses, angles de rotation entre deux déplacements et sinuosité des déplacements...). (Marchand et al., 2015).

Dès lors, les données récoltées lors des suivis biologiques par GPS peuvent couvrir des finalités diverses. Quelques exemples sont donnés ci-dessous:

- Comprendre les réponses spatiales des cerfs vis-à vis d'activités humaines, telles que la chasse (Chassagneux et al. Faune sauvage octobre-novembre 2020) ou le tourisme (Licoppe et al., 2005).
- Etudier l'impact d'éléments anthropiques sur le comportement et le domaine vital des animaux, comme une décharge (González-Gallina et al., 2018) ou des chemins de randonnées (Duparc et al., 2017).
- Analyser les comportements dans l'optique d'une résolution des conflits entre l'homme et la faune sauvage (Douglas-Hamilton et al., 2005 ; Henderson et al., 2018).
- Optimiser la sélection des endroits les plus pertinents pour aménager des ouvrages de franchissement de la faune (Bastille-Rousseau et al., 2018).
- Evaluer l'utilisation générale de l'habitat par des espèces exotiques envahissantes telles que le raton-laveur (Meulemans, 2021) ou le python birman (Hart et al., 2015).
- Etudier la dispersion de graines de fruits forestiers par les animaux (Kays et al., 2011 ; Oleksy et al., 2015 ; Saïd et al. 2019)
- Analyser les interactions entre proies et prédateurs (Mosnier et al., 2008 ; Eriksen et al., 2009).
- Etudier la dispersion des espèces (Stoner et al., 2008 ; Prévot & Licoppe, 2013 ; Walton et al., 2018).
- Evaluer la dynamique de maladies épizootiques en fonction des facteurs environnementaux (Dellicour et al., 2020) ou la corrélation entre l'abondance de parasites et la fréquentation de l'habitat par son hôte (Qviller et al., 2016).
- Etudier l'utilisation de différents habitats en fonction des périodes de la journée (Salvatori et al., 2022) ou de facteurs environnementaux (Cornélis et al., 2011).
- Utiliser les informations de déplacements en tant que données de prédiction météorologiques (Treep et al., 2016).
- Evaluer le succès de réintroduction et de translocation d'espèces animales (Flanagan et al., 2016 ; Yiu et al. 2016).
- Faciliter l'éradication de certaines espèces exotiques envahissantes en marquant des individus pour localiser les sites de rencontre avec leurs conspécifiques (Aschehoug et al., 2004 ; Woolnough et al., 2006 ; Smith et al., 2016).

Certains auteurs mettent en évidence la nécessité de réaliser des études plus quantitatives du mouvement des espèces pour en évaluer les capacités de dispersion (ex : Morelle et al., 2014). D'autres soulignent l'importance de continuer à mieux étudier les conséquences des structures anthropiques dans le but d'apporter tous les éléments écologiques nécessaires pour l'aménagement des territoires (Marchand et al., 2015).

Il apparaît évident que les suivis GPS ont encore un grand potentiel à déployer pour apporter des informations sur l'écologie des espèces et sur les outils à mettre en œuvre pour la gestion de leurs habitats (Marchand et al., 2015). Certains colliers peuvent également être accompagnés de différents types de capteurs (accéléromètre, capteur de température, etc.) permettant de compléter les données de géolocalisation.

2.2.3. De nouveaux challenges

2.2.3.1. *Une complexité croissante*

La facilité de mise en place et la démocratisation de la télémétrie GPS ont significativement complexifié le travail relatif à la gestion des données. Afin de pouvoir être traités efficacement, ces grands jeux de données nécessitent désormais des compétences spécifiques en programmation (Dwyer et al., 2015). La gestion de base de données de tailles conséquentes devient ainsi une compétence essentielle pour les écologistes, au même titre que les statistiques et le traitement SIG (Urbano et al., 2015).

Au sein de la communauté scientifique, la formation à la programmation et au traitement de bases de données complexes n'est pas toujours considéré comme une priorité dans le cadre des recherches en écologie (Urbano et al., 2015). Ceci s'explique probablement à cause d'une certaine réticence à modifier des habitudes de travail, mais aussi à cause des investissements initiaux que peuvent nécessiter la mise en place de nouvelles solutions ou la formation à de nouvelles compétences (Urbano et al., 2015). Plusieurs techniques d'analyse avancées de données de télémétrie animale nécessitent des compétences spécifiques ou des ordinateurs disposant d'une grande puissance de calcul (Cagnacci et al., 2010).

2.2.3.2. *Une standardisation nécessaire des analyses*

L'analyse de données de télémétrie animale fait également face à une difficulté mise en évidence par plusieurs auteurs, particulièrement lorsqu'elle se concentre sur l'estimation des domaines vitaux. Suite à l'examen de 161 publications sur le sujet, Laver & Kelly (2008) avaient remarqué une grande disparité dans la manière de quantifier le domaine vital d'un animal, ce qui engendrait des problèmes de reproductibilité des données et de validité dans les comparaisons entre études.

Quelques années auparavant, certains auteurs (Lawson & Rodgers, 1997) mettaient déjà en évidence ce phénomène et insistaient sur la nécessité que l'estimation d'un domaine vital soit non-seulement accompagnée de la méthode utilisée, mais également de l'ensemble des paramètres ayant servi à le calculer.

L'harmonisation des analyses du mouvement des animaux est un facteur essentiel pour la comparaison des études entre espèces et lieux. (Dwyer et al., 2015). Le développement de méthodes standardisées est impératif afin de réaliser des comparaisons d'études qui ont du sens (Pebsworth et al., 2012).

2.2.3.3. *La nécessité de nouveaux outils*

Face à ces difficultés de traitement et de standardisation, plusieurs auteurs (Coyne & Godley, 2005 ; Urbano et al., 2015) ont souligné l'importance de déployer des solutions informatiques en ligne pour faciliter l'analyse des données de télémétrie animale. L'apparence conviviale de ces logiciels est un facteur important, simplifiant fortement l'usage de la technologie GPS (Klein et al., 2006). Toutefois, l'absence de logiciels conviviaux peut constituer un frein au traitement de grands jeux de données de télémétrie (Michelot et al., 2016).

C'est pourquoi la recherche scientifique doit se focaliser sur la création de logiciels alternatifs et « user friendly » permettant de réaliser des traitements et des estimations de domaines vitaux sur des bases de données de taille conséquente (>10.000 géolocalisations) (Walter et al., 2011). Ces solutions informatiques ne devraient pas se limiter à la gestion et au traitement de données mais également à la visualisation et au partage des informations (Dwyer et al., 2015).

2.2.4. La multiplication des solutions logicielles

Au début des années 90, de nombreux logiciels étaient déjà disponibles pour l'analyse du domaine vital des animaux. Cependant, la nécessité de mettre en œuvre des programmes plus complets et plus modulables (Larkin & Halkin, 1994) ou aux interfaces moins compliquées (Rodgers et al., 2007) se faisait déjà ressentir. Par la suite, de multiples solutions informatiques furent développées pour archiver, traiter et partager les données de télémétrie animale (Rodgers et al., 2007) avec par exemple: « *CALHOME* » (Kie et al., 1996), « *KERNELHR* » (Seaman et al., 1998), « *OpenJump HoRAE* » (Steiniger & Hunter, 2012), « *Hawth's tool* » (Beyer, 2004), etc.

Parmi ces applications, le logiciel « *Animal Movement Analyst Extension* » (Hooge et al., 2001) permet de combiner l'utilisation d'une application de traitement SIG à ArcView© en intégrant une cinquantaine de fonctions destinées à l'exploration du mouvement et de l'habitat des animaux. Elle sera suivie en 2007 par l'extension logicielle « *Home range tools* » (Rodgers et al., 2007) qui permet aux utilisateurs de fournir les options offertes par les programmes antérieurs pour faire varier les différents paramètres.

Le logiciel STAT (Coyne & Godley, 2005) constitue un programme pionnier dans l'analyse en ligne de données de télémétrie. Initialement prévue pour travailler avec le système de géolocalisations ARGOS, cette application gratuite permet la récupération de données, leur traitement et la mise en lien avec de nombreuses couches environnementales.

Le package ADEhabitat pour R (Calenge, 2006) permet quant à lui, l'exploration et l'analyse des interactions entre un animal et son environnement ainsi que le calcul d'estimateurs du domaine vital. Cet outil est encore aujourd'hui largement utilisé par la communauté scientifique (ex : Enriquez-Mercado et al., 2018 ; Reading et al., 2019 ; Abner, 2020 ; Acácio et al., 2021 ; Salvatori et al., 2022).

Plus récemment, le package pour R « *rhr* » (Signer & Balkenhol, 2015) fut développé pour faciliter le calcul des domaines vitaux des animaux suivis par télémétrie GPS. Répondant ainsi aux recommandations effectuées sur la nécessité de pouvoir reproduire les résultats des estimations (Lawson & Rodgers, 1997 ; Laver & Kelly, 2008), ce package génère un rapport comprenant l'ensemble des paramètres utilisés pour le calcul des domaines vitaux. Il permet la mise en œuvre d'une petite interface faisant abstraction de compétences spécifiques en programmation.

Ces dernières années, les logiciels pour la télémétrie animale ont continué à se multiplier, en évoluant davantage vers des solutions en ligne. De manière non-exhaustive, il est possible de

citer : « *MoveBank* » (Kranstauber et al., 2011), « *Game online* » (Jarolímek et al., 2014), « *UvA Virtual Lab* » (<https://www.uva-bits.nl/>), « *Wildlife Tracker* » (<https://www.gis4-wildlife.com/>)

Parmi ces récentes solutions, « *ZoaTrack* » (Hunter et al., 2013) fournit une infrastructure qui soutient la gestion, le pré-traitement, l'analyse et la visualisation des données (Hunter et al. 2013). Initialement baptisé « *OzTrack* », le logiciel se focalise sur la région Australasienne et permet la possibilité de superposer de nombreuses couches thématiques. Dans son fonctionnement, l'application se base essentiellement sur le package R ADE habitat.

2.3. Exploitation et analyse de données : quels besoins de la communauté?

2.3.1. Pré-traitement des données

2.3.1.1. Gestion des « outliers »

L'élimination des localisations erronées ou suspectes est un processus courant en traitement de données GPS (ex : Ngene et al., 2016 ; Meulemans, 2021; Acácio et al., 2021). Ces « *outliers* » peuvent se détecter visuellement ou sur base de vitesses considérées anormales entre deux points successifs (ex : Acácio et al., 2021) ou encore en utilisant la valeur DOP (« *Dilution of Precision* ») (ex : Schofield et al., 2007 ; Teng et al., 2021). Plusieurs méthodes plus élaborées (ex : Bjørneraa et al., 2010 ; Hirakawa et al., 2022) ont été développées pour identifier ces localisations anormales dans les jeux de données de télémétrie animale.

2.3.1.2. Rééchantillonnage

Le rééchantillonnage des données est un traitement régulièrement utilisé afin de disposer de localisations régulièrement espacées dans le temps, permettant d'éviter un biais dans les analyses. Il peut se faire par sous-échantillonnage (ex : Meulemans, 2021 ; Schofield et al., 2007; Athreya et al., 2014) ou par sur-échantillonnage au moyen de l'interpolation de points manquants (Lyver et al., 2011). D'autres techniques plus avancées peuvent encore être utilisées : interpolation par Random Walk Model (Bergman & al., 2000) ou au moyen de modèles de Markov (Zhang et al., 2019).

2.3.1.3. Sélection temporelle de données

Certaines plages de dates peuvent être volontairement supprimées, notamment pour ne pas biaiser le suivi avec un comportement anormal de stress lié à la capture et au marquage des individus (ex : Athreya et al., 2014 ; Gupte et al., 2022). Ce filtre temporel peut également s'appliquer sur des plages horaires définies (ex : Gupte et al., 2022 ; Zengeya et al., 2011).

2.3.1.4. Sélection spatiale de données

Il est parfois souhaitable de filtrer les données se situant à l'intérieur ou à l'extérieur d'une zone géographique particulière, soit parce que les données concernées ne sont pas intéressantes pour la problématique ou parce qu'elles peuvent être erronées ou biaisées (ex : Gupte et al., 2022).

2.3.2. Estimation du domaine vital

2.3.2.1. *Introduction*

Le domaine vital d'un animal peut être défini comme « la zone traversée par un individu au cours de ses activités normales telles que la recherche de nourriture, la reproduction, l'élevage des jeunes, etc. » (Burt 1943, cité par Powell 2000). La définition de ce concept a toutefois fait l'objet d'une évolution au cours des dernières décennies, allant de pair avec les méthodes permettant de procéder à son estimation (Fieberg & Kochanny, 2005). Il existe plusieurs méthodes d'analyse du domaine vital d'un animal mais il n'y a pas de consensus au sein de la communauté scientifique sur la méthode la plus adaptée (Laver & Kelly, 2008).

2.3.2.2. *Polygones minimum convexes*

Parmi les estimateurs du domaine vital, les polygones minimums convexes (MCP) sont une méthode largement acceptée (Burgman & Fox, 2003) et sans doute la technique plus répandue (Laver, 2005). La méthode consiste à utiliser les points les plus extérieurs d'un jeu de données de localisations et de les connecter pour former un polygone sans angle concave (Kie et al., 1996)(Figure 2).

La simplicité d'utilisation des MCP constitue un avantage (Burgman & Fox, 2003). Malgré leur popularité, ils sont parfois qualifiés de mauvais estimateurs car ils ont tendance à surestimer le domaine vital d'un individu. En effet, par l'inclusion de grandes zones qui ne sont pas utilisées ou utilisables (Pimley et al., 2005 ; Seidel et al., 2018), ils sont sujets à des biais peu prévisibles (Börger et al., 2006). Les MCP sont particulièrement instables car ils varient de manière importante en fonction du nombre et de la distribution des points (Downs & Horner, 2008). En revanche, ils sont plutôt robustes à l'autocorrélation induite par des intervalles rapprochés (Fieberg, 2007, cité par Olsen 2011).

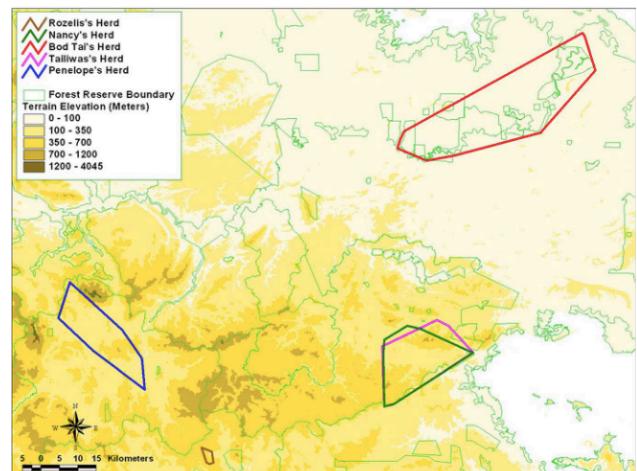
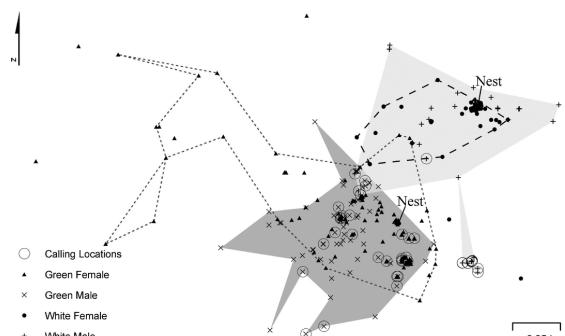


Figure 2: Domaine vital de plusieurs individus par MCP (Alfred et al., 2012)

Malgré ces critiques, les MCP restent un estimateur du domaine vital très fréquemment rencontré dans la littérature scientifique, particulièrement avec un isoplète de 95% permettant d'écartier les géolocalisations les plus éloignées du centroïde (ex : Frone et al., 2009 ; Recio et al., 2010 ; Pebsworth et al., 2012 ; Athreya et al., 2014 ; González-Gallina et al., 2018 ; Fisher et al., 2020).

2.3.2.3. Characteristic Hull Polygon (CHP)



Les Characteristic Hull Polygon (CHP) sont une variante des MCP (Downs & Horner, 2009). Contrairement à ces derniers, ils peuvent avoir des bords concaves (Figure 3), être disjoints ou présenter des zones de vides. Ils sont construits sur base d'une triangulation de Delaunay, à partir de laquelle ne sont conservés qu'un certain pourcentage des plus petits triangles. Ils semblent bien adaptés pour l'étude des animaux qui occupent de manière intensive des emplacements assez éloignés (Downs & Horner, 2009).

2.3.2.4. Kernels de densité

L'analyse du domaine vital au moyen des kernels de densité (KDE) est une méthode statistique non-paramétrique qui permet l'estimation de la densité de probabilité sur base d'un ensemble de points (Rodgers et al., 2007). Cette méthode est utilisée depuis les années 1990 pour l'analyse des domaines vitaux (Worton, 1989, cité par Rodgers et al., 2007). De manière vulgarisée, cet estimateur du domaine vital permet de quantifier la probabilité de localiser un individu à un endroit déterminé (Rodgers et al., 2007). L'utilisation de différents seuils de probabilité permet d'identifier des zones exploitées de manière plus intensive, souvent appelées « zones noyaux » (Figure 4).

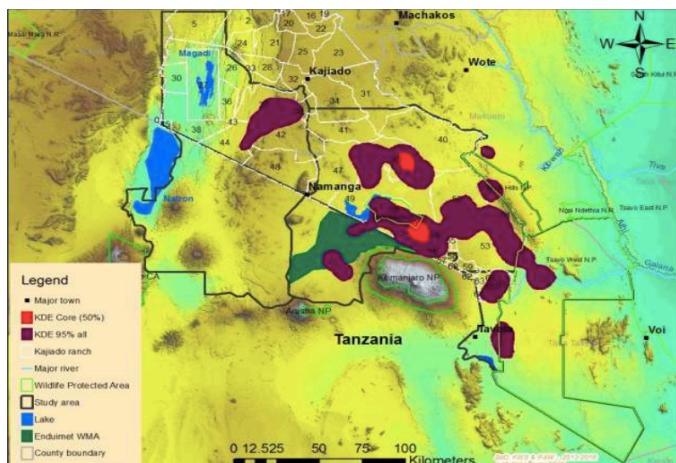


Figure 4 : Représentation d'un kernel de densité avec les zones noyaux en rouge (Ngene et al., 2016)

Les kernels de densité ont fait l'objet de critiques car ils ne convergent pas vers l'utilisation réelle de l'habitat avec l'augmentation du nombre de données. Lorsque ces derniers sont utilisés sur des animaux utilisant de vastes territoires, ils peuvent inclure des zones qui ne sont pas réellement utilisées. Dans certains cas, ces zones peuvent même être identifiées à tort comme noyaux (Getz & Wilmers, 2004).

Toutefois, certains auteurs ont proposé des adaptations permettant d'augmenter significativement la pertinence biologique des kernels de densité. Ces améliorations visent généralement à intégrer des données sur les mouvements et l'activité des individus (Benhamou & Cornelis, 2010) ou à identifier des zones exploitées de manière plus intensives (Benhamou & Riotte-Lambert, 2012). Les kernels de densité (Figure 4) n'en restent pas moins un estimateur de densité encore largement utilisé (ex : Beauchamp et al., 2018 ; Pillay et al., 2018, Zhang et al., 2019 ; Rodriguez et al., 2021 ; Salvatori et al., 2022).

2.3.2.5. Grid Cell Method

Cette méthode vise à superposer un quadrillage sur la zone d'étude et à définir le domaine vital comme l'ensemble des mailles qui recoupent la position d'un individu (Horner & Powell, 1990, cité par Powell, 2000). Chaque maille peut obtenir un score défini par le nombre de localisations qu'elle intègre. Il en résulte une surface composée d'un assemblage de carrés permettant d'évaluer l'utilisation de l'habitat par l'animal et d'identifier des zones de plus forte fréquentation (Figure 5).

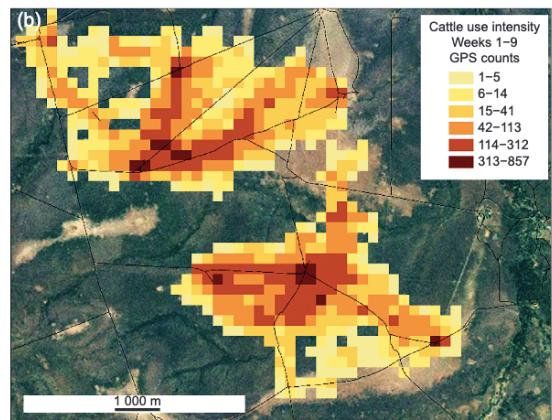


Figure 5 : Représentation de l'utilisation de l'habitat par Grid Cell (Schieltz et al., 2017)

La résolution de la grille est déterminante car la taille des cellules va influencer fortement l'interprétation des résultats (Stark et al., 2017). La maille se détermine généralement en fonction de la surface de la zone d'étude ou des comportements de l'animal. (Lloyd, 1967 ; Vandermeer, 1981, cité par Powell, 2000). Si la résolution des mailles est trop fine, l'estimation peut aboutir à un domaine vital comprenant beaucoup de surfaces disjointes (Powell, 2000). L'avantage de la méthode est de permettre une analyse détaillée des zones fortement utilisées et de réaliser une approche complémentaire intéressante à un autre estimateur (Stark et al., 2017). Par exemple, si le MCP donne de meilleures estimations sur les domaines mensuels et saisonniers chez une espèce de primate, le Grid Cell permet cependant de réaliser une estimation plus fine à l'échelle annuelle (Grueter et al., 2009 ; Eraud et al., 2019).

2.3.2.6. Local Convex Hull

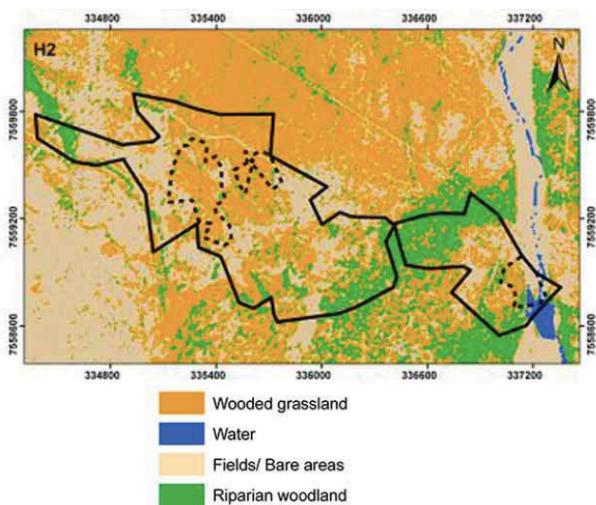


Figure 6 : Représentation de l'utilisation de l'habitat par LoCoH (Zengetya et al., 2011)

Proposée par Getz & Wilmers (2004), la méthode du K-NNCH (K-Nearest Neighbour Convex Hull, rebaptisée plus tard K-LoCoH) est considérée comme une extension du MCP classique avec l'ajout d'un paramètre supplémentaire. La méthode du K-LoCoH considère l'union progressive des MCP réalisés autour de chaque point du jeu de données et de ses $k-1$ voisins. Cette méthode, permet d'obtenir un ou plusieurs polygones aux bords concaves avec inclusion de zones creuses (Figure 6). Le LoCoH converge vers l'utilisation réelle du territoire avec l'augmentation du nombre de données et permet l'identification de zones barrières (Getz & Wilmers, 2004).

La variante du R-LoCoh (Getz et al., 2007) suit un principe identique mais pour chaque localisation, réalise un MCP sur base des points inclus dans un cercle de rayon déterminé. Une autre variante, le A-LoCoH (Getz et al., 2007) inclut les localisations jusqu'à ce que la somme cumulée des distances entre le point de référence et les points intégrés soit inférieure à la valeur a . Cette dernière variante semble donner de meilleurs résultats (Getz & Wilmers, 2004). Une autre variante, le T-LoCoH (Lyons et al., 2013), permet de prendre en compte la composante temporelle dans l'utilisation de l'habitat. Ceci permet de modéliser la variation spatio-temporelle saisonnière de l'utilisation de l'habitat, les réponses de la faune par rapport aux modifications de l'habitat (Cross et al., 2020).

Utilisée dans plusieurs études (ex : Loveridge et al., 2009 ; Castellanos, 2011 ; Scull et al., 2012), cette méthode du LoCoH (Figure 6) semble particulièrement performante dans les environnements où surviennent des barrières naturelles et anthropiques (Scull et al., 2012) car elle met en évidence les lieux potentiellement évités ou non-accessibles aux individus (Lichti et Swihart, 2011).

2.3.2.7. *L'intérêt de croiser plusieurs estimateurs*

Une estimation correcte de l'habitat et des ressources utilisées par un animal est importante à des fins de gestion et de conservation (Benhamou & Cornelis, 2010). Le choix d'un estimateur repose avant tout sur des considérations écologiques (Cumming & Cornélis, 2012). Pour une même espèce, il est intéressant d'utiliser plusieurs estimations du domaine vital car la méthode la plus appropriée est en lien direct avec le comportement de l'animal, mais aussi avec le nombre de données collectées (Boyle et al., 2008) et l'environnement (Scull et al., 2012).

Pour un même estimateur, les paramètres utilisés peuvent faire varier considérablement le résultat et mener à des interprétations biologiques différentes (Seider et al. 2018). Il est donc essentiel de garder à l'esprit que tous les estimateurs du domaine vital ne sont qu'une approximation de la réalité (Scull et al., 2012). Certains estimateurs semblent peu appropriées pour certains taxons spécifiques. Par exemple, les kernels de densité paraissent peu adaptés pour l'étude de l'herpétofaune (Row & Blouin-Demers, 2008) tandis que d'autres auteurs suggèrent que les MCP ne sont pas indiqués pour l'étude des primates (Grueter et al., 2009).

Au sein d'une même étude, il est donc fréquent de considérer la superficie de plusieurs estimateurs du domaine vital, afin de comparer leurs superficies relatives ou communes (ex : Rodriguez et al., 2021). Le KDE et le MCP sont les estimateurs du domaine vital les plus largement utilisés (Otten et al., 2021) et souvent utilisés de manière conjointe (ex : Kauhala & Auttila, 2010 ; Enriquez-Mercado et al., 2018 ; Ross et al., 2019 ; Vieira et al., 2019 ; Otten et al., 2021 ; Rodriguez et al., 2021).

2.3.3. Visualisation de données

La projection des géolocalisations des points GPS sur un fond de carte est une pratique courante pour visualiser l'information spatiale. Il est alors possible de distinguer les localisations en fonction de la période du jour par une couleur différente (Klein et al., 2006), une icône différente (Athreya et al. ; 2014) ou selon les individus (ex : Meulemans, 2021 ; Dürr & Ward, 2014)(Figure 7).

Cette représentation des localisations étant généralement peu informative en raison du grand nombre de points, il est plus fréquent de retrouver les surfaces correspondantes aux domaines vitaux (ex : Meulemans, 2021).

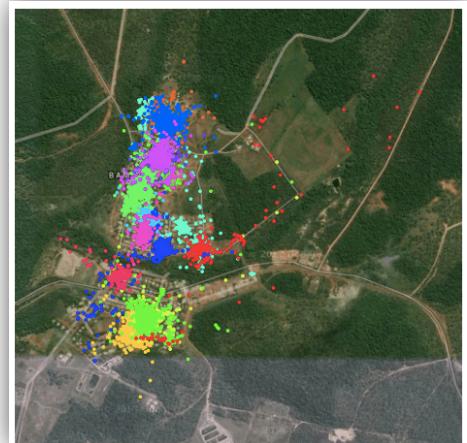


Figure 7: Représentation directe des géolocalisations (Dürr & Ward, 2014)

Ceux-ci peuvent être déclinés de manière à représenter des sous-ensembles temporels (ex : Stoner et al., 2008) ou différents isoplèthes (ex : Pfeiffer et Meyburg, 2015). Ceci permet d'apprécier la variation de superficie du domaine vital en fonction des paramètres choisis (Figure 8).

La superposition graphique de différents estimateurs du domaine vital (Figure 9) permet d'évaluer leurs recouvrements (ex : Grueter et al., 2009 ; Richard-Hansen et al., 2014). Cette représentation permet de mettre visuellement en évidence les similarités spatiales entre les estimations, complétant ainsi l'information de superficie.

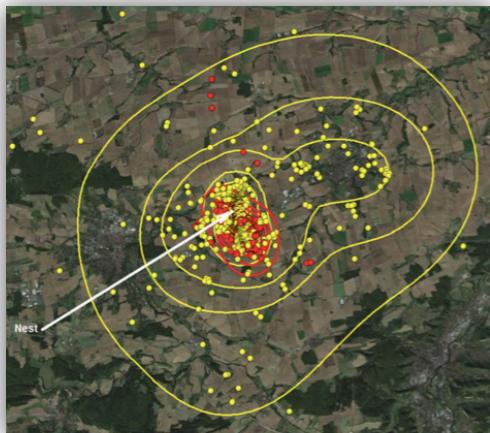


Figure 8 : Isoplèthes d'un kernel de densité, (Pfeiffer & Meyburg, 2015)

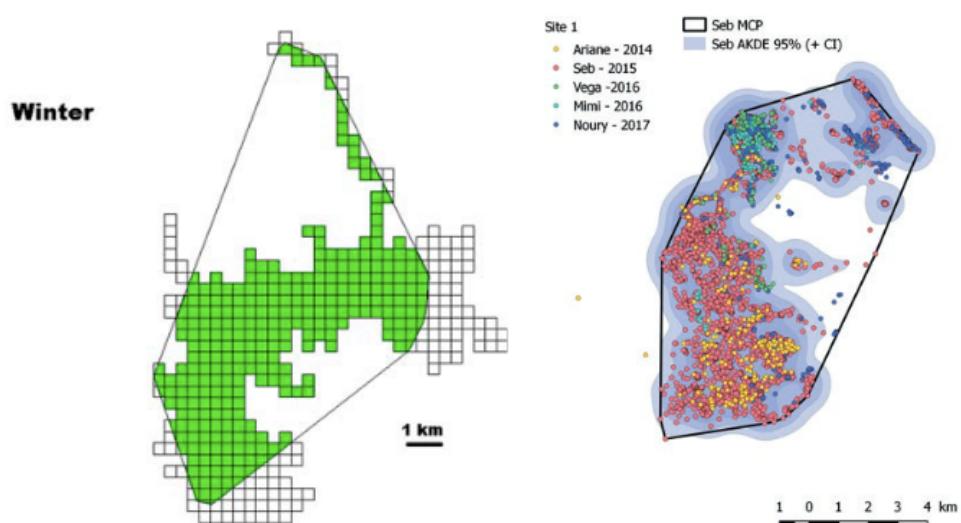


Figure 9 : Superposition des estimations d'un domaine vital par MCP avec un Grid Cell (gauche- Grueter et al, 2009) et un kernel de densité (droite - Richard-Hansen et al., 2014)

Occasionnellement, les trajectoires peuvent être représentées sous forme de lignes entre points consécutifs (Stoner et al., 2008 ; Stanley et al., 2021). Parfois une flèche indique le sens de déplacement (Jarolimek, 2014). Cette représentation des trajectoires (Figure 10) est généralement moins répandue et habituellement réservée pour illustrer des migrations (ex : Baert et al., 2018 ; Wolfson et al., 2017) ou des individus en dispersion (ex : Walton et al., 2018). (Figure). En effet, l'information devient rapidement peu informative en raison de l'enchevêtrement d'un nombre conséquent de lignes.

Dans le cas des individus dispersants, un vecteur entre la position initiale et la position finale peut être représenté. En terme d'analyse, il permet de comparer la distance vectorielle et la distance réelle de la dispersion (Stoner et al., 2008).

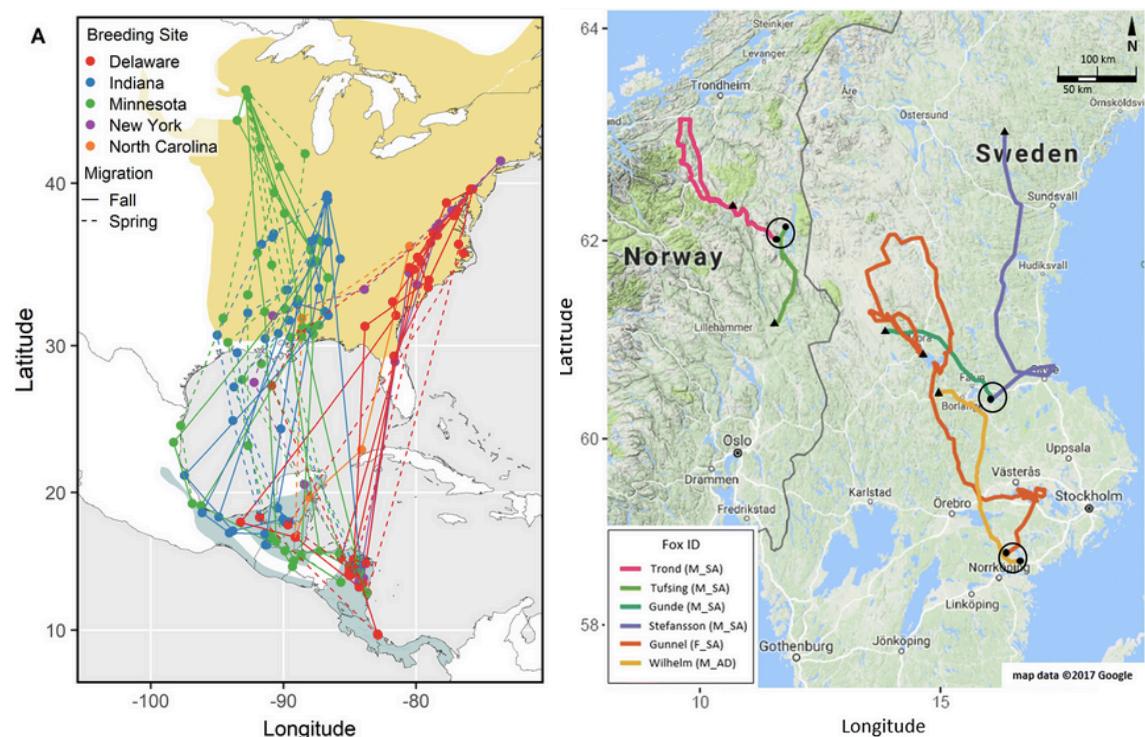


Figure 10 : Trajectoires d'un oiseau migrateur (Gauche : Stanley et al., 2021), dispersion de renards (Droite : Walton et al., 2018)

2.3.4. Analyses graphiques

2.3.4.1. Activité d'un animal

L'activité d'un individu peut s'interpréter à travers une multitude de représentations graphiques. Ainsi, les proportions de géolocalisations peuvent se comparer selon plusieurs facteurs de regroupement : classes d'altitude (Mosnier et al., 2008), distances à des points de référence tels que des chemins (ex : Licoppe, 2005) ou des villages (Athreya et al. 2014).

La vitesse moyenne constitue un autre moyen de visualiser l'activité d'un individu. Celle-ci peut être déclinée selon la période du jour (Peris, 2020)(Figure 11), de l'heure de la journée (Merill et al., 2003) ou la température ambiante (Krone et al., 2009). De manière alternative, l'activité peut également être représentée par le déplacement médian de l'individu en fonction de l'heure de la journée (de Melo et al., 2006).

La distance parcourue constitue un indicateur fréquemment utilisé à différentes échelles de temps. L'évolution de la distance mensuelle ou journalière parcourue par un individu (ex : Stoner et al., 2008 ; Baxley & Qualls, 2009) permet de percevoir l'évolution de l'activité d'un animal au cours de l'année (Figure 12).

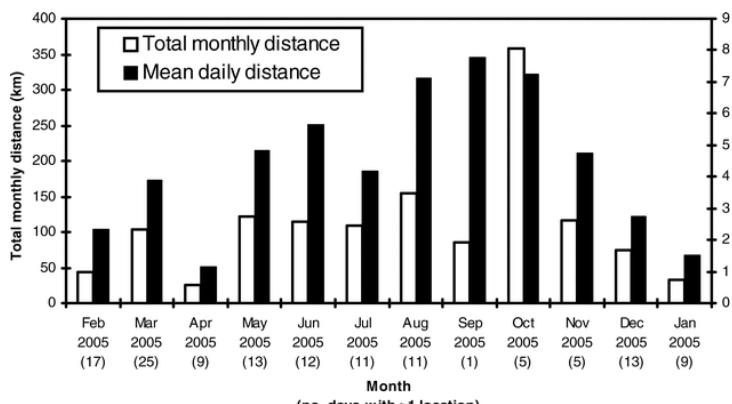


Figure 12: Représentation de la distance mensuelle totale et moyenne quotidienne parcourue (Stoner, 2008)

L'évolution de la distance entre l'individu et un point de référence peut aussi être visualisée par le « *Net Square Displacement* » (NSD), qui consiste à mesurer la distance en ligne droite entre un point de référence et l'ensemble des localisations suivantes (Bunnefeld et al., 2011, Morelle et al., 2017 ; Walton et al., 2018). L'évolution de cette valeur au fil du temps permet de rattacher l'animal à un comportement sédentaire, migrateur ou nomade (Figure 13).

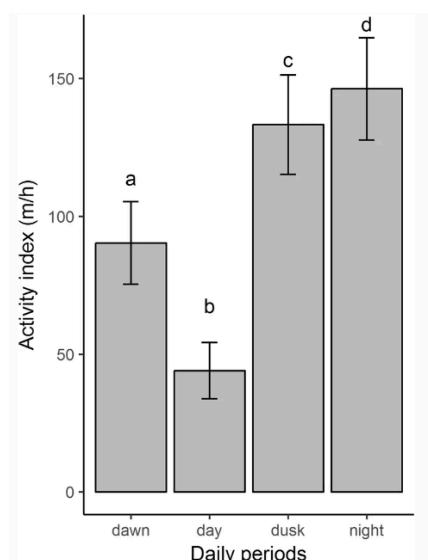


Figure 11: Vitesse moyenne selon la période du jour (Peris et al., 2020)

La distance maximum journalière quotidienne définie par le plus grand écart entre deux localisations de la même journée peut se décliner par une moyenne mensuelle ou selon un graphique de fréquences (Acácio et al., 2021). La densité de fréquence des distances parcourues par intervalle de temps (ex : Meulemans, 2021) permet de comparer l'activité de différents individus.

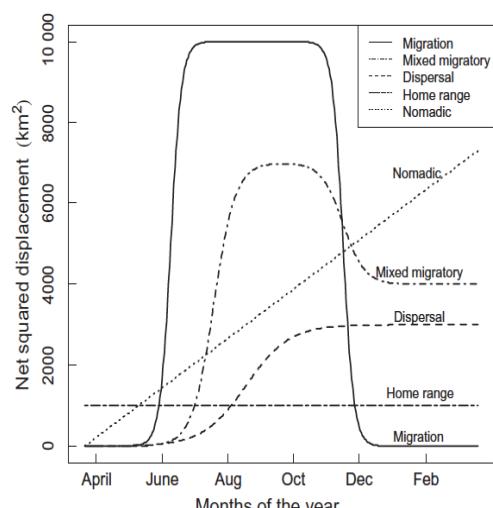


Figure 13: Représentation du NSD (Bunnefeld et al., 2011)

2.3.4.2. Evolution de la superficie des domaines vitaux

La courbe d'évolution de la surface cumulée de l'estimateur du domaine vital en fonction du temps écoulé depuis le début du suivi permet d'évaluer si la superficie du domaine vital atteint une valeur stable avec les données disponibles (ex : Girard et al., 2002, Richard-Hansen et al., 2014 ; Oleksy et al., 2015, Yiu et al., 2016) ou en fonction du nombre de géolocalisations acquises (ex : Bandeira de Melo et al., 2007). Ceci permet de vérifier la convergence de la superficie vers une valeur asymptotique (Figure 14).

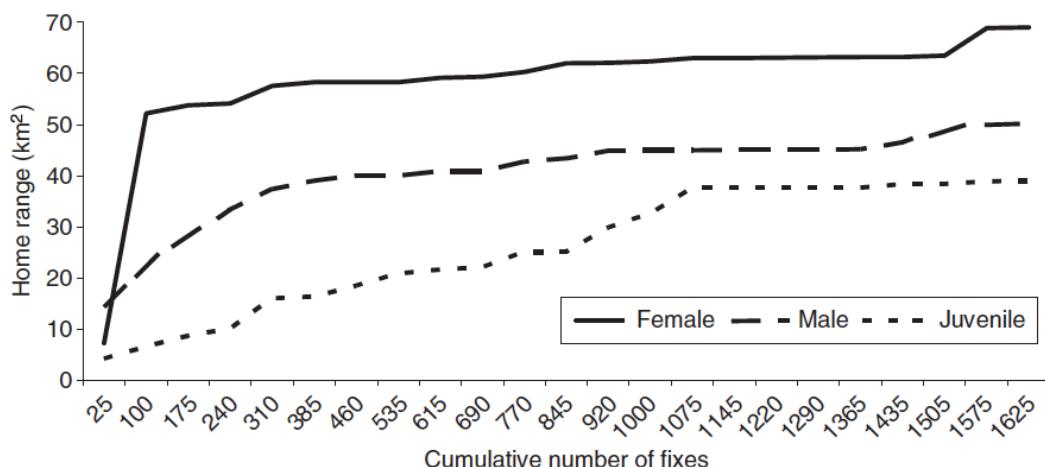


Figure 14 : Evolution de la taille du domaine vital en fonction du nombre de géolocalisations (Bandeira de Melo et al., 2007)

Cette évolution de la superficie du domaine vital peut aussi s'évaluer à travers l'examen de la superficie du domaine vital journalier au cours du temps (ex : Bandeira de Melo et al., 2006). De même, il est possible de comparer les superficies des domaines vitaux sur différents cycles temporels tels que les années (Scull et al., 2012), saisons (Deacon & Smit, 2017, Reading et al., 2019 ; Teng et al., 2021), semaines (Peiffer & Meyburg, 2015) (Figure 15). Cette évolution peut aussi inclure d'autres facteurs tels que le sexe (ex : Miller et al., 2012) ou la classe d'âge (ex : Meulemans, 2021) afin de faire apparaître d'éventuelles différences significatives.

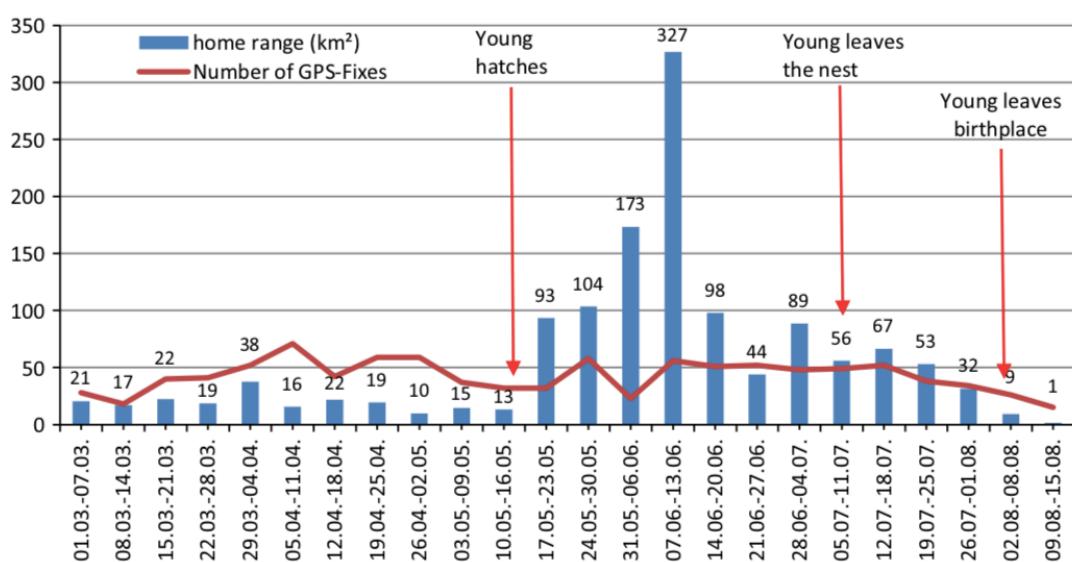


Figure 15: Evolution du domaine vital hebdomadaire d'un milan royal (Peiffer & Meyburg, 2015)

2.3.4.3. Utilisation de l'habitat

La projection des localisations peut être effectuée sur un fond de carte reprenant les différents types d'habitats (ex : Henderson et al, 2018)(Figure 16).

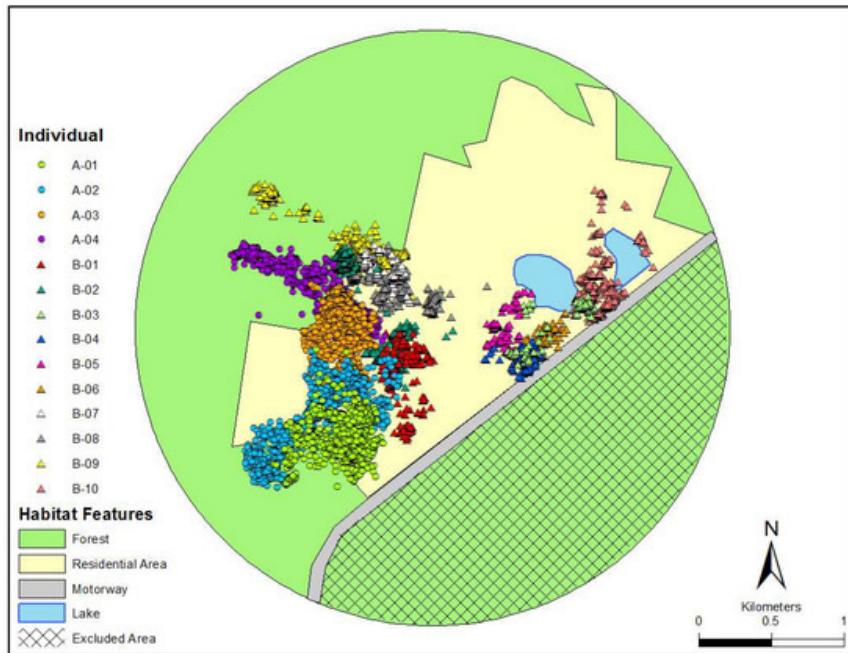


Figure 16 : Projection des points sur un shapefile reprenant les différents types d'habitats (Henderson et al., 2018)

La proportion de la surface occupée par chaque type d'habitat pour un même individu au moyen de différents estimateurs (ex : Kauhala & Auttila ; 2010) permet d'évaluer la fiabilité de l'analyse. Le même principe peut être utilisé pour comparer plusieurs individus entre eux, en examinant la composition des domaines vitaux calculés au moyen d'un même estimateur (ex : Alfred et al., 2012)(Figure 17). Cette proportion des habitats dans le domaine vital peut aussi être déclinée mensuellement, via l'utilisation éventuelle d'une moyenne dans le cas d'un suivi de plusieurs individus (Meulemans, 2021).

Une comparaison graphique entre l'habitat utilisé et disponible, permet de visualiser l'attractivité d'un habitat donné sur un animal (ex : Richard-Hansen et al., 2014 ; Lesmeister et al., 2009)(Figure 18).

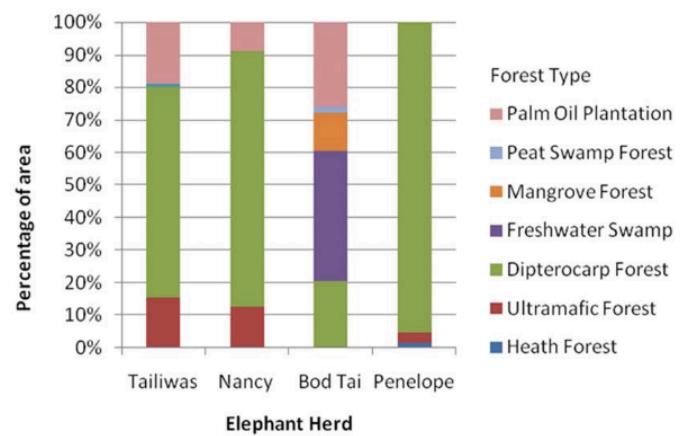


Figure 17 : Comparaison de la proportion d'habitats dans le domaine vital (Alfred et al., 2012)

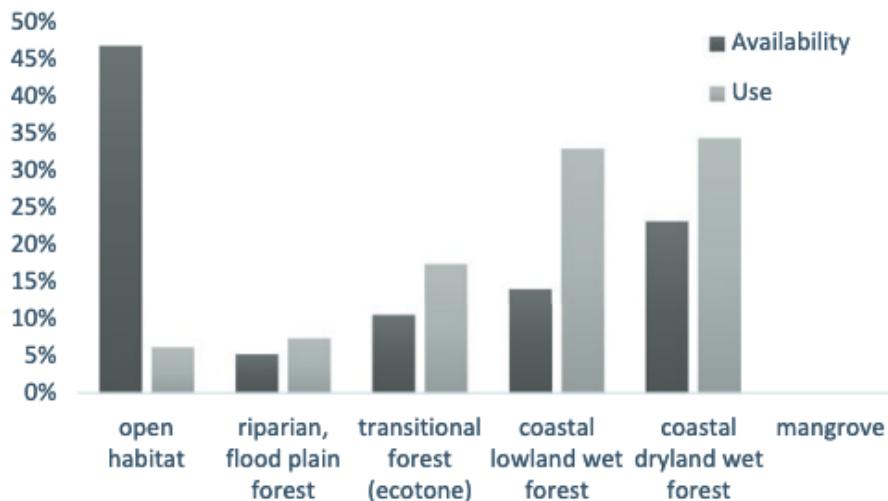


Figure 18 : Comparaison par habitat des superficies disponibles et utilisées dans le domaine vital d'un pécari (Richard-Hansen et al., 2014)

2.3.4.4. Distances par rapport à un élément spécifique

La courbe cumulée de l'utilisation de l'espace par rapport à un point permet de mettre en évidence l'influence de cet élément dans la répartition spatiale des individus (ex : Cornélis et al., 2011). Cette représentation peut également utiliser la fréquence des géolocalisations en regard d'une distance à un élément spécifique tel qu'un nid (ex : Peiffer & Meyburg, 2015)(Figure 19), une plage (Schofield et al., 2007) ou des aires protégées (Blakey et al., 2018).

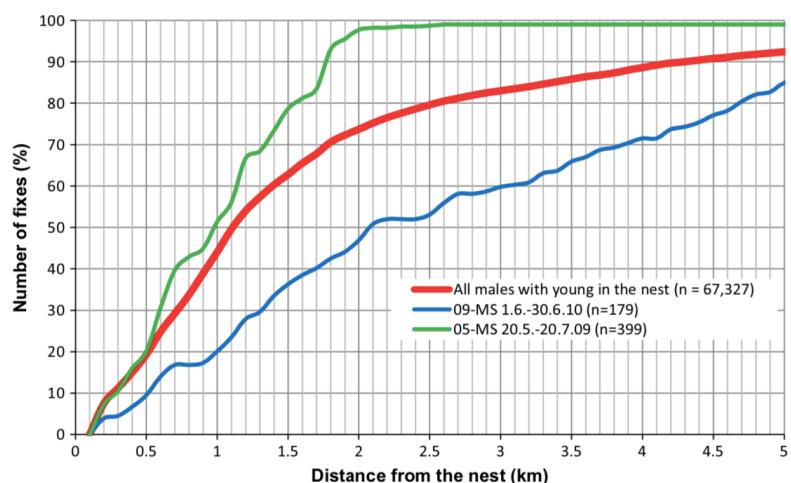


Figure 19 : Fréquences cumulées des localisations en fonction de la distance au nid (Peiffer & Meyburg, 2015)

III. Méthodes

3.1. Définition du concept général

La première étape de ce travail a consisté à définir le concept global de l'application et ses principales caractéristiques. La finalité était de produire une approche visuelle et interactive des données afin de faciliter la compréhension des déplacements et de l'utilisation de l'habitat des animaux suivis par GPS.

Il semblait donc essentiel que le logiciel puisse garantir un traitement rapide des données et permettre une certaine flexibilité tant sur le fond que sur la forme : choix des paramètres pour les calculs, des estimateurs de l'habitat, définition des couleurs, de l'affichage, etc.

Afin de valoriser ce programme en dehors de son utilisation au DEMNA, sa mise en ligne sous la forme d'une application web a été imaginée. En effet, en tant que Service Public, il semblait opportun de pouvoir ouvrir son utilisation à des équipes de recherche, des universités, des gestionnaires, des ASBL, des ONG, etc. Le choix de se diriger vers une application en ligne se justifiait pour plusieurs raisons : maintenance centralisée, pas d'installation à effectuer, limitation des problèmes de compatibilité entre les différents systèmes d'exploitation, etc.

En cours de développement, le programme fut baptisé « *TracksMining* ». « *Tracks* » faisant référence aux déplacements des animaux et « *Mining* » à l'exploration de ces données. L'utilisation d'un nom anglais semblait plus cohérente avec l'idée que le programme puisse être diffusé auprès d'un plus large public.

3.2. Prototypage

La deuxième phase du développement du projet impliquait de réaliser un prototype du logiciel. L'objectif était de pouvoir rapidement valider le concept avec un programme basique disposant de quelques fonctionnalités élémentaires. Cette étape a permis d'appréhender la faisabilité du projet et d'évaluer le temps nécessaire pour l'implémentation des fonctionnalités prévues. Ceci a permis de définir des délais réalistes pour la tenue d'un calendrier et d'évaluer si les ressources disponibles seraient suffisantes pour mener à bien ce travail : jeux de données pouvant être mis à disposition, ressources informatiques, adéquation entre le temps estimé et le temps professionnellement imparti pour ce travail.

Le prototypage a commencé au mois de septembre 2021 pour être présenté aux co-promoteurs à la fin du mois de novembre de la même année. Cette réunion a permis de présenter le concept du projet sur base d'éléments concrets et de bénéficier d'une vue globale assez claire sur le résultat pouvant être atteint au terme de l'échéance.

3.3. Définition des critères d'évaluation

En accord avec les promoteurs, il fut convenu que sur le plan technique, l'objectif pour l'évaluation était d'arriver à une première version stable et fonctionnelle de l'application. Cet objectif général fut traduit en 8 critères auxquels le logiciel devait satisfaire dans les délais impartis (Table 1).

Table 1 : Définition des critères permettant d'évaluer l'atteinte de l'objectif

Critère	Le logiciel :	Explication
1	est fonctionnel dans un contexte professionnel.	L'application doit pouvoir être utilisée dans le cadre des missions du DEMNA afin de générer les cartes interactives, principalement destinées à communiquer aux agents du DNF les déplacements des animaux marqués sur leurs cantonnement.
2	intègre une première série de traitements élémentaires	L'application doit permettre la visualisation des éléments les plus utiles pour la compréhension du déplacement des animaux, incluant au minimum : géolocalisations, trajectoires, animation temporelle, estimations du domaine vital par MCP et KDE.
3	est validé par des utilisateurs potentiels	Le logiciel et plus spécifiquement les cartes interactives, doivent avoir été présentés à des utilisateurs potentiels afin de recueillir des avis externes et objectifs sur leur intérêt et leur facilité d'emploi.
4	est validé par rapport à ses résultats	Les résultats fournis par l'application doivent être cohérents avec ceux fournis par d'autres applications déjà validées par la communauté scientifique.
5	est validé par rapport à ses performances	L'application doit permettre un traitement rapide des données, et être optimisée de manière à minimiser les bugs d'exécution ou la survenue de lenteurs extrêmes.
6	est testé sur plusieurs modèles biologiques	L'application doit être testée avec succès sur plusieurs jeux de données différents afin de limiter le sur-ajustement à des espèces ou individus spécifiques. Dans la mesure du possible, les algorithmes utilisés doivent pouvoir s'extraire à un maximum d'espèces animales et de contextes géographiques.
7	est structuré de manière cohérente et logique	L'application doit disposer d'une architecture logicielle ouverte à l'implémentation de nouvelles fonctionnalités et compréhensible pour d'autres développeurs souhaitant s'investir dans le projet. La syntaxe d'écriture doit être claire et bien documentée.
8	dispose d'une interface utilisateur	Une interface pratique et intuitive doit permettre à un utilisateur d'activer les fonctionnalités du programme et de modifier les paramètres des analyses sans devoir recourir à une modification du code.

3.4. Développement de l'application

3.4.1. Introduction

Ce chapitre présente l'application dans son fonctionnement interne, en dehors de l'interface accessible à l'utilisateur. Ces informations ont pour objectif de :

- Mettre en évidence les avantages et certaines limites de l'application.
- Justifier certains choix techniques inhérents à la rédaction du logiciel.
- Clarifier l'organisation interne des différents éléments pour faciliter l'implémentation des extensions futures.

Ce chapitre est volontairement vulgarisé afin de rester accessible à un lecteur non-initié aux concepts de base en programmation informatique.

3.4.2. Choix du langage de programmation

Le choix de Python (Van Rossum & Drake, 2009) s'est rapidement imposé pour le développement de TracksMining. Au-delà d'un attrait personnel, ce choix s'est justifié par plusieurs éléments :

- Une syntaxe d'écriture particulièrement lisible permettant de rédiger des programmes clairs à petite ou grande échelle (Kuhlman, 2009, cité par Srinath, 2017).
- Python est reconnu comme étant le langage informatique présentant la croissance la plus rapide de ces dernières années (Srinath, 2017).
- L'existence d'un typage dynamique permet de ne pas devoir déclarer à l'avance les types des variables qui sont utilisées dans le programme et permet une gestion automatique de la mémoire (Srinath, 2017).
- Plusieurs paradigmes de programmation sont supportés par Python, parmi lesquels la programmation impérative, fonctionnelle, procédurale ou orientée objet (Srinath, 2017).

3.4.3. Syntaxe

TracksMining est caractérisé par une programmation orientée objet (POO). Ce terme désigne un paradigme de développement informatique dans lequel les algorithmes qui constituent un programme sont représentés par des briques logicielles appelées « objets ». Ces derniers contiennent des variables propres (attributs) et des fonctions spécifiques (méthodes) qui définissent leur comportement. Ces objets interagissent entre eux et s'utilisent les uns les autres pour mener le programme à son objectif (Boucheny, 2020).

Cette approche fut choisie pour permettre une maintenance plus aisée du programme, grâce à de nombreuses stratégies spécifiques à ce style d'écriture. Par exemple, les différentes fonctions nécessitent très peu de paramètres car elles accèdent directement aux variables pertinentes stockées sous forme d'attributs. De même, l'héritage entre classes permet de réduire drastiquement le nombre de lignes de code en recyclant certaines parties pour éviter les répétitions.

3.4.4. Utilisation de modules externes

L’application fait appel à des modules externes contenant des blocs de code extérieurs au programme principal (cfr Table 2). Au-delà de permettre un gain de temps significatif dans la rédaction du code, ces modules présentent l’avantage d’être optimisés pour permettre les traitements de données les plus rapides possibles.

« *Folium* » est un module qui permet la manipulation et la visualisation de données géographiques, en les projetant sur une carte interactive. *Folium* constitue une interface entre Python et la bibliothèque JavaScript *Leaflet*. (Source : <https://python-visualization.github.io/folium/index.html>)

« *Pandas* » est un module qui permet l’implémentation de structures de données sur lesquelles il est possible de travailler rapidement et facilement. *Pandas* combine les performances élevées de *Numpy* avec la flexibilité de manipulation de feuilles de calcul et de bases de données relationnelles. Ce module fournit des fonctionnalités sophistiquées permettant notamment de découper, synthétiser, agréger et sélectionner des données (McKinney, 2012).

« *Shapely* » est un module qui permet l’analyse et la manipulation et le traitement d’entités géométriques planaires (Gillies et al., 2007).

« *Pyproj* » est utilisé pour la projection et les calculs géodésiques sur des flux de données géospatiales (McFerren & Van Zyl, 2016).

« *Scikit-Learn* » est un module *Python* qui intègre une large gamme d’algorithmes d’apprentissage automatique destinés à effectuer des classifications supervisées et non-supervisées. L’accent est mis sur la facilité d’utilisation, les performances du module, la documentation et la cohérence de l’API. Il possède un minimum de dépendances et est distribué sous forme d’une license BSD, ce qui encourage son utilisation commerciale ou académique (Pedregosa et al., 2011).

« *Numpy* » est la principale bibliothèque de programmation vectorielle pour Python. Ce module joue un rôle essentiel dans les pipelines d’analyse et de recherche dans différents domaines tels que la physique, la chimie, l’astronomie, les géosciences, la biologie, la finance, etc. (Harris et al., 2020). Dans le language Python, *Numpy* constitue la représentation standard des données numériques et permet l’implémentation efficace de calculs numériques (Van der Walt et al., 2011).

« *Matplotlib* » est un module graphique utilisé pour développer des applications, des scripts interactifs ainsi que pour générer des graphiques et images de qualité au travers de différentes interfaces et systèmes d’exploitation (Hunter, 2007).

« *Flask* » est un module qui permet l’implémentation d’un micro-framework pour le développement de sites et d’applications internet (Grinberg, 2018).

D’autres modules sont également utilisés de manière plus anecdotique : « *Suntime* » (Stopa et al., 2019), « *Fiona* » (Gillies et al., 2008), « *colors* » (<https://pypi.org/project/Colors/>), etc.

Table 2 : Résumé des modules externes utilisés par TracksMining

Module	Utilité principale dans l'application
Folium	Construction de la carte interactive et des éléments d'interaction avec l'utilisateur (popups, tooltips, etc.).
Pandas	Structure générale des données et opérations vectorielles.
Shapely	Construction et manipulation des entités géométriques (Estimateurs de l'habitat, localisations, trajectoires, etc).
Pyproj	Conversion des localisations géographiques entre différents CRS.
Scikit-Learn	Algorithme des plus proches voisins, Kernel Density et K-means.
Numpy	Conversion des données numériques entre Pandas vers Scikit-Learn.
Matplotlib	Création de graphiques pour visualiser le résultat des analyses.
Flask	Création de l'interface utilisateur
Suntime	Détermination des éphémérides locales.
Colors	Création des gradients de couleurs.
Fiona	Conversion du format .geojson vers .shp

3.4.5. Architecture logicielle

L'architecture logicielle de TracksMining suit le motif Modèle-Vue-Contrôleur (MVC). Ce modèle de conception en programmation est très utile pour l'architecture de systèmes logiciels interactifs (Leff & Rayfield, 2001). Depuis sa création, le MVC a été largement adopté comme une approche de développement d'applications internet contenant une composante de programmation côté serveur (Heidke et al., 2008). Le modèle MVC cherche à isoler l'interface utilisateur de la logique des modèles qui composent le programme (Heidke et al., 2008).

3.4.6. Logique générale

La logique générale de TracksMining, hébergée dans le contrôleur, consiste à faire intervenir différentes classes de traitement qui vont successivement traiter le jeu de données brut (ex : phase de nettoyage des données, phase de détection des *outliers*, phase de calcul, etc.).

En parallèle, les classes construisent progressivement un objet (« *Logger* ») qui référence l'ensemble des résultats calculés lors de chaque étape. En stockant les données, il agit comme une mémoire cache et permet aux différentes classes de traitement de ne pas avoir à recalculer des éléments communs. Ceci permet d'augmenter significativement la vitesse de calcul.

Son utilisation permet également un découplage des classes qui ne doivent plus transférer des arguments entre elles ni respecter un ordre d'appel bien défini. Par conséquent, il devient facile d'insérer, d'interchanger ou de supprimer des classes de traitement au fur et à mesure de l'évolution du programme.

3.4.7. Classes de traitement

Les classes de traitement constituent la base du modèle de TracksMining. Chaque classe encapsule une série de méthodes bien spécifiques destinées au traitement ou à l'analyse des données et regroupées selon une logique commune en fonction de leur objectif (ex : nettoyer les données, calculer un MCP, etc.).

Au sein de chaque classe de traitement, il peut exister des classes plus spécifiques qui héritent des méthodes et attributs de la classe d'origine en y ajoutant des spécificités. Un bon exemple est donné avec la classe LocohComputer(), dont héritent les deux classes K-Locoh() et R-Locoh(), deux variantes de l'estimateur du domaine vital. Les méthodes communes se retrouvent dans la classe d'origine tandis que les spécificités aux deux algorithmes se trouvent dans les classes héritières.

Certaines classes de traitement sont essentielles et par conséquent nécessaires au bon déroulement du flux d'exécution. D'autres classes sont optionnelles et peuvent être activées ou désactivées selon les besoins de l'analyse. Tous les traitements sont modulés par des paramètres qui se définissent via l'interface utilisateur (ex : valeur isoplèthe, résolution, couleurs, etc.).

3.4.8. Fichiers produits

L'application génère un fichier compressé qui se décompose en plusieurs éléments (Figure 20):

- Une carte interactive (.html) pour visualiser les données dans un navigateur.
- Des graphiques au format (.jpg ou .png) pour les analyses.
- Les données brutes (.txt) nettoyées et accompagnées de nouvelles variables, pour d'éventuels traitements ultérieurs.
- Un répertoire complet (.json) comprenant l'ensemble des résultats calculés par l'application. Ces données correspondent au logger.
- L'ensemble des paramètres (.json) utilisés pour le processus.
- Les objets géométriques (.shp ou .geojson) correspondant aux estimateurs de l'habitat souhaités pour d'éventuels traitements complémentaires par SIG.

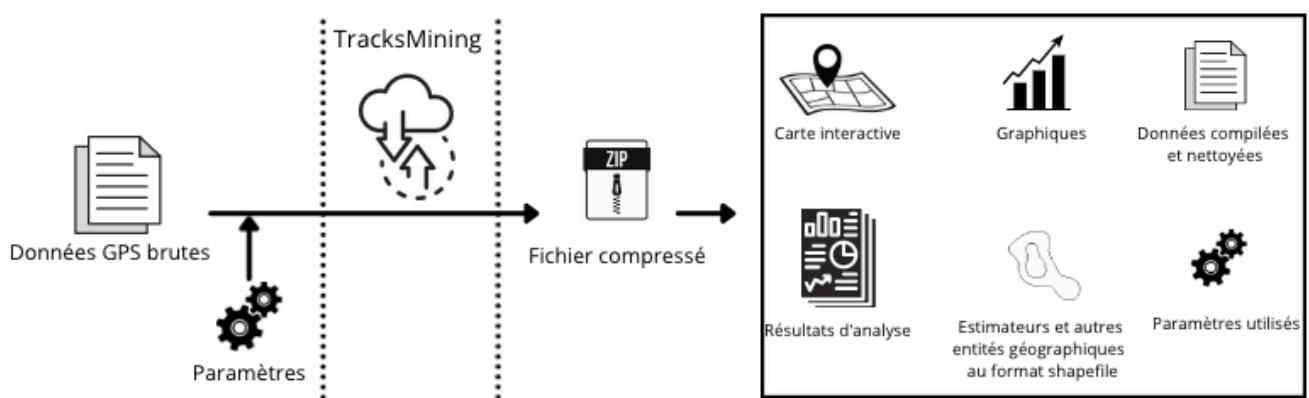


Figure 20 : Vue d'ensemble du processus

IV. Résultats

Ce chapitre reprend, de manière non-exhaustive, les fonctionnalités de l'application. Le code source détaillé au moyen d'un Jupyter Notebook ainsi que des tutoriels vidéos qui couvrent plus en détails les différentes fonctionnalités de l'application, sont disponibles en libre accès à l'adresse suivante :

www.github.com/QuentinWATTHEZ/TracksMining

4.1. Fonctionnalités de l'application

4.1.1. Accès à l'interface utilisateur

Dans sa version actuelle, l'accès à l'interface utilisateur se réalise à partir d'un fichier python « *TracksMining_localserver.py* ». L'exécution de ce fichier crée un serveur local pour héberger l'interface. Lors de l'exécution de ce fichier, une adresse IP est affichée dans la console et redirige l'utilisateur vers l'application, qui s'ouvre dans un navigateur internet.

4.1.2. Chargement des données

Sur la page d'accueil, un explorateur de fichier (Figure 21) permet la sélection du fichier de données sur le disque local de l'utilisateur. L'application permet le chargement d'un fichier au format .txt ou .csv. Pour être valide, le fichier doit comprendre au minimum 4 colonnes dans un ordre indifférent et nommées : « LAT », « LON », « TIME », « DOP » (Table 3).

Après sélection, l'utilisateur peut valider le bouton « continuer » afin de passer au paramétrage.

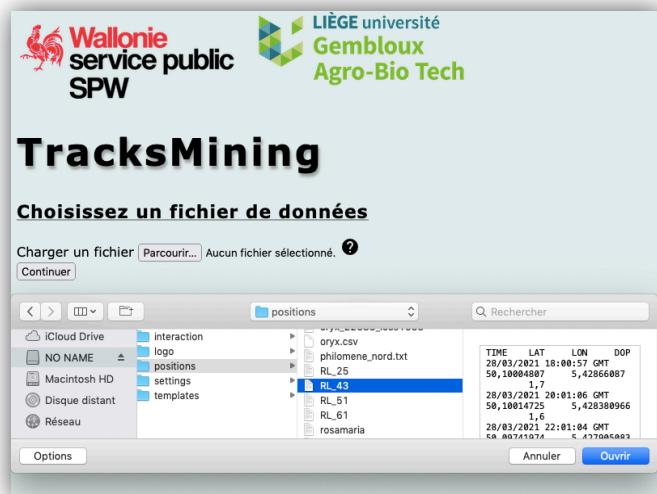


Figure 21 : Page d'accueil et explorateur de fichier

Table 3 : Récapitulatif des colonnes essentielles à un fichier valide

Colonne	Contenu
LAT	Latitude des géolocalisations dans le CRS WGS84 (ex : 50,547232)
LON	Longitude des géolocalisations dans le CRS WGS84 (ex : 4,547232)
TIME	Date et l'heure de la géolocalisation.
DOP	Indicateur numérique de la précision de la géolocalisation.

Le nombre maximum de lignes de données n'est pas restreint mais certaines fonctionnalités du logiciel vont nécessiter un minimum d'information pour pouvoir être activées. La décimale des colonnes numériques (LAT, LON, DOP) peut être indifféremment symbolisée par une virgule ou un point. Pour la colonne « TIME », TracksMining convertit automatiquement la date et l'heure depuis une chaîne de caractères vers une donnée temporelle, ce qui évite de devoir préciser le format des données. Ce système n'étant toutefois pas infaillible, il est recommandé de respecter l'un des 9 formats de données recommandés (Table 4).

Table 4 : Formats de dates recommandés (exemple du 18 avril 2021 à 15h53)

Format ISO 8601	Exemple
jj/mm/aa 'T' hh:mm	18/04/21 15:53
jj-mm-aa 'T' hh:mm	18-04-21 15:53
jj/mm/aa 'T' hh:mm:ss	18/04/21 15:53:32
jj-mm-aa 'T' hh:mm:ss	18-04-21 15:53:32
jj/mm/aaaa 'T' hh:mm:ss	18/04/2021 15:53:32
aaaa-mm-jj 'T' hh:mm:ss	2021-04-18 15:53:32
jj/mm/aaaa 'T' hh:mm:ssZ	18/04/2021 15:53:32 GMT
jj/mm/aa 'T' hh:mm:ssZ	18/04/21 15:53:32 GMT
aaaa/mm/jj 'T' hh:mm:ssZ	2021-04-18 15:53:32 GMT

4.1.3. Scan du fichier

Après validation du fichier, un premier morceau de code s'exécute afin de réaliser les actions suivantes :

- Vérifier que le fichier de données contient les 4 colonnes essentielles.
- Compter les données.
- S'assurer que les données soient au bon format.
- Remettre les données dans l'ordre chronologique si nécessaire.
- Vérifier la première et la dernière date disponible.
- Faire une première évaluation de présence de géolocalisations potentiellement anormales.

L'exécution de ce script est très rapide (<0.5s pour 20.000 données) et passe inaperçue dans la transition entre les pages. Si le fichier est valide, la transition s'opère sur la page de configuration des paramètres. Dans le cas contraire, un message d'erreur s'affiche sur la page pour inviter l'utilisateur à charger un autre fichier ou à effectuer les modifications requises.

4.1.4. Configuration des paramètres

La page de configuration des paramètres contient un formulaire interactif permettant de définir les préférences pour le traitement de ses données (Figure 22). Cette interface permet à l'utilisateur de paramétriser les cartes et les traitements selon des valeurs spécifiques, d'appliquer des filtres sur les données, de personnaliser le style de la carte interactive, etc. Le scan du fichier permet de pré-remplir les champs avec les valeurs les plus adaptées, de suggérer des actions (ex : nettoyer les *outliers*) ou de limiter les entrées à certaines valeurs (ex : dates disponibles pour le filtre temporel).

4.1.5. Options et paramètres

4.1.5.1. Nom de l'animal

Cette donnée est utilisée pour nommer les fichiers et pour personnaliser certains éléments de la carte interactive. Dans le cas où le fichier de données brut contient une colonne nommée « ID », cette information est récupérée automatiquement et pré-remplie dans le formulaire.

4.1.5.2. Gestion des outliers :

L'utilisateur peut préciser comment il souhaite gérer les géolocalisations qui semblent incohérentes :

- « *Conserver* » : aucune donnée n'est écartée.
- « *Signaler* » : les données suspectes restent incluses sur la carte et dans les analyses, mais font l'objet d'une distinction visuelle sur la carte interactive (Figure 23).
- « *Supprimer* » : l'application supprime les positions identifiées comme *outliers*.

Figure 22 : Formulaire des paramètres

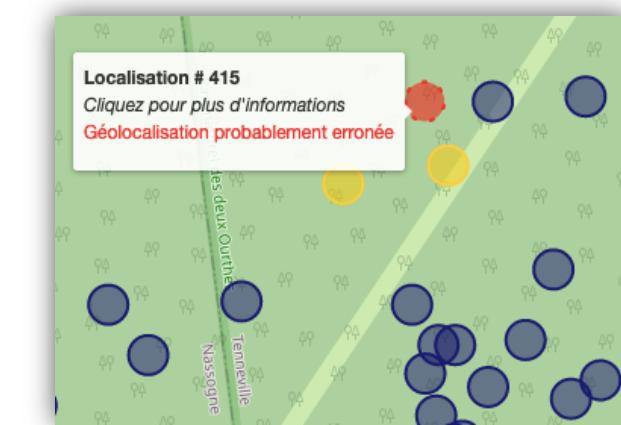


Figure 23 : Signalement visuel d'un outlier potentiel

4.1.5.3. *Filtrer les données*

Cette option permet d'appliquer et combiner des filtres, afin de sélectionner un échantillon du jeu de données original.

- « *Filtrer selon une plage de date* » : les données sont conservées entre une date initiale et une date finale.
- « *Filtrer selon une valeur de DOP* » : les données supérieures à la valeur indiquée sont écartées.
- « *Filtrer selon une plage horaire* » : les données sont conservées entre une heure de début et de fin.
- « *Filtrer selon les coordonnées* » : les données sont conservées si elles sont situées à l'intérieur d'un rectangle défini par une latitude et longitude minimale et maximale.

4.1.5.4. *Couche de fond*

Cette option permet de définir la couche qui sera utilisée pour le fond de carte de l'application (Figure 24).

- « *OpenStreetMap* » : carte du monde sous licence libre.
- « *StamenTerrain* » : topographie et noms de rues.
- « *CartoDBPositron* » : couche de fond au design très épuré.
- « *Aucune couche de fond* » : les données sont projetées sur fond gris uniforme.

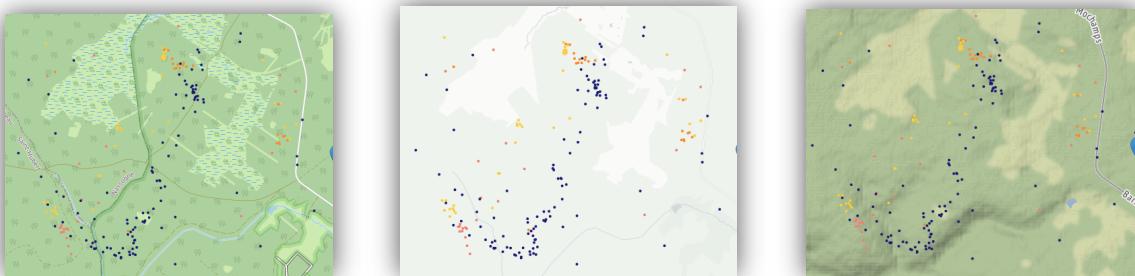


Figure 24 : de gauche à droite : *OpenStreetMap*, *CartoDBPositron*, *StamenTerrain*

4.1.5.5. *Rythme circadien*

Cette option permet de définir si les géolocalisations doivent se distinguer visuellement selon la période du jour à laquelle elles correspondent.

- « *Uniforme* » : aucune distinction n'est faite entre les points selon la période du jour.
- « *Bimodal* » : l'application distingue les géolocalisations en fonction de la période diurne ou nocturne.
- « *Quadrimal* » : l'application distingue les géolocalisations en fonction de la période diurne ou nocturne mais distingue également l'aube et le crépuscule.

La classification est réalisée selon les éphémérides réelles et relatives à la localisation de l'individu.

4.1.5.6. Choix des couleurs

Cette option permet de définir les différentes couleurs qui seront utilisées pour les géolocalisations selon leur classification en fonction de la période de la journée (Figure 25).

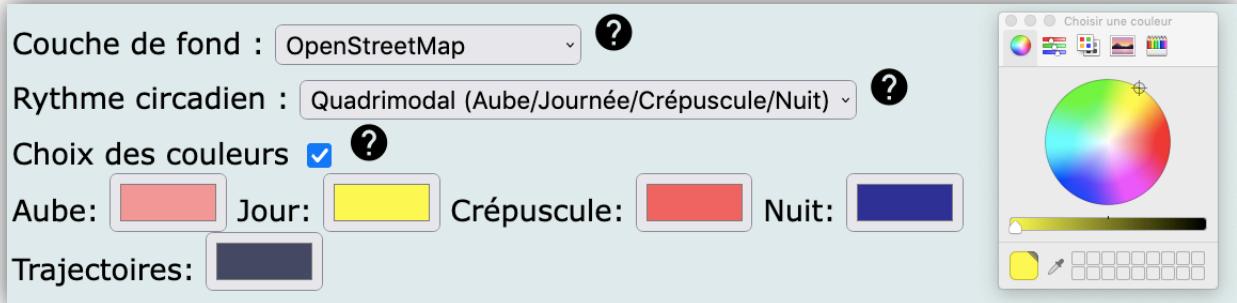


Figure 25 : Vue des principales options de personnalisation du style

4.1.6. Téléchargement des résultats

Lorsque l'ensemble des traitements sont exécutés, un fichier compressé .zip est proposé au téléchargement sur le disque local de l'utilisateur (Figure 26).

Ce fichier est nommé selon la structure suivante : *Identifiant de l'animal + date de début + date de fin*.

Exemple : *Cecil_from_20150101_to_20150701*

Après décompression, le dossier se décompose en 6 sous-dossiers (Table 5):



Figure 26 : Téléchargement du dossier compressé

Table 5 : Récapitulatif du contenu des sous-dossiers

Dossier	Contenu
Charts	Graphiques au format .png
Data_computed	Base de données originale avec les nouvelles variables.
Geometries	Entités géographiques au format .shp ou .geojson
Map	Carte interactive au format .html
Metrics	Répertoire de l'ensemble des données calculées
Settings	Répertoire des valeurs renseignées en tant que paramètres

4.2. Carte interactive et fonctionnalités

4.2.1. Ouverture du document et vue générale

La cartographie interactive au format .html s'ouvre dans un navigateur internet. A l'ouverture, la carte se centre sur le centroïde des données et adapte son niveau de zoom à l'ensemble des géolocalisations de l'individu. Celles-ci, qui constituent les données de base, sont automatiquement affichées à l'ouverture du document (Figure 27).

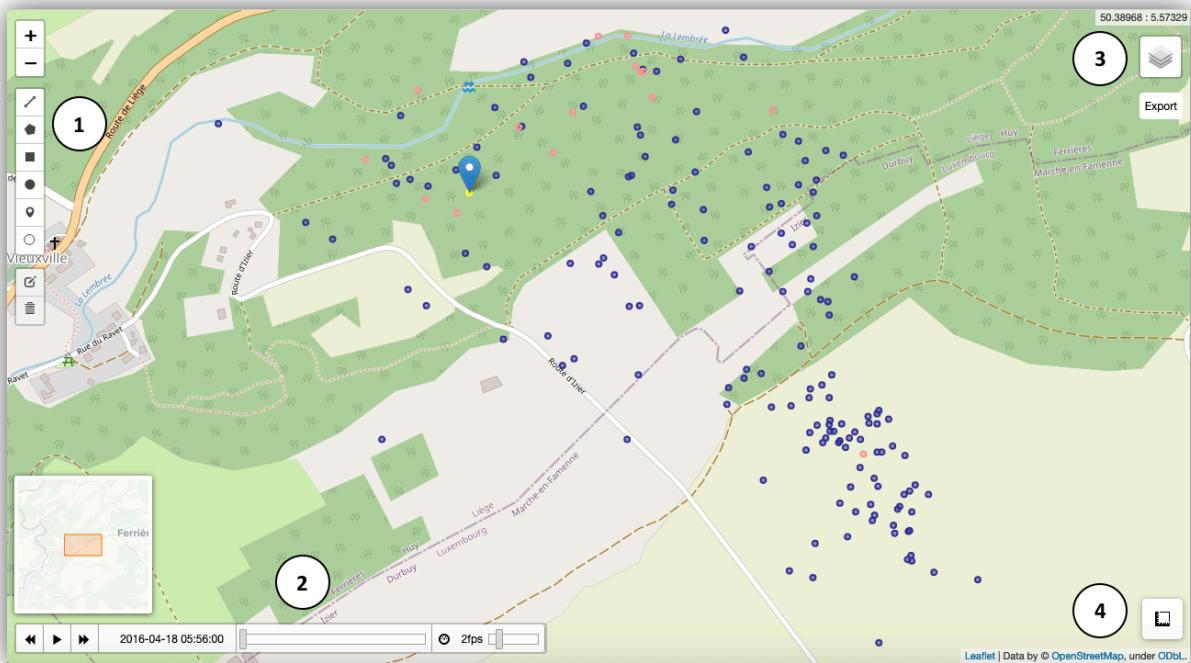


Figure 27 : Vue générale de la carte avec : 1) L'outil de dessin 2) La barre d'animation 3) Le panneau de navigation 4) L'outil de mesure

4.2.2. Navigation entre les couches

Les différentes couches d'informations sont accessibles via l'icône en haut à droite de l'écran. Cette icône active le panneau de navigation entre les différents éléments disponibles et permet de visualiser les informations requises au moment de la configuration des paramètres (Figure 28).

La carte fonctionne via un système de couches superposées. La dernière couche sélectionnée sera amenée à l'avant plan. Cette dernière peut donc masquer les fonctionnalités (*popups*, *tooltips*, etc.) des autres couches situées en arrière-plan.

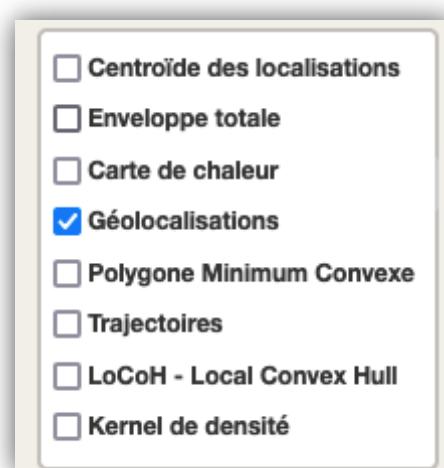


Figure 28 : Panneau de navigation

4.2.3. Description des couches de bases

4.2.3.1. Géolocalisations

Au survol de la souris, chaque localisation affiche un index qui correspond à son numéro d'ordre dans la série chronologique des données. En cliquant sur chaque point, on affiche les données de cette géolocalisation précise: latitude, longitude, date et heure (Figure 29).

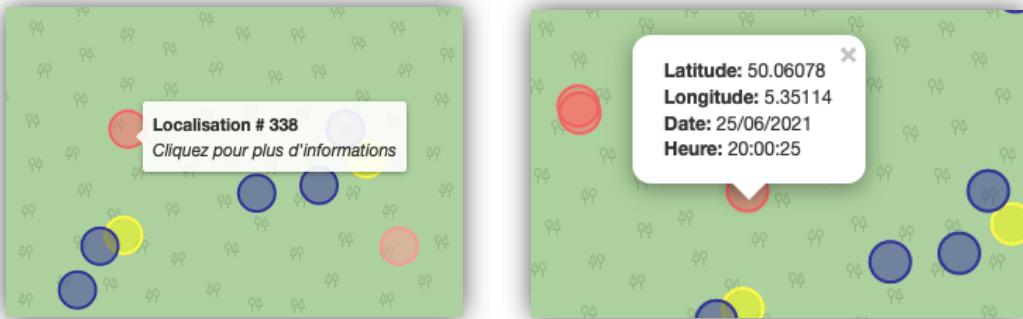


Figure 29 : Tooltip au survol (à gauche) et popup au clic de souris (à droite)

4.2.3.2. Trajectoires

Les connexions vectorielles entre les points peuvent également être affichées au moyen de la couche « Trajectoires ». Elles permettent de mieux comprendre les liens entre géolocalisations successives visualisées à petite échelle (Figure 30).

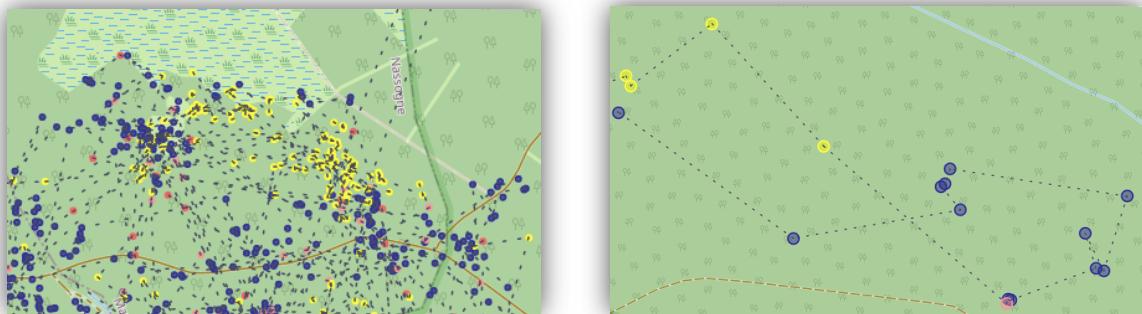


Figure 30 : Les trajectoires à large échelle (gauche) et à petite échelle (droite)

A l'instar des géolocalisations, les trajectoires permettent l'affichage d'informations via un clic de souris : coordonnées de départ et d'arrivée, temps écoulé entre les localisations (Figure 31).

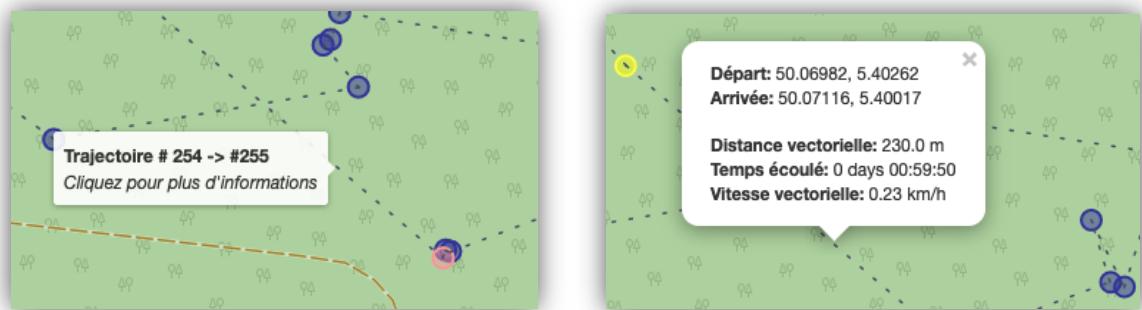


Figure 31 : Tooltip au survol (à gauche) et popup au clic de souris (à droite)

La distance entre géolocalisations successives correspond à la longueur de la ligne droite entre les deux points. Par conséquent, cette valeur est toujours inférieure ou égale à la distance réellement parcourue par l'animal. La mesure de vitesse correspond à la vitesse moyenne à laquelle l'animal a dû se déplacer pour relier ces deux points.

4.2.3.3. *Centroïde des géolocalisations*

Le centroïde des localisations s'affiche sous la forme d'un plus gros cercle de couleur différente (Figure 32). Il représente le « centre de gravité » du nuage de points. Il a peu d'intérêt à priori sur la carte car il ne représente pas une géolocalisation réelle de l'animal. Toutefois, il reste disponible à titre informatif car le centroïde constitue un élément central dans le calcul du polygone minimum convexe et de l'enveloppe totale.



Figure 32 : Centroïde des géolocalisations

4.2.3.4. *Enveloppe totale*



Figure 33 : Enveloppe totale des géolocalisations

4.2.3.5. *Carte de chaleur*

La carte de chaleur représente la densité des géolocalisations. Elle affiche un gradient de couleurs qui représente la densité du nuage de points et s'ajuste à chaque niveau de zoom. Cette fonction permet de visualiser où se situent les grandes concentrations de localisations (Figure 34).



Figure 34 : Carte de chaleur

4.2.3.6. Polygone Minimum Convexe (MCP)

Le Polygone Convexe Minimum (MCP) utilise le même principe que l'enveloppe totale mais permet d'inclure un certain pourcentage de données les plus proches du centroïde (Figure 35). Cette valeur isoplète, habituellement fixée à 95%, peut toutefois se préciser au moment de la configuration des paramètres.



Figure 35 : MCP 95%

4.2.3.7. Kernels de Densité (KDE)

Les kernels de densité permettent d'estimer le domaine vital d'un animal en fonction de la distribution de l'utilisation. La configuration permet d'identifier deux types de zones, correspondant au domaine vital et aux zones « cœur » (Figure 36), d'utilisation plus intense. Par défaut, ces valeurs sont définies respectivement à 95% et 50% de probabilité mais sont entièrement paramétrables.

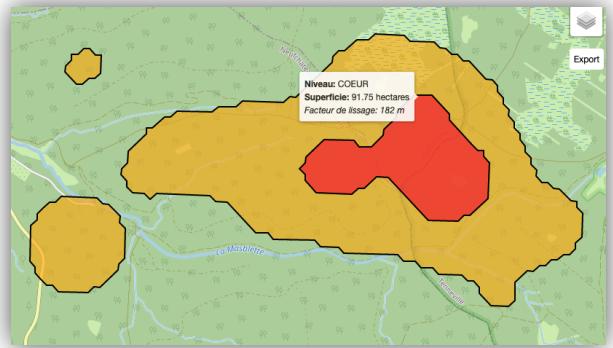


Figure 36 : KDE avec isoplèthes 95% (orange) et 50%

4.2.3.8. Local Convex Hull (LoCoH)

Le Local Convex Hull (LoCoH) permet l'estimation du domaine vital au moyen d'un assemblage de MCP réalisés sur des sous-ensembles de points. La valeur isoplète représente le pourcentage de points qui sont inclus dans l'estimateur. Deux versions de cet estimateur sont disponibles : le K-LoCoH qui construit ses polygones sur un nombre de plus proches voisins fixés (Figure 37) et le R-LoCoH qui inclut les localisations dans un rayon fixé.

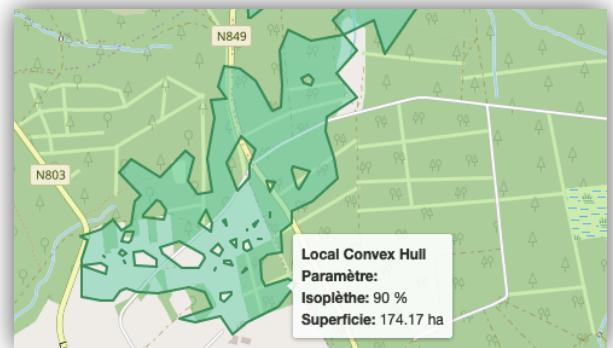


Figure 37 : K-LoCoH avec isoplète de 90%

4.2.3.9. Grid Cell

Cet estimateur du domaine vital superpose une grille aux géolocalisations avec des mailles carrées de largeur définies par l'utilisateur. Chaque maille est colorée selon la densité de points recouverts, ce qui permet de mettre en évidence les zones les plus utilisées par l'animal (Figure 38).

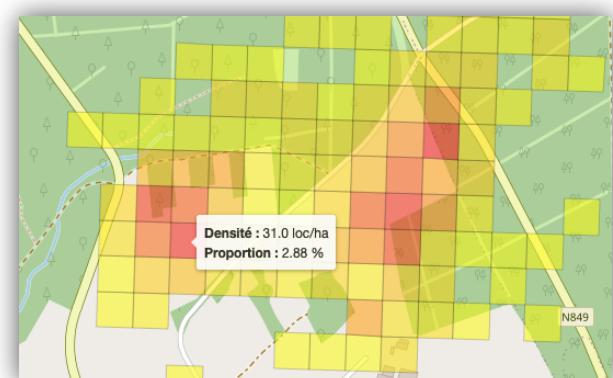


Figure 38 : Grid Cell avec mailles de 100m

4.2.3.10. Interaction

L'interaction consiste à croiser l'information issue des géolocalisations d'un individu avec une autre source de données géographiques (Figure 39). Ces entités spatiales peuvent correspondre à un parcellaire forestier, à des habitats, à des types d'utilisation du sol, etc.

Chaque entité géographique peut contenir un identifiant (ex : compartiment forestier) ou une information de classe (ex : forêt, champ, futaie, pessière...). L'application va examiner la distribution des localisations au sein de chaque entité afin de déterminer l'utilisation de l'espace par l'animal. Cette analyse porte sur plusieurs variables : nombre et densité des localisations au sein de chaque entité mais également dates et récurrence des visites, relation entre espace utilisé et disponible, etc.

De manière analogue au Grid Cell, la couleur de chaque entité sera déterminée par la densité de points recouverts (Figure 38). Chaque objet comportant au minimum une localisation peut afficher un *popup* reprenant la fréquentation mensuelle déclinée selon la période du jour (Figure 40).

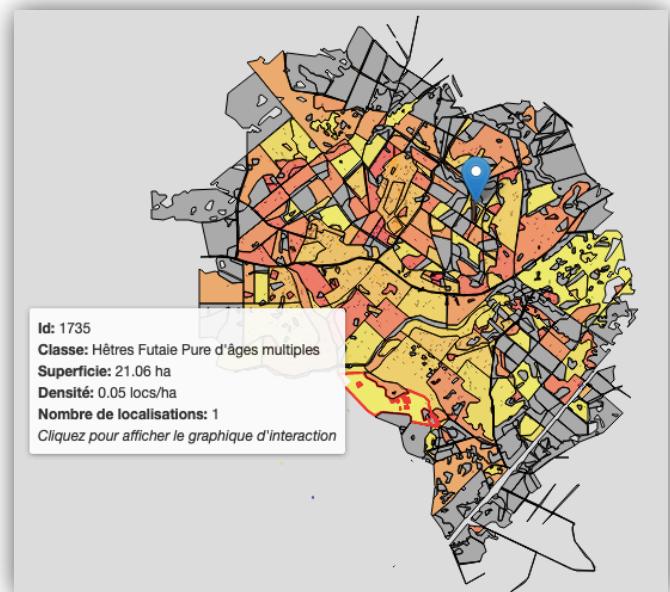


Figure 39 : Interaction d'un raton-laveur avec un parcellaire forestier

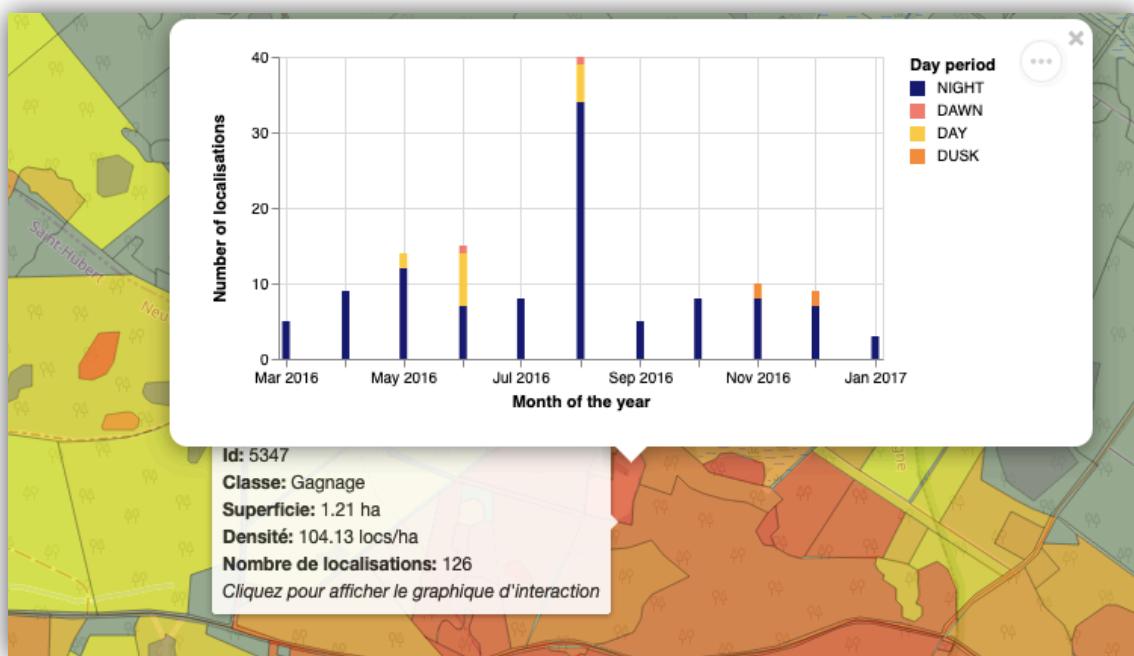


Figure 40 : Histogramme d'interaction mensuelle disponible en popup pour chaque entité

4.2.3.11. Animation des données

L'animation des données fait défiler séquentiellement les géolocalisations, selon des pas de temps et des regroupements temporels déterminés par l'utilisateur (Figure 41). L'enchaînement des séquences peut être manuel ou automatique. Dans ce dernier cas, la vitesse de défilement (frames par seconde) peut être réglée sur la carte. Il est ainsi possible de faire succéder les localisations une par une, ou par paquets (ex : toutes les localisations d'une même journée).

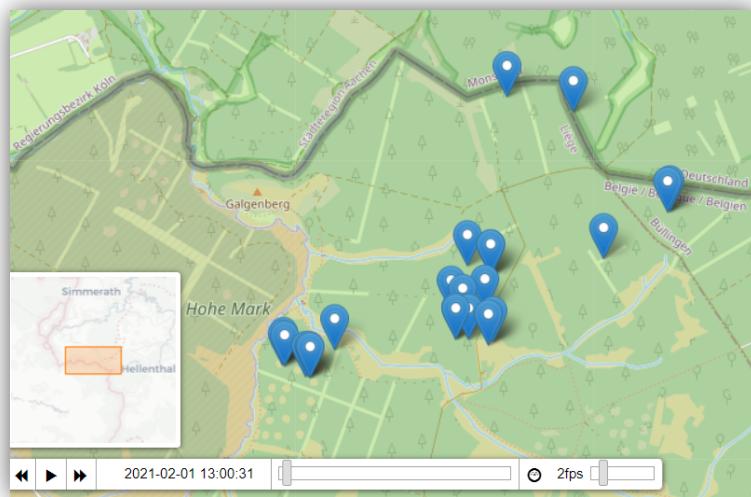


Figure 41 : Animation temporelle, faisant défiler les localisations par groupes de 24h

4.2.3.12. Outil de dessin et de mesure

L'outil de dessin (Figure 42) permet de générer des formes géométriques sur la carte (points, lignes, polygones,...). Ces dernières peuvent être exportées via le bouton 'Export' en haut à droite et téléchargées au format .geojson. Elles peuvent alors être directement utilisées dans un logiciel SIG tel que QGIS ou converties au format .shp pour une utilisation vers ArcGIS. L'outil de mesure, symbolisé par une équerre en bas à droite de la carte, permet de mesurer des distances ou des surfaces sur la carte.

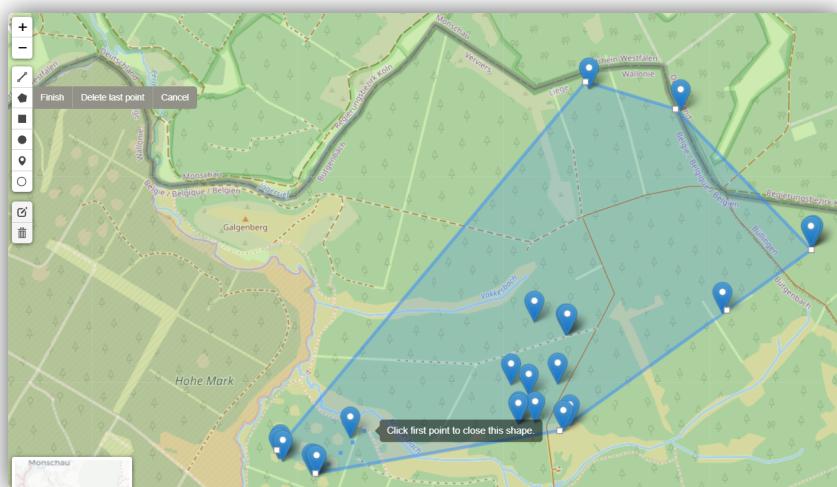


Figure 42 : Outil de dessin, utilisé pour créer un MCP sur les localisations d'une même journée

4.3. Analyses Graphiques

En complément de la carte interactive, l'application peut également produire les résultats sous forme graphique. Cette fonctionnalité peut être activée via l'interface utilisateur, en précisant les graphiques souhaités. Quelques exemples disponibles sont détaillés ci-dessous.

4.3.1. **Distance parcourue**

La distance mensuelle parcourue par un individu est représentée sous forme d'un graphique en bâtonnets qui permet de visualiser l'évolution des déplacements au cours de l'année et l'importance relative de ceux-ci en fonction de la période du jour (Figure 43).

Ce même graphique peut être réalisé selon un découpage quotidien.

4.3.2. **Distance à un point de référence**

La distance peut aussi être représentée de manière relative à un point de l'espace qui sert de référence pour le calcul des distances relatives (Figure 44). Cette visualisation des données permet une première appréciation du comportement de l'animal sur la période étudiée : sédentaire, migrateur ou dispersant. A court terme, la localisation de ce point de référence pourra être définie dans les paramètres.

4.3.3. **Evolution de la superficie du domaine vital**

Le choix d'une valeur isoplète dans le cas d'un MCP ou d'un seuil de probabilité dans le cas du KDE, constitue un paramètre arbitraire pour calculer un estimateur de l'habitat. Cependant cette valeur peut faire varier de manière significative la superficie du domaine vital et par conséquent, modifier les interprétations qui en découlent. Ce graphique permet de visualiser l'effet de ce paramètre sur la superficie de l'estimation du domaine vital (Figure 45). Cette estimation peut superposer les estimateurs activés lors de la configuration comme le MCP, le KDE, le R-LoCoH ou le K-LoCoH.

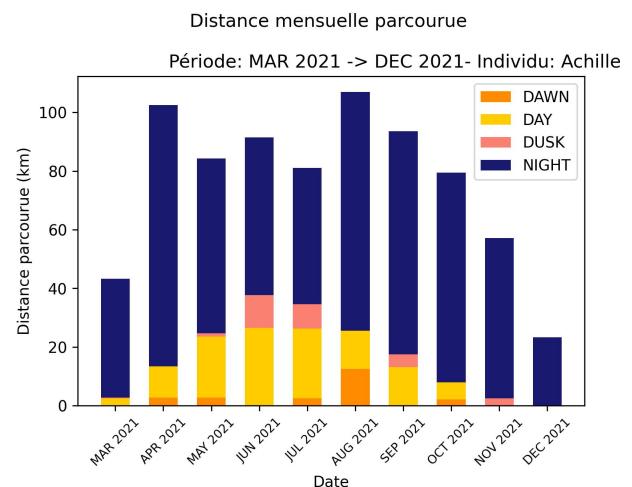


Figure 43: Distance mensuelle parcourue avec déclinaison selon la période du jour

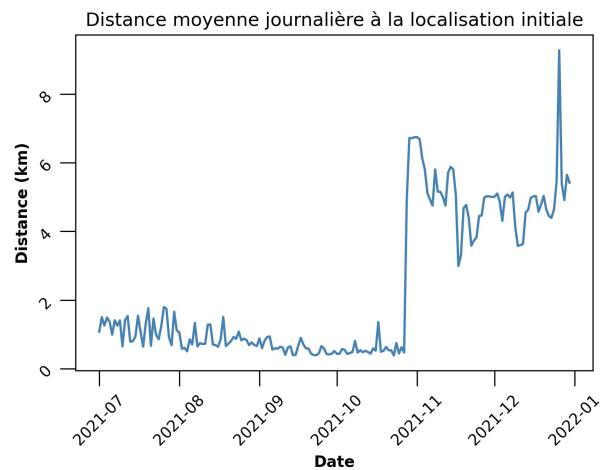


Figure 44: Distance relative au point initial

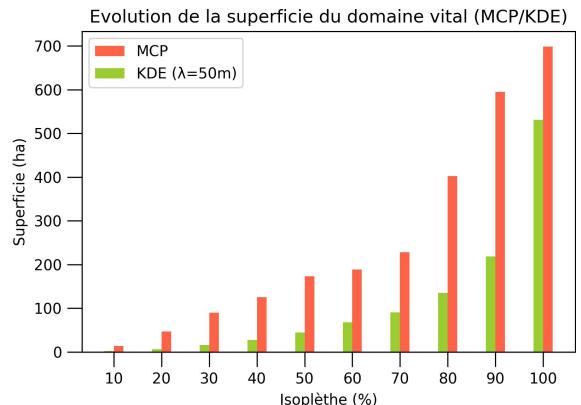


Figure 45: Comparaison de la superficie d'un MCP et d'un KDE pour différentes valeurs

La superposition de plusieurs estimateurs permet également de visualiser les écarts entre valeurs données par les différentes méthodes. Il est à préciser que pour le MCP et le KDE, les estimations pour chaque valeur (de 10 à 100%, par pas de 10%) peuvent être calculées sans augmentation significative du temps de calcul par rapport à un seule valeur.

4.3.4. Utilisation de l'habitat

Ce graphique est disponible si la fonctionnalité de l'interaction a été activée. Le graphique représente alors un diagramme permettant de visualiser la répartition relative des localisations au sein des différentes classes présentes dans le fichier d'interaction (Figure 46). Pour des raisons de lisibilité, les classes dont la fréquence est inférieure à 2% sont aggrégées.

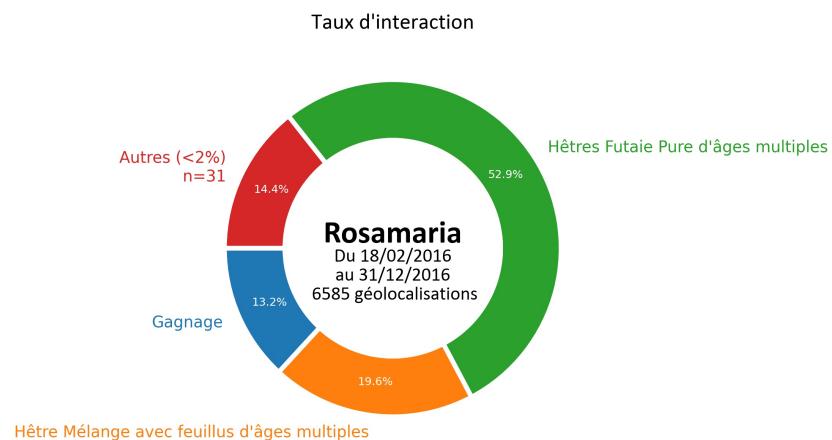


Figure 46 : Graphique de distribution des localisations au sein des habitats

Cette utilisation de l'habitat peut également être représentée de manière à confronter les classes en termes de surface utilisée et de surface disponible afin de mettre en évidence des habitats plus attractifs pour l'animal (Figure 47).

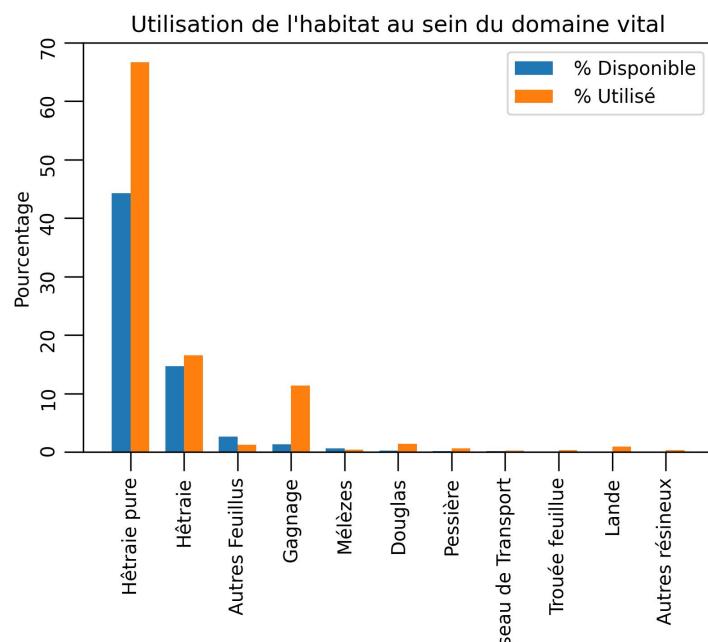


Figure 47 : Confrontation habitat disponible / utilisé au sein du domaine vital

V. Discussion

5.1. Evaluation de l'atteinte des objectifs

5.1.1. Opérationnalité dans un contexte professionnel

Depuis le mois d'avril 2022, TracksMining est utilisé pour générer les cartes interactives qui sont transmises mensuellement aux agents DNF des triages sur lesquels se déroulent des suivis d'animaux par télémétrie.

Au début de l'utilisation du programme, le travail de production des cartes interactives actualisées pour cinq animaux prenait initialement une heure environ, en raison de la préparation manuelle des données. L'ajout d'une classe de traitement supplémentaire pour effectuer ce travail de nettoyage a permis de raccourcir l'ensemble du processus à quelques minutes, ce qui engendre un gain de temps significatif.

5.1.2. Intégration des traitements élémentaires

Au terme de la première phase de développement de l'application, plusieurs fonctionnalités sont disponibles:

- Application de différents filtres sur les données.
- Détection et suppression des *outliers*.
- Affichage des géolocalisations, du centroïde et des trajectoires.
- Carte de chaleur
- Animation temporelle
- Polygone Minimum Convexe
- Kernels de Densité
- LoCoH (type K et type R)
- Grid Cell Method
- Interaction avec l'habitat
- Génération de graphiques

L'ensemble de ces fonctionnalités est détaillé dans le chapitre 4.2.

5.1.3. Appréciation du logiciel par des utilisateurs potentiels

5.1.3.1. Présentation du logiciel dans le cadre du projet Nassonia

Le développement du logiciel s'est articulé autour d'un projet professionnel dont l'objectif était de synthétiser les données relatives aux déplacements de plusieurs biches équipées de colliers GPS sur le massif forestier de Saint-Michel-Freyr. L'analyse devait servir à la rédaction du futur plan d'aménagement de Nassonia afin d'intégrer une composante de gestion spécifique à la grande faune. C'est dans le cadre de ce projet que le 31 mars 2022, le logiciel fut présenté pour la première fois à des utilisateurs externes.

La mise à disposition d'un parcellaire forestier détaillé a notamment permis de mettre en œuvre la fonctionnalité relative à l'interaction. Celle-ci a permis d'illustrer la manière dont les animaux étudiés utilisaient les différentes unités de gestion sur base d'une composante spatiale et temporelle.

Grâce aux nombreuses données disponibles accumulées depuis plusieurs années, des exemples bien spécifiques furent préparés pour illustrer les déplacements des biches marquées. Cette première séance de présentation de l'outil fut particulièrement productive. En effet, les données présentées faisaient appel à des animaux déjà relativement bien connus des agents forestiers. Leurs connaissances des comportements permettait donc une forme de validation des données.

Les retours sur cette première version du logiciel semblaient très positifs et ont permis de récolter directement de nombreuses critiques constructives. Les principales suggestions formulées étaient les suivantes :

- Permettre un export de la fenêtre de carte interactive vers une image de type .jpg ou .pdf dans le but de pouvoir les intégrer dans des rapports papier.
- Pouvoir exporter les trajectoires sous format .shp.
- Pouvoir scinder l'estimation du Grid Cell en une phase diurne et nocturne.
- Avoir la possibilité de subdiviser les données des géolocalisations selon deux modalités (jour/nuit) plutôt que quatre, pour ne pas tenir compte de l'aube et du crépuscule.
- Augmenter la taille de la police utilisée pour les *popups* et les *tooltips*.
- Pouvoir ajouter une légende qui illustre les valeurs du gradient d'interaction.
- Prendre en compte la disponibilité de l'habitat (surface occupée par une classe d'occupation par rapport à la surface totale du domaine vital).
- Mettre en place une technique permettant d'identifier des barrières éventuelles aux déplacements des individus.

Suite à cette réunion et la mise en œuvre d'une partie des corrections souhaitées, l'ensemble des cartographies interactives reprenant les données annuelles à partir de 2010 pour chaque biche furent transmises aux agents.

5.1.3.2. *Réalisation d'une enquête complémentaire*

Au terme de la phase de développement fin juin, une enquête en ligne fut envoyée à de nouveaux agents forestiers ainsi qu'à quelques autres personnes externes au SPW (34 personnes contactées). Cette démarche poursuivait un double objectif. D'une part, mieux cerner la plus-value que pouvait apporter un tel outil de visualisation et d'analyse des données de télémétrie dans le travail quotidien des personnes sondées. D'autre part, récolter de nouveaux avis sur la cartographie interactive et d'éventuelles améliorations à mettre en place.

Le formulaire est renseigné en annexe de ce travail. L'enquête n'a pas eu le taux de participation attendu (n=5) mais les résultats ont toutefois mis en évidence les éléments suivants :

- La carte interactive apporte une plus-value par rapport aux cartes statiques pdf.
- La plupart des personnes sont intéressées par une mise à disposition de l'outil.
- La carte est perçue comme esthétique et globalement intuitive.

- La carte ne semble pas poser de problèmes techniques pour l'affichage.
- La majorité des répondants souhaitent qu'il soit possible de réaliser un suivi simultané de plusieurs individus.
- La barre de défilement de l'animation temporelle ne semble pas assez pratique pour certains utilisateurs.
- Certains utilisateurs souhaitent pouvoir isoler sur des couches séparées les localisations selon la période diurne, nocturne ou crépusculaire.

L'organisation d'une deuxième séance de présentation directe de l'outil dans les prochains mois serait probablement plus productive pour récolter davantage d'avis.

5.1.4. Validation des résultats

Afin de vérifier la cohérence des résultats produits par TracksMining, ces derniers furent confrontés avec les résultats obtenus par ADEHabitat (Calenge, 2007), un module R couramment utilisé par la communauté scientifique.

La cohérence des résultats sur un même jeu de données ($n=1000$ données) et sur les mêmes paramètres a pu être vérifiée pour l'ensemble des estimateurs de l'habitat. Les deux programmes ont abouti à des résultats identiques. Quelques précisions sont cependant intéressantes à mentionner.

Dans le cas du MCP, cette validation a mis en évidence qu'il était indispensable de préciser le nombre de points inclus et écartés dans chaque estimateur lorsqu'un arrondi est nécessaire. En effet la surface de cet estimateur peut varier de manière significative avec une différence d'une seule localisation. Afin d'éviter un problème de non-reproductibilité des résultats, le nombre de géolocalisations effectivement incluses dans chaque estimateur est consultable au sein du fichier « metrics ». Dans son algorithme, TracksMining réalise un arrondi à l'entier le plus proche.

Le calcul du KDE a présenté des résultats très similaires entre les deux applications. La superposition des zones coïncide mais une légère différence subsiste (Figure 48). Cette différence est liée à un lissage de la forme finale dans ADEHabitat que ne réalise pas TracksMining. L'effet peut engendrer une différence moyenne de l'ordre de 1 à 3% sur la surface estimée qui dépend de la résolution, de la forme et de l'étendue du domaine vital.

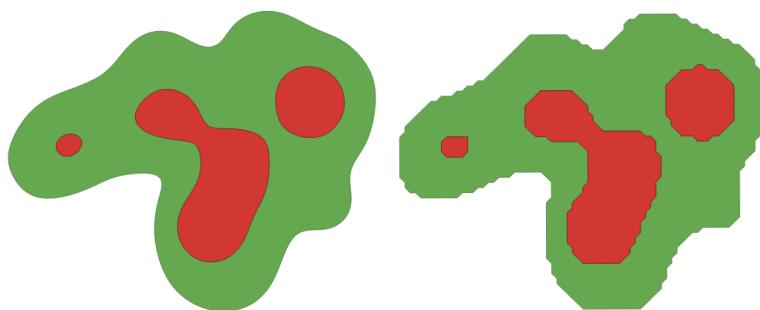


Figure 48 : Kernel de densité par ADEHabitat (gauche) et TracksMining (droite)

Les estimations du domaine vital par K-Locoh aboutissent à des résultats identiques. La seule différence concerne la représentation des polygones constitutifs dans l'estimateur final. Ceux-ci sont préservés dans ADEHabitat et fusionnés dans TracksMining (Figure 49).

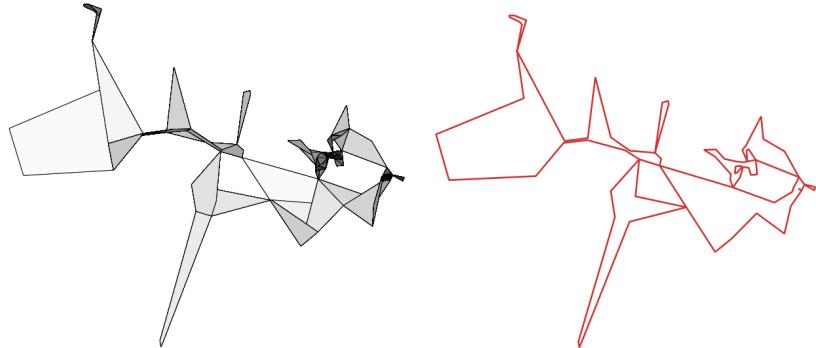


Figure 49 : K-LoCoH par ADEHabitat (gauche) et TracksMining (droite)

Il reste toutefois utopique de penser que le programme est à ce stade complètement exempt d'erreurs. Cependant, celles-ci seront identifiées et corrigées au fur et à mesure de l'utilisation de l'outil et d'un processus de validation plus approfondi.

5.1.5. Validation des performances

Le temps d'exécution moyen fut calculé pour le traitement de 3 jeux de données. La performance de vitesse est fournie sur base d'une fourchette comprise entre une valeur minimum et maximum pour un ensemble de paramètres identiques. La valeur minimum prend en compte le temps nécessaire pour traiter le jeu de données avec les classes essentielles, tandis que la valeur maximum rajoute l'ensemble des classes optionnelles (Table 6).

Table 6 : Performances absolues (secondes) de l'application – (* = classe optionnelle)

Processus	1000 données	5.000 données	10.000 données
Scanner les données	0,041	0,092	0,187
Lire les données	0,014	0,028	0,049
Nettoyer les données	0,012	0,027	0,053
Convertir les coordonnées	0,129	0,677	1,542
Gérer les <i>outliers</i>	0,058	0,191	0,336
Calculer les nouvelles variables	0,895	3,824	8,342
Calculer les entités géométriques	0,370	1,697	4,176
Créer la carte interactive	1,389	5,210	11,262
*Calculer l'interaction (x1000 entités)	2,296	11,296	46,790
*Calculer le MCP (x10 isoplèthes)	0,122	0,368	0,722
*Calculer le KDE (x10 isoplèthes)	0,429	0,732	1,115
*Calculer le Grid Cell (L=250m)	0,072	0,147	0,347
*Calculer un K-LoCoH (k=10)	0,443	10,174	39,005
*Créer les graphiques (x5 graphiques)	1,976	5,831	17,318
*Exporter les données	4,631	13,770	28,338
Temps minimum (secondes)	3	12	26
Temps maximum (secondes)	14	55	160

Les données sont présentées à titre indicatif. En effet, ces dernières vont être modulées de manière plus ou moins significative par une multitude de facteurs : emprise spatiale des données, valeur de certains paramètres, puissance de la machine utilisée (Table 7), etc. Les paramètres les plus influents sur le temps de traitement sont essentiellement la résolution spatiale souhaitée, le nombre d'entités pour le calcul de l'interaction, etc.

Table 7 : Caractéristiques de l'ordinateur utilisé pour le calcul des performances

Modèle	MacBook Air (13-inch, early 2015)
Système	macOS Mojave version 10.14.6
Processeur	1,6 GHz Intel Core i5
Mémoire	8 Go 1600MHz

La validation des performances s'est également effectuée de manière comparative avec ADEHabitat (Calenge, 2007). La comparaison a porté sur les 3 estimateurs de l'habitat précédemment testés au point 5.1.4, sur base de la vitesse d'exécution médiane obtenue sur 10 itérations (Table 8). Les vitesses d'exécution du calcul du MCP et KDE ont donné lieu à des performances d'ordre similaire. En revanche, TracksMining s'est révélé bien plus performant pour le calcul de l'estimateur du domaine vital par K-LoCoH. En effet sur un jeu de données de 1000 localisations, le processus a été réalisé en moins de 0,5 secondes par TracksMining contre plus de 2 minutes pour ADEHabitat.

Table 8 : Comparaison de la vitesse d'exécution médiane (secondes) pour 3 estimateurs

Processus	ADEHabitat	TracksMining
MCP	0,035	0,016
KDE	0,091	0,111
K-LoCoH	+ de 2 min	0,461

Il convient toutefois de prendre un certain recul par rapport à une comparaison directe des performances entre les deux programmes. D'une part, ils s'exécutent dans des environnements différents (R et Python). D'autre part, le contexte d'appel de la fonction est important à considérer. En effet, TracksMining agit comme un flux d'exécution global au sein duquel chaque estimateur augmente ses performances en accédant à des ressources déjà calculées en amont, disponibles au sein de l'objet Logger.

Plusieurs stratégies sont mises en place pour limiter la survenue de lenteurs dans l'exécution. Pour explorer les données, l'utilisateur a notamment la possibilité d'activer un ajustement automatique. Cette option donne la liberté au programme d'adapter certains paramètres pour éviter un temps de calcul trop important. Par exemple, une maille de résolution trop petite par rapport à l'emprise spatiale du jeu de données pourra être automatiquement agrandie.

Ces résultats mettent en évidence plusieurs points importants. Premièrement, TracksMining ne présente pas de difficultés à traiter des jeux de données de taille conséquente. En effet, 10.000 données représentent plus d'une année de suivi si l'on considère une fréquence d'une localisation GPS par heure. Le traitement complet peut alors être effectué en moins de 3 minutes.

Ensuite, la comparaison avec un logiciel de référence reconnu au sein de la communauté scientifique a permis de valider l'intérêt de TracksMining au niveau de ses performances de calcul, particulièrement pour l'estimateur du LoCoH. Il convient aussi de souligner le fait que l'interface utilisateur évite de consacrer du temps à rédiger des scripts et à modifier la valeur des paramètres au sein de ceux-ci.

Enfin, le temps de travail nécessaire avec l'augmentation de la taille du jeu de données évolue différemment selon les classes de traitement. Il sera donc essentiel d'améliorer les performances de calcul des classes présentant une croissance exponentielle pour garantir un processus fluide sur des jeux de données plus conséquents. Un profilage complet permettra d'analyser la consommation du programme en ressources informatiques. Ce travail servira à identifier les fonctions les plus chronophages et définir des priorités en termes d'optimisation du code.

5.1.6. Extrapolation à plusieurs modèles biologiques

La calibration et le développement du logiciel ont été réalisés sur des données collectées par le DEMNA, relatives à des suivis de cervidés marqués dans le massif forestier de Saint-Michel-Freyr. Toutefois, l'ensemble des fonctions de l'application a été testé sur de nouveaux jeux de données, relatifs à des animaux ou des espèces qui n'avaient pas été utilisées dans le processus de calibration (sanglier, blaireaux, rats-laveurs).

Ce dernier a permis d'identifier quelques faiblesses dans les traitements. Par exemple, l'utilisation du logiciel sur le suivi d'un blaireau a mis en évidence que l'animal générait uniquement des données nocturnes ou crépusculaires, ce qui perturbait la fluidité de l'animation temporelle.

L'application fut également testée sur d'autres jeux de données d'animaux situés dans des contextes géographiques différents (données issues de MoveBank, 2022). Les résultats se sont avérés cohérents, mais ont fait apparaître qu'une optimisation était nécessaire lorsque ces espèces utilisent un très large territoire afin d'accélérer la vitesse du processus de calcul.

L'analyse du déplacement de cigognes noires (données fournies par Gérard Jadoul, 2022) a démontré qu'il subsiste un gros challenge pour l'extrapolation aux espèces migratrices. En effet, les grandes distances parcourues par les individus et l'éloignement des clusters d'activité des animaux implique encore d'adapter les algorithmes de calcul ainsi que la représentation des éléments sur les cartes.

En contrepartie, cette démarche a permis de conforter l'idée que moyennant un travail de généralisation plus approfondi, ce programme pouvait théoriquement s'appliquer à de nombreuses espèces d'animaux. Ces modifications feront partie de la suite du développement de l'application.

5.1.7. Structuration cohérente et logique

Dans les premières étapes de sa phase de développement, le projet fut confronté à certaines difficultés informatiques : enchevêtrement de structures conditionnelles complexes, conflits entre noms de variables, encombrement de l'espace mémoire, répétitions de code, etc.

A ce stade, chaque nouvelle fonctionnalité implémentée dans le programme s'accompagnait de comportements inattendus et de résultats erronés. Ceci induisait une lenteur dans l'exécution mais aussi dans le bon déroulement du projet. Etant donné que le logiciel a été imaginé dans une perspective d'amélioration continue, il était impératif de pouvoir le maintenir facilement au fur et à mesure de ses extensions et de le rendre compréhensible pour d'autres développeurs. C'est suite à ce constat qu'une restructuration complète du programme s'est mise en place, en adoptant la programmation orientée objet (cfr chapitre 3.5.3) et l'architecture logicielle de type Modèle-Vue-Contrôleur (cfr chapitre 3.5.5).

L'ensemble du code tend à respecter les normes de la PEP 8 (Python Extension Proposal – Van Rossum et al., 2001). Cet ensemble de bonnes pratiques spécifique à Python permet de rendre le code clair, propre et structuré à travers des conventions d'écriture bien définies : taille des indentations et des espaces, règles d'import de modules, encodage, nombre de caractères par ligne, commentaires et annotations en anglais, présence de « *docstring* » explicative pour chaque fonction. Ce soin apporté à la forme trouve tout son sens avec l'idée que d'autres personnes pourraient potentiellement continuer à travailler sur le développement de l'outil.

5.1.8. Mise à disposition d'une interface utilisateur

L'application est fonctionnelle à partir d'une interface accessible par un navigateur via un serveur local. L'accent a été mis sur la fonctionnalité de l'interface mais son design reste pour l'instant peu attractif. Le choix d'une architecture logicielle de type Modèle-Vue-Contrôleur s'est avéré judicieux puisque l'interface est complètement indépendante du modèle. Par conséquent, il est possible de retravailler son apparence sans risquer de modifier le comportement de l'application. Une nouvelle interface est en cours de réalisation (Figure 50).

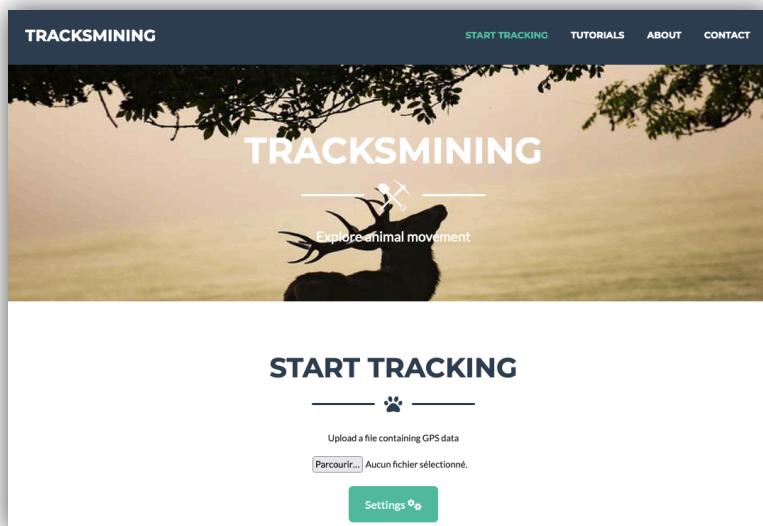


Figure 50 : Aperçu de la nouvelle interface utilisateur en cours de développement

En accord avec les promoteurs, il fut convenu que la mise en ligne du logiciel ne ferait pas partie du travail de fin d'études. D'une part, parce que le projet devra encore s'améliorer avant d'être rendu public. Ensuite parce que la mise en ligne fait appel à des compétences qui relèvent d'un informaticien professionnel, notamment pour résoudre des questions de sécurité informatique. Il conviendra aussi d'éclaircir les modalités d'accès à l'application ainsi que celles liées à l'hébergement sur un serveur disposant d'une puissance de calcul suffisante.

Une solution alternative sera de diffuser le code sous forme d'un module distribué à travers le Python Package index, un dépôt de modules en libre accès (<https://pypi.org/>). Les algorithmes de TracksMining pourraient également être partagés sur la plateforme MoveApp (<https://www.moveapps.org>) puisqu'une possibilité de développement collaboratif au moyen de Python est prochainement annoncée sur le site internet.

5.2. Quelle plus-value pour TracksMining ?

Comme évoqué dans la synthèse bibliographique, de nombreux logiciels sont déjà disponibles pour le traitement des données de télémétrie animale. Dès lors il est légitime de s'interroger sur l'intérêt de développer une nouvelle application informatique plutôt que de greffer les besoins du DEMNA sur un programme déjà existant et éprouvé. Ce choix peut se justifier pour plusieurs raisons.

5.2.1. Minimiser les dépendances

Aussi pratiques peuvent s'avérer les solutions externes déjà disponibles, le développement d'un outil de travail sur mesure rend la cellule Faune indépendante de toute entité extérieure. Cette dépendance peut se révéler problématique d'un point de vue technique : panne sur le serveur, abandon de maintenance du site ou du programme, etc. Ce cas de figure fut déjà rencontré par le passé au sein de notre service (Licoppe, communication personnelle).

La création d'un programme spécifique permet donc de garder la main sur l'intégralité du programme et d'y intégrer les traitements spécifiques aux besoins de nos études. Gupte et al. (2022) voient d'ailleurs la diversité croissante des études et des jeux de données comme une incitation pour les utilisateurs à développer leurs propres systèmes d'analyse. Bien que certaines applications en ligne (ex : ZoaTrack) permettent une transparence totale sur le code source, l'accès à celui-ci est généralement compliqué (Dwyer et al., 2015). La marge de manœuvre en termes de personnalisation du contenu des solutions existantes reste donc inévitablement limitée, tant sur le fond que sur la forme.

La sécurité des données est un autre point à ne pas négliger, particulièrement dans le cas d'espèces protégées ou présentant un risque de braconnage. TracksMining permet ainsi de travailler localement sur des données sans les faire transiter ni les stocker sur un serveur externe à l'organisation.

5.2.2. Mettre à disposition publique une solution complémentaire

Bien que TracksMining soit initialement développé pour répondre à des besoins professionnels du DEMNA, il existe une volonté de partager cet outil avec la communauté scientifique. Malgré la diversité de l'offre, les recherches bibliographiques semblent indiquer qu'aucun module Python ne soit encore disponible pour traiter spécifiquement des données de télémétrie animale. En s'appuyant sur des modules particulièrement performants tels que Scikit-Learn (Pedregosa et al., 2011), ou MovingPandas (Graser, 2019), TracksMining pourrait ainsi constituer une alternative intéressante aux programmes équivalents déjà disponibles sur R. Ceci est attesté par les performances de vitesse obtenues pour le calcul de certains estimateurs, tels que le LoCoH (cfr chapitre 5.1.5.).

5.2.3. Proposer une approche globale

La logique de TracksMining consiste à proposer une chaîne d'analyse globale visant à intégrer un maximum des traitements correspondants aux besoins les plus fréquents de la communauté scientifique (cfr chapitre 2.3). Cette approche permet d'éviter tant que possible qu'un utilisateur doive recourir à des logiciels complémentaires (ex : Fisher et al., 2020). En effet, les procédures d'import et d'export de données vers des logiciels différents sont chronophages et constituent une source d'erreur et de réPLICATION des données (Urbano et al., 2010). De plus, l'application a pour objectif de pouvoir se généraliser à une multitude de modèles biologiques. Dwyer et al. (2015) avaient mis en évidence que les algorithmes sous-jacents de certains logiciels étaient axés sur un groupe d'espèces en particulier, ce qui limitait leur utilisation par une frange plus large de la communauté scientifique.

5.2.4. Créer un logiciel accessible

Le programme se veut simple à utiliser pour un utilisateur ne disposant d'aucune compétence en programmation. En effet, l'activation des fonctionnalités et la définition des paramètres se réalisent dans une interface externe au code source. En revanche, des utilisateurs plus chevronnés peuvent importer directement des classes de traitement ou des méthodes spécifiques pour les intégrer dans leurs propres scripts Python.

5.2.5. Favoriser la communication

Au-delà des fonctions d'exploration et d'analyse, TracksMining a pour ambition de vulgariser l'information relayée par les cartes. Par exemple, des *popups* expliquent à quoi correspondent les différentes couches d'informations. La transmission de ces cartes doit être facile, d'où l'intérêt de produire des fichiers légers et embarqués pouvant se transmettre en une pièce jointe ou s'insérer aisément sur le site internet. Ces derniers évitent la connexion à un site tiers pour pouvoir être consultés.

Enfin, la diffusion de données de télémétrie peut être utile à des fins de communication auprès du grand public (Coyne & Godley, 2005). Cette affirmation prend sens avec le fait que le DEMNA est régulièrement amené à communiquer avec la presse ou à mettre à jour son site internet. Le format de cartes interactives s'insère parfaitement dans un « *iframe* », ce qui en rend la publication extrêmement simple.

VI. Conclusions et perspectives

6.1. Conclusions

Ce travail de fin d'études visait à initier le développement d'une application destinée à l'analyse et à l'exploration des données d'animaux suivis par balises GPS. Le développement d'une telle solution technique répondait à une volonté d'automatiser un travail récurrent et chronophage, tout en améliorant la communication de ces données dynamiques et complexes.

La première version de TracksMining a atteint la plupart de ses objectifs en étant capable de traiter automatiquement et rapidement un jeu de données de taille conséquente. Le programme en extrait alors de nombreuses informations et peut les restituer de manière visuelle au travers d'une cartographie interactive et de graphiques d'analyse.

Les retours des premiers utilisateurs semblent indiquer que la mise en place de cet outil améliore fortement le processus de communication des données de télémétrie animale entre le DEMNA et ses partenaires. Le logiciel a également permis de réduire significativement le temps de travail destiné à la production des cartes. Initialement compris entre 15 et 30 minutes en moyenne, le temps de travail pour production d'une carte est désormais de moins d'une minute.

Moyennant un travail de validation plus approfondi ainsi qu'une optimisation des fonctionnalités et du paramétrage, cette application pourra être mise à disposition d'utilisateurs externes. TracksMining intègre de nombreuses possibilités de traitement et d'interprétation qui répondent tant que possible aux besoins de la communauté scientifique. La centralisation des fonctionnalités, les bonnes performances et la standardisation des traitements pourraient rendre l'application attractive dans le domaine de l'écologie spatiale.

Pour un public plus large, la présence d'une interface simple et intuitive permet de s'affranchir de toute compétence en traitement SIG ou en programmation. Les utilisateurs avertis garderont cependant la possibilité d'extraire des fonctions spécifiques qui s'intégreront dans des scripts plus complexes. La forme finale que prendra la mise à disposition publique de l'application reste toutefois à définir. Celle-ci sera conditionnée par différentes contraintes techniques qui se précisieront dans les phases ultérieures du développement de ce projet.

Disposant d'une grande marge de progression d'un point de vue technique et fonctionnel, TracksMining continuera son développement au cours des prochaines années. Le logiciel tentera d'implémenter progressivement de nouvelles fonctionnalités et d'optimiser celles qui sont déjà existantes. Ces améliorations pourront être dictées par les besoins du contexte professionnel ou suggérées par de nouvelles avancées scientifiques et technologiques. Une attention particulière sera également portée sur l'amélioration de l'interactivité mais aussi de l'aspect esthétique de la carte et de l'interface utilisateur.

6.2. Perspectives

De nombreuses perspectives d'amélioration du programme sont possibles, tant dans des aspects fonctionnels que techniques. Quelques-unes de ces perspectives sont détaillées ci-dessous.

6.2.1. Permettre le suivi simultané de plusieurs individus

Suite à des suggestions d'utilisateurs, il est apparu que la possibilité de visualiser les déplacements et les domaines vitaux de plusieurs individus sur une même carte serait particulièrement intéressante. Ce suivi simultané permettrait entre autres de vérifier les superpositions des domaines vitaux en fonction des périodes de l'année. Cette fonctionnalité nécessitera de solutionner une multitude de problèmes sous-jacents : augmentation significative du poids de la cartographie, multiplication importante des couches, incohérences dans la barre de gestion de l'animation temporelle, etc. Ceci fera partie des évolutions prioritaires de l'application.

6.2.2. Adapter le programme aux données d'observations

Au vu des bases de données conséquentes disponibles au DEMNA, il semble pertinent de permettre l'utilisation de données d'observation ponctuelles qui ne sont pas liées à un individu unique suivi par télémétrie mais relatives à des observations aléatoires et opportunistes (Figure 51). Certains estimateurs de l'habitat s'adaptent bien sur des données et permettent de quantifier la distribution ou l'expansion d'espèces au fil du temps (ex : Boscherini et al., 2020). Plusieurs classes de traitement seraient directement applicables, tandis que d'autres nécessiteraient des adaptations plus approfondies pour garantir des résultats cohérents.

6.2.3. Elargir les possibilités d'analyse

Le module « *Scikit-Learn* » offre de nombreuses perspectives intéressantes pour l'évolution du programme et permettrait la mise en place de techniques avancées de classification ou de prédiction. Par exemple, les données de géolocalisations pourraient être regroupées sur base des valeurs des différentes variables afin de tenter de les classifier selon des comportements spécifiques (repos, déplacements, prédatation, etc.). A cet usage, la méthode de partition au moyen de l'algorithme des « *K-Means* » semble particulièrement indiquée (Schwager et al., 2007).

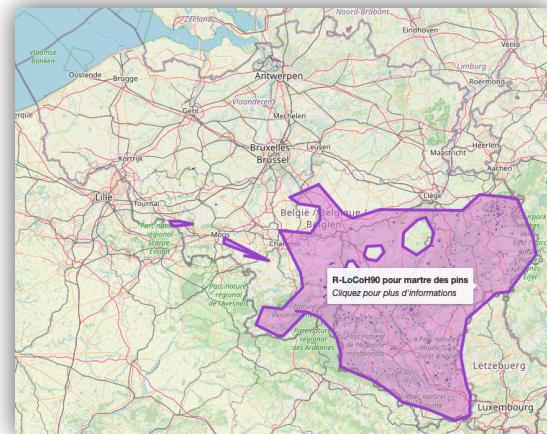


Figure 51 : R-LoCoH ($r=15\text{km}$) appliqué aux observations ponctuelles de la martre des pins entre 2001 et 2021

Comme évoqué dans la bibliographie, de nombreuses possibilités d'analyse graphiques sont possibles pour visualiser les données. Il serait utile de créer l'ensemble des classes de traitement qui permettent de couvrir les représentations graphiques les plus couramment rencontrées dans la littérature afin de disposer d'un catalogue au sein duquel l'utilisateur pourrait définir des styles personnalisés (police, couleur, etc..) .

Les informations relatives à l'analyse du mouvement mériteraient d'être davantage exploitées. De nombreuses variables pourraient être calculées pour caractériser chaque déplacement, ce qui ouvrirait la voie à une multitude d'analyses (cfr Gurarie et al., 2009 ; Edelhoff et al., 2016 ; Morelle et al. 2017).

Enfin, dans le cadre des calculs d'interaction, il serait intéressant de permettre l'interopérabilité de l'application avec des couches externes disponibles en ligne (cfr Salvatori et al., 2022)

6.2.4. Améliorer la détection automatique des localisations erronées

Les géolocalisations anormales peuvent affecter de manière significative les performances et la qualité des estimations. Par conséquent, il convient de les détecter de la manière la plus efficace possible. Actuellement, TracksMining permet une détection des *outliers* sur base d'une méthode personnelle qui reste à valider et devra probablement se compléter ou se substituer par d'autres approches validées pour les suivis d'animaux (Bjørneraa et al., 2010 ; Hirakawa et al., 2022).

6.2.5. Implémenter de nouveaux estimateurs de l'habitat

La superposition des différents estimateurs du domaine vital sur la carte interactive constitue l'un des points d'intérêt de TracksMining. Il pourrait donc être intéressant de proposer de nouvelles méthodes d'estimation du domaine vital comme par exemple : le modèles des ponts Browniens (Horne et al., 2007), le « *Time-LoCoH* » (Lyons et al., 2013) ou encore le « *Biased Random Walk Model* » (Benhamou, 2011).

Pour les estimateurs du domaine vital déjà implémentés, il serait judicieux de pouvoir proposer des déclinaisons plus avancées telles que la possibilité de les décomposer selon la saison afin de mieux prendre en compte l'aspect temporel dans leurs calcul (Dettki & Ericsson, 2008). Au niveau des KDE, il serait intéressant de proposer de nouvelles méthodes de détermination du facteur de lissage le plus adapté, actuellement uniquement disponible avec la formule de Bowman & Foster (1993).

6.2.6. Augmenter les performances générales

Si les performances du programme sont satisfaisantes pour son utilisation actuelle, il est essentiel de garantir une exécution rapide pour rendre le programme agréable à l'utilisation et attractif au regard des autres solutions disponibles. Cette optimisation passe notamment par une augmentation de la vitesse d'exécution ainsi qu'une réduction de l'utilisation de la mémoire vive : vectorisation des traitements, « *multi-threading* », etc.

VII. Bibliographie

Abner, D. (2020). Home range size and movement of Northern Bobwhite in relation to cattle at the Blue Grass Army Depot, Madison County, Kentucky.

Acácio, M., Mullers, R. H., Franco, A., Willems, F. J., & Amar, A. (2021). Changes in surface water drive the movements of Shoebills. *Scientific Reports*, 11(1), 1-12

Alfred, R., Ahmad, A. H., Payne, J., Williams, C., Ambu, L. N., How, P. M., & Goossens, B. (2012). Home range and ranging behaviour of Bornean elephant (*Elephas maximus borneensis*) females. *PLoS One*, 7(2), e31400.

Aschekoug, E. T., Scott, C. A., & Van Vuren, D. H. (2004). A test of the Judas technique as a method for eradicating feral pigs. *Transactions of the Western Section of the Wildlife Society*, 40, 120-126.

Athreya, V., Navya, R., Punjabi, G. A., Linnell, J. D., Odden, M., Khetarpal, S., & Karanth, K. U. (2014). Movement and activity pattern of a collared tigress in a human-dominated landscape in central India. *Tropical Conservation Science*, 7(1), 75-86.

Baert, J. M., Stienen, E. W., Heylen, B. C., Kavelaars, M. M., Buijs, R. J., Shamoun-Baranes, J., ... & Müller, W. (2018). High-resolution GPS tracking reveals sex differences in migratory behaviour and stopover habitat use in the Lesser Black-backed Gull *Larus fuscus*. *Scientific reports*, 8(1), 1-11.

Bastille-Rousseau, G., Wall, J., Douglas-Hamilton, I., & Wittemyer, G. (2018). Optimizing the positioning of wildlife crossing structures using GPS telemetry. *Journal of Applied Ecology*, 55(4), 2055-2063.

Baxley, D. L., & Qualls, C. P. (2009). Black pine snake (*Pituophis melanoleucus lodingi*): Spatial ecology and associations between habitat use and prey dynamics. *Journal of Herpetology*, 43(2), 284-293.

Beauchamp, J. S., Hart, K. M., Cherkiss, M. S., & Mazzotti, F. J. (2018). Variation in home range size and patterns in adult female American crocodiles *Crocodylus acutus*. *Endangered Species Research*, 36, 161-171.

Benhamou, S. (2011). Dynamic approach to space and habitat use based on biased random bridges. *PLoS one*, 6(1), e14592.

Benhamou, S., & Cornélis, D. (2010). Incorporating movement behavior and barriers to improve kernel home range space use estimates. *The Journal of Wildlife Management*, 74(6), 1353-1360.

Benhamou, S., & Riotte-Lambert, L. (2012). Beyond the Utilization Distribution: Identifying home range areas that are intensively exploited or repeatedly visited. *Ecological Modelling*, 227, 112-116.

Bergman, C. M., Schaefer, J. A., & Luttich, S. N. (2000). Caribou movement as a correlated random walk. *Oecologia*, 123(3), 364-374.

Beyer, H. (2004). Hawth's analysis tools for ArcGIS. <http://www.spatialecology.com/htools>.

Bjørneraas, K., Van Moorter, B., Rolandsen, C. M., & Herfindal, I. (2010). Screening Global Positioning System Location Data for Errors Using Animal Movement Characteristics. *Journal of Wildlife Management*, 74(6), 1361.

Blakey, R. V., Siegel, R. B., Webb, E. B., Dillingham, C. P., Bauer, R. L., Johnson, M., & Kesler, D. C. (2019). Space use, forays, and habitat selection by California Spotted Owls (*Strix occidentalis occidentalis*) during the breeding season: New insights from high resolution GPS tracking. *Forest Ecology and Management*, 432, 912-922.

Börger, L., Franconi, N., De Michele, G., Gantz, A., Meschi, F., Manica, A., ... & Coulson, T. I. M. (2006). Effects of sampling regime on the mean and variance of home range size estimates. *Journal of Animal Ecology*, 1393-1405.

Boscherini, A., Mazza, G., Menchetti, M., Laurenzi, A., & Mori, E. (2020). Time is running out! Rapid range expansion of the invasive northern raccoon in central Italy. *Mammalia*, 84(1), 98-101.

Boucheny V. (2020). Apprendre la Programmation Orientée Objet avec le langage Python. ENI Editions.

Bowman, A. W., & Foster, P. J. (1993). Adaptive smoothing and density-based tests of multivariate normality. *Journal of the American Statistical Association*, 88(422), 529-537.

Boyle, S. A., Lourenço, W. C., Da Silva, L. R., & Smith, A. T. (2008). Home range estimates vary with sample size and methods. *Folia Primatologica*, 80(1), 33-42.

Bunnefeld, N., Börger, L., van Moorter, 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(2), 466-476.

Burgman, M. A., & Fox, J. C. (2003, February). Bias in species range estimates from minimum convex polygons: implications for conservation and options for improved planning. In *Animal Conservation Forum* (Vol. 6, No. 1, pp. 19-28). Cambridge University Press.

Cagnacci, F., Boitani, L., Powell, R. A., & Boyce, M. S. (2010). Animal ecology meets GPS-based radiotelemetry: a perfect storm of opportunities and challenges. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1550), 2157-2162.

Calenge, C. (2007). Exploring habitat selection by wildlife with adehabitat. *Journal of statistical software*, 22, 1-19.

Calenge, C. (2011). Home range estimation in R: the adehabitatHR package. *Office national de la classe et de la faune sauvage: Saint Benoist, Auffargis, France*.

Calenge, C., Dray, S., & Royer-Carenzi, M. (2009). The concept of animals' trajectories from a data analysis perspective. *Ecological informatics*, 4(1), 34-41.

Castellanos, A. (2011). Andean bear home ranges in the Intag region, Ecuador. *Ursus*, 22(1), 65-73.

Chassagneux, A., Calenge, C., Marchand, P., Richard, E., Guillaumat, E., Baubet, E., & Saïd, S. (2020). Should I stay or should I go? Determinants of immediate and delayed movement responses of female red deer (*Cervus elaphus*) to drive hunts. *Plos one*, 15(3), e0228865.

Cornélis, D., Benhamou, S., Janeau, G., Morellet, N., Ouedraogo, M., & De Visscher, M. N. (2011). Spatiotemporal dynamics of forage and water resources shape space use of West African savanna buffaloes. *Journal of Mammalogy*, 92(6), 1287-1297.

Coyne, M. S., & Godley, B. J. (2005). Satellite Tracking and Analysis Tool (STAT): an integrated system for archiving, analyzing and mapping animal tracking data. *Marine Ecology Progress Series*, 301, 1-7.

Crane, M., Silva, I., Marshall, B. M., & Strine, C. T. (2021). Lots of movement, little progress: A review of reptile home range literature. *PeerJ*, 9, e11742.

Cross, S. L., Tomlinson, S., Craig, M. D., & Bateman, P. W. (2020). The Time Local Convex Hull method as a tool for assessing responses of fauna to habitat restoration: a case study using the perentie (*Varanus giganteus*: Reptilia: Varanidae). *Australian Journal of Zoology*, 67(1), 27-37.

Cumming, G. S., & Cornélis, D. (2012). Quantitative comparison and selection of home range metrics for telemetry data. *Diversity and Distributions*, 18(11), 1057-1065.

Curry, E., & Grace, P. (2008). Flexible self-management using the model-view-controller pattern. *IEEE software*, 25(3), 84-90.

de Melo, L. B., Sabato, M. L., Magni, E. V., Young, R. J., & Coelho, C. M. (2006). Secret lives of maned wolves (*Chrysocyon brachyurus* Illiger 1815): as revealed by GPS tracking collars.

Deacon, F., & Smit, N. (2017). Spatial ecology and habitat use of giraffe (*Giraffa camelopardalis*) in South Africa. *Basic and Applied Ecology*, 21, 55-65.

Dellicour, S., Desmecht, D., Paternostre, J., Malengreaux, C., Licoppe, A., Gilbert, M., & Linden, A. Unravelling the dispersal dynamics and ecological drivers of the African swine fever outbreak in Belgium.

Dettki, H., & Ericsson, G. (2008). Screening radiolocation datasets for movement strategies with time series segmentation. *The Journal of Wildlife Management*, 72(2), 535-542.

Douglas-Hamilton, I., Krink, T., & Vollrath, F. (2005). Movements and corridors of African elephants in relation to protected areas. *Naturwissenschaften*, 92(4), 158-163.

Downs, J. A., & Horner, M. W. (2008). Effects of point pattern shape on home-range estimates. *The Journal of Wildlife Management*, 72(8), 1813-1818.

Downs, J. A., & Horner, M. W. (2009). A characteristic-hull based method for home range estimation. *Transactions in GIS*, 13(5-6), 527-537.

Duparc, A., Amblard T., Garel M., Marchand P., Perrin-Malterre C., Dubray D., Maillard D. & Loison, A. 2017. Réponse comportementale face au dérangement dans un espace fortement fréquenté par les randonneurs: Le cas d'une population de chamois dans le massif des Bauges. *Faune sauvage* n°316 : 20-25.

Dürr, S., & Ward, M. P. (2014). Roaming behaviour and home range estimation of domestic dogs in Aboriginal and Torres Strait Islander communities in northern Australia using four different methods. *Preventive veterinary medicine*, 117(2), 340-357.

Dwyer, R. G., Brooking, C., Brimblecombe, W., Campbell, H. A., Hunter, J., Watts, M., & Franklin, C. E. (2015). An open Web-based system for the analysis and sharing of animal tracking data. *Animal Biotelemetry*, 3(1), 1-11.

Edelhoff, H., Signer, J., & Balkenhol, N. (2016). Path segmentation for beginners: an overview of current methods for detecting changes in animal movement patterns. *Movement ecology*, 4(1), 1-21.

Enríquez-Mercado, I., Montiel-Ugalde, A., Aparicio, Á., GAONA-MURILLO, E., Butterfield, T., & Macip-Ríos, R. (2018). Population ecology and home range of the Mexican Rough-footed Mud Turtle (*Kinosternon hirtipes murrayi*) in Central Mexico. *Acta Herpetologica*, 13(2), 109-115.

Eraud C., Wolff A., Menard C., Mars B., Audouit Q. & Villers, A. 2015. Intérêt du suivi individuel pour la gestion du risque animalier: Le cas de l'outarde canepetière sur la plateforme aéroportuaire de Marseille-Provence. *Faune sauvage* n°323 : 10-16.

Eriksen, A., Wabakken, P., Zimmermann, B., Andreassen, H. P., Arnemo, J. M., Gundersen, H., ... & Storaas, T. (2009). Encounter frequencies between GPS-collared wolves (*Canis lupus*) and moose (*Alces alces*) in a Scandinavian wolf territory. *Ecological research*, 24(3), 547-557.

Evans, L. J., Davies, A. B., Goossens, B., & Asner, G. P. (2017). Riparian vegetation structure and the hunting behavior of adult estuarine crocodiles. *PLoS one*, 12(10), e0184804.

Fieberg, J., & Kochanny, C. O. (2005). Quantifying home-range overlap: The importance of the utilization distribution. *The Journal of Wildlife Management*, 69(4), 1346-1359.

Fisher, M., Muth, A., & Johnson, R. F. (2020). A long-term study of home range of Coachella fringe-toed lizards, *Uma inornata*. *Journal of Herpetology*, 54(2), 174-182.

Flanagan, S. E., Brown, M. B., Fennessy, J., & Bolger, D. T. (2016). Use of home range behaviour to assess establishment in translocated giraffes.

Getz, W. M., & Wilmers, C. C. (2004). A local nearest-neighbor convex-hull construction of home ranges and utilization distributions. *Ecography*, 27(4), 489-505.

Getz, W. M., Fortmann-Roe, S., Cross, P. C., Lyons, A. J., Ryan, S. J., & Wilmers, C. C. (2007). LoCoH: nonparameteric kernel methods for constructing home ranges and utilization distributions. *PloS one*, 2(2), e207.

Gillies, S., Bierbaum, A., Lautaportti, K., & Tonnhofer, O. (2007). Shapely: manipulation and analysis of geometric objects. *toblerity.org*.

Girard, I., Ouellet, J. P., Courtois, R., Dussault, C., & Breton, L. (2002). Effects of sampling effort based on GPS telemetry on home-range size estimations. *The Journal of wildlife management*, 1290-1300.

González-Gallina, A., Hidalgo-Mihart, M. G., Pérez-Garduza, F., Iglesias-Hernández, J. A., de Ita, A. O., Chacón-Hernández, A., & Vázquez-Zúñiga, O. (2018). Home range of a male jaguar spatially associated with the landfill of the city of Playa del Carmen, Mexico. *Mammalia*, 82(1), 54-61.

Graser, A. (2019). MovingPandas: Efficient Structures for Movement Data in Python. *GI_Forum – Journal of Geographic Information Science* 2019, 1-2019, 54-68.

Grinberg, M. (2018). *Flask web development: developing web applications with python*. "O'Reilly Media, Inc.".

Grueter, C.C., Li, D., Ren, B. *et al.* Choice of analytical method can have dramatic effects on primate home range estimates. *Primates* 50, 81 (2009). <https://doi.org/10.1007/s10329-008-0113-2>

Gupte, P. R., Beardsworth, C. E., Spiegel, O., Lourie, E., Toledo, S., Nathan, R., & Bijleveld, A. I. (2022). A guide to pre-processing high-throughput animal tracking data. *Journal of Animal Ecology*, 91(2), 287-307.

Gurarie, E., Andrews, R. D., & Laidre, K. L. (2009). A novel method for identifying behavioural changes in animal movement data. *Ecology letters*, 12(5), 395-408.

Harris, C. R., Millman, K. J., Van Der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., ... & Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825), 357-362.

Hart, K. M., Cherkiss, M. S., Smith, B. J., Mazzotti, F. J., Fujisaki, I., Snow, R. W., & Dorcas, M. E. (2015). Home range, habitat use, and movement patterns of non-native Burmese pythons in Everglades National Park, Florida, USA. *Animal Biotelemetry*, 3(1), 1-13.

Heidke, N., Morrison, J., & Morrison, M. (2008, April). Assessing the effectiveness of the model view controller architecture for creating web applications. In *Midwest instruction and computing symposium, Rapid City, SD*.

Henderson, T., Vernes, K., Körtner, G., & Rajaratnam, R. (2018). Using GPS technology to understand spatial and temporal activity of kangaroos in a peri-urban environment. *Animals*, 8(6), 97.

Hirakawa, H., Muramatsu, D., Gordo, M., Takii, A., & Izumiya, S. (2022). GAP method: a gap-based technique for outlier detection and its application to wildlife GPS data.

Hofman, M. P., Hayward, M. W., Heim, M., Marchand, P., Rolandsen, C. M., Mattisson, J., ... & Balkenhol, N. (2019). Right on track? Performance of satellite telemetry in terrestrial wildlife research. *PloS one*, 14(5), e0216223.

Hooge, P. N., Eichenlaub, W. M., & Solomon, E. K. (2001). Using GIS to analyze animal movements in the marine environment. *Spatial Processes and Management of Marine Populations. Alaska Sea Grant College Program, Anchorage Alaska*, 37-51.

Horne, J. S., Garton, E. O., Krone, S. M., & Lewis, J. S. (2007). Analyzing animal movements using Brownian bridges. *Ecology*, 88(9), 2354-2363.

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in science & engineering*, 9(03), 90-95.

Hunter, J., Brooking, C., Brimblecombe, W., Dwyer, R. G., Campbell, H. A., Watts, M. E., & Franklin, C. E. OzTrack—e-Infrastructure to support the Management, Analysis and Sharing of Animal Tracking Data.

Jarolímek, J., Vaněk, J., Ježek, M., Masner, J., & Stočes, M. (2014). The telemetric tracking of wild boar as a tool for field crops damage limitation. *Plant, Soil and Environment*, 60(9), 418-425.

Kabatha, P. (2018). An open source web GIS tool for analysis and visualization of elephant GPS telemetry data, alongside environmental and anthropogenic variables. *Master Thesis in Geographical Information Science*.

Kauhala, K., & Auttila, M. (2010). Estimating habitat selection of badgers—a test between different methods. *Folia Zoologica*, 59(1), 16-25.

Kays, R., Jansen, P. A., Knecht, E. M., Vohwinkel, R., & Wikelski, M. (2011). The effect of feeding time on dispersal of Virola seeds by toucans determined from GPS tracking and accelerometers. *Acta Oecologica*, 37(6), 625-631.

Kie, J. G., Baldwin, J. A., & Evans, C. J. (1996). CALHOME: a program for estimating animal home ranges. *Wildlife Society Bulletin*, 342-344.

Klein, F., Gendner, J. P., Storms, D., Hamann, J. L., Saïd, S., Michallet, J., & Pfaff, E. (2006). Le GPS et l'étude des ongulés sauvages. *Gibier Faune Sauvage*, 272, 31-38.

Kranstauber, B., Cameron, A., Weinzerl, R., Fountain, T., Tilak, S., Wikelski, M., & Kays, R. (2011). The Movebank data model for animal tracking. *Environmental Modelling & Software*, 26(6), 834-835.

Krone, O., Berger, A., & Schulte, R. (2009). Recording movement and activity pattern of a White-tailed Sea Eagle (*Haliaeetus albicilla*) by a GPS datalogger. *Journal of Ornithology*, 150(1), 273-280.

Larkin, R. P., & Halkin, D. (1994). A review of software packages for estimating animal home ranges.

Laver, P. (2005). ABODE. *Book ABODE. Virginia: Department of Fisheries and Wildlife Sciences, Virginia Tech.*

Laver, P. N., & Kelly, M. J. (2008). A critical review of home range studies. *The Journal of Wildlife Management*, 72(1), 290-298.

Lawson, E. J. G., & Rodgers, A. R. (1997). Differences in home-range size computed in commonly used software programs. *Wildlife Society Bulletin*, 721-729.

Leff, A., & Rayfield, J. T. (2001, September). Web-application development using the model/view/controller design pattern. In *Proceedings fifth ieee international enterprise distributed object computing conference* (pp. 118-127). IEEE.

Lesmeister, D. B., Gompper, M. E., & Millspaugh, J. J. (2009). Habitat selection and home range dynamics of eastern spotted skunks in the Ouachita Mountains, Arkansas, USA. *The Journal of Wildlife Management*, 73(1), 18-25.

Lichti, N. I., & Swihart, R. K. (2011). Estimating utilization distributions with kernel versus local convex hull methods. *The Journal of Wildlife Management*, 75(2), 413-422.

Licoppe, A., Lievens, J., & Verheyden, H. (2005). Mesure par télémétrie de l'impact du dérangement humain sur le cerf (*Cervus elaphus L.*) en parcours libre. *Parcs et Réserves*, 60(3), 10-21.

Loveridge, A. J., Valeix, M., Davidson, Z., Murindagomo, F., Fritz, H., & Macdonald, D. W. (2009). Changes in home range size of African lions in relation to pride size and prey biomass in a semi-arid savanna. *Ecography*, 32(6), 953-962.

Lyons, A. J., Turner, W. C., & Getz, W. M. (2013). Home range plus: a space-time characterization of movement over real landscapes. *Movement Ecology*, 1(1), 1-14.

Lyver, P. B., MacLeod, C. J., Ballard, G., Karl, B. J., Barton, K. J., Adams, J., ... & Wilson, P. R. (2011). Intra-seasonal variation in foraging behavior among Adélie penguins (*Pygoscelis adeliae*) breeding at Cape Hallett, Ross Sea, Antarctica. *Polar biology*, 34(1), 49-67.

Marchand, P., Itty, C., Duhayer, J., Garel, M., Bourgouin, G. & Dubray, D. 2015. Dérangements par le tourisme et la chasse : quelles réponses du mouflon ? Bilan de suivis par colliers GPS dans le massif du Caroux-Espinouse (34). *Faune sauvage* n°306 : 4-10.

McFerren, G., & Van Zyl, T. (2016). Geospatial data stream processing in Python using FOSS4G components. *ISPRS*.

McGregorA, H. W., LeggeB, S., PottsC, J., JonesA, M. E., & JohnsonA, C. N. Density and home range of feral cats in north-western Australia.

McKinney, W. (2012). *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython.* " O'Reilly Media, Inc.".

McLellan, B. N., & Hovey, F. W. (2001). Habitats selected by grizzly bears in a multiple use landscape. *The Journal of wildlife management*, 92-99.

Merrill, S. B., & Mech, L. D. (2003). The usefulness of GPS telemetry to study wolf circadian and social activity. *Wildlife Society Bulletin*, 947-960.

Michelot, T., Langrock, R., & Patterson, T. A. (2016). moveHMM: an R package for the statistical modelling of animal movement data using hidden Markov models. *Methods in Ecology and Evolution*, 7(11), 1308-1315.

Miller, G. J., Smith, L. L., Johnson, S. A., & Franz, R. (2012). Home range size and habitat selection in the Florida pine snake (*Pituophis melanoleucus mugitus*). *Copeia*, 2012(4), 706-713.

Mitchell, B. R. (2006). Comparison of programs for fixed kernel home range analysis. *Remotely Wild*, 21, 1-7.

Morelle, K., Bunnefeld, N., Lejeune, P., & Oswald, S. A. (2017). From animal tracks to fine-scale movement modes: a straightforward approach for identifying multiple spatial movement patterns. *Methods in Ecology and Evolution*, 8(11), 1488-1498.

Morelle, K., Lehaire, F., & Lejeune, P. (2014). Is wild boar heading towards movement ecology? A review of trends and gaps. *Wildlife biology*, 20(4), 196-205.

Mosnier, A., Boisjoly, D., Courtois, R., & OUELLET, J. P. (2008). Extensive predator space use can limit the efficacy of a control program. *The Journal of Wildlife Management*, 72(2), 483-491.

Ngene, S., Okello, M. M., Mukeka, J., Muya, S., Njumbi, S., & Isiche, J. (2017). Home range sizes and space use of African elephants (*Loxodonta africana*) in the Southern Kenya and Northern Tanzania borderland landscape.

Oleksy, R., Racey, P. A., & Jones, G. (2015). High-resolution GPS tracking reveals habitat selection and the potential for long-distance seed dispersal by Madagascan flying foxes *Pteropus rufus*. *Global Ecology and Conservation*, 3, 678-692.

Olsen, J. (2011). Home-range size and territorial calling of Southern Boobooks (*Ninox novaeseelandiae*) in adjacent territories. *Journal of Raptor Research*, 45(2), 136-142.

Otten, J. G., Hulbert, A. C., Berg, S. W., & Tamplin, J. W. (2021). Home Range, Site Fidelity, and Movement Patterns of the Wood Turtle (*Glyptemys insculpta*) at the Southwestern Edge of Its Range. *Chelonian Conservation and Biology*, 20(2), 231-241.

Pebsworth, P.A., Morgan, H.R. & Huffman, M.A. Evaluating home range techniques: use of Global Positioning System (GPS) collar data from chacma baboons. *Primates* 53, 345–355 (2012). <https://doi.org/10.1007/s10329-012-0307-5>

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 12, 2825-2830.

Peris, A., Closa, F., Marco, I., Acevedo, P., Barasona, J. A., & Casas-Díaz, E. (2020). Towards the comparison of home range estimators obtained from contrasting tracking regimes: the wild boar as a case study. *European journal of wildlife research*, 66(2), 1-10.

Pfeiffer, T., & Meyburg, B. U. (2015). GPS tracking of Red Kites (*Milvus milvus*) reveals fledgling number is negatively correlated with home range size. *Journal of Ornithology*, 156(4), 963-975.

Pillay, K. R., Streicher, J., & Downs, C. T. (2018). Home range and habitat use of feral cats in an urban mosaic in Pietermaritzburg, KwaZulu-Natal, South Africa. *Urban Ecosystems*, 21(5), 999-1009.

Pimley, E. R., Bearder, S. K., & Dixson, A. F. (2005). Home range analysis of *Perodicticus potto edwardsi* and *Sciurocheirus cameronensis*. *International Journal of Primatology*, 26(1), 191-206.

Powell, R. A. (2000). Animal home ranges and territories and home range estimators. *Research techniques in animal ecology: controversies and consequences*, 442, 65-110.

Prévot, C., & Licoppe, A. (2013). Comparing red deer (*Cervus elaphus* L.) and wild boar (*Sus scrofa* L.) dispersal patterns in southern Belgium. *European Journal of Wildlife Research*, 59(6), 795-803.

Qviller, L., Viljugrein, H., Loe, L. E., Meisingset, E. L., & Mysterud, A. (2016). The influence of red deer space use on the distribution of *Ixodes ricinus* ticks in the landscape. *Parasites & vectors*, 9(1), 1-9.

Reading, R. P., Bradley, J., Hancock, P., Garbett, R., Selebatso, M., & Maude, G. (2019). Home-range size and movement patterns of Hooded Vultures *Necrosyrtes monachus* in southern Africa. *Ostrich*, 90(1), 73-77.

Recio, M. R., Mathieu, R., Maloney, R., & Seddon, P. J. (2010). First results of feral cats (*Felis catus*) monitored with GPS collars in New Zealand. *New Zealand Journal of Ecology*, 288-296.

Richard-Hansen, C., Surugue, N., Khazraie, K., Le Noc, M., & Grenand, P. (2014). Long-term fluctuations of white-lipped peccary populations in French Guiana. *Mammalia*, 78(3), 291-301.

Rodgers, A. R., Carr, A. P., Beyer, H. L., Smith, L., & Kie, J. G. (2007). HRT: home range tools for ArcGIS.

Rodriguez, D., Reyes, A., del Pilar Tarquino-Carbonell, A., Restrepo, H., & Reyes-Amaya, N. (2021). Space use by a male Andean bear (*Tremarctos ornatus*) tracked with GPS telemetry in the Macizo Chingaza, Cordillera Oriental of the Colombian Andes. *Notas sobre Mamíferos Sudamericanos*, 3.

Ross, J. P., Bluett, R. D., & Dreslik, M. J. (2019). Movement and home range of the smooth softshell turtle (*Apalone mutica*): Spatial ecology of a river specialist. *Diversity*, 11(8), 124.

Row, J. R., & Blouin-Demers, G. (2006). Kernels are not accurate estimators of home-range size for herpetofauna. *Copeia*, 2006(4), 797-802.

Said, S., Jégoux, F., Chassagneux, A., Liehrmann, O., Shukla, U., Guillaumat, E., & Baltzinger, C. (2019). Dispersion des graines par le cerf élaphe (*Cervus elaphus*) pendant la saison de chasse: mise en place d'une expérimentation dans le Domaine national de Chambord.

Salvatori, M., De Groot, J., van Loon, E., De Baets, B., Morellet, N., Focardi, S., ... & Cagnacci, F. (2022). Day versus night use of forest by red and roe deer as determined by Corine Land Cover and Copernicus Tree Cover Density: assessing use of geographic layers in movement ecology. *Landscape Ecology*, 37(5), 1453-1468.

Schieltz, J. M., Okanga, S., Allan, B. F., & Rubenstein, D. I. (2017). GPS tracking cattle as a monitoring tool for conservation and management. *African journal of range & forage science*, 34(3), 173-177.

Schofield, G., Bishop, C. M., MacLean, G., Brown, P., Baker, M., Katselidis, K. A., ... & Hays, G. C. (2007). Novel GPS tracking of sea turtles as a tool for conservation management. *Journal of experimental marine biology and ecology*, 347(1-2), 58-68.

Schwager, M., Anderson, D. M., Butler, Z., & Rus, D. (2007). Robust classification of animal tracking data. *Computers and Electronics in Agriculture*, 56(1), 46-59.

Scull, P., Palmer, M., Frey, F., & Kraly, E. (2012). A comparison of two home range modeling methods using Ugandan mountain gorilla data. *International Journal of Geographical Information Science*, 26(11), 2111-2121.

Seaman, D. E., Griffith, B., & Powell, R. A. (1998). KERNELHR: a program for estimating animal home ranges. *Wildlife Society Bulletin*, 95-100.

Seidel, D. P., Dougherty, E., Carlson, C., & Getz, W. M. (2018). Ecological metrics and methods for GPS movement data. *International Journal of Geographical Information Science*, 32(11), 2272-2293.

Shenk, J., Byttner, W., Nambusubramaniyan, S., & Zoeller, A. (2021). Traja: A Python toolbox for animal trajectory analysis. *Journal of Open Source Software*, 6(63), 3202.

Signer, J., & Balkenhol, N. (2015). Reproducible home ranges (rhr): A new, user-friendly R package for analyses of wildlife telemetry data. *Wildlife Society Bulletin*, 39(2), 358-363.

Simcharoen, S., Barlow, A. C., Simcharoen, A., & Smith, J. L. (2008). Home range size and daytime habitat selection of leopards in Huai Kha Khaeng Wildlife Sanctuary, Thailand. *Biological Conservation*, 141(9), 2242-2250.

Smith, B. J., Cherkiss, M. S., Hart, K. M., Rochford, M. R., Selby, T. H., Snow, R. W., & Mazzotti, F. J. (2016). Betrayal: radio-tagged Burmese pythons reveal locations of conspecifics in Everglades National Park. *Biological Invasions*, 18(11), 3239-3250.

Srinath, K. R. (2017). Python—the fastest growing programming language. *International Research Journal of Engineering and Technology*, 4(12), 354-357.

Stanley, C. Q., Dudash, M. R., Ryder, T. B., Shriver, W. G., Serno, K., Adalsteinsson, S., & Marra, P. P. (2021). Seasonal variation in habitat selection for a Neotropical migratory songbird using high-resolution GPS tracking. *Ecosphere*, 12(3), e03421.

Stark, D. J., Vaughan, I. P., Ramirez Saldivar, D. A., Nathan, S. K., & Goossens, B. (2017). Evaluating methods for estimating home ranges using GPS collars: A comparison using proboscis monkeys (*Nasalis larvatus*). *PLoS one*, 12(3), e0174891.

Steiniger, S., & Hunter, A. J. (2012). OpenJUMP HoRAE—A free GIS and toolbox for home-range analysis. *Wildlife Society Bulletin*, 36(3), 600-608.

Stoner, D. C., Rieth, W. R., Wolfe, M. L., Mecham, M. B., & Neville, A. (2008). Long-Distance Dispersal of a Female Cougar in a Basin and Range Landscape. *Journal of Wildlife Management*, 72(4), 933.

Teng, Y., Shupei, T. A. N. G., Menghe, D., Wu, L., Zhiqing, H. A. N., Yingying, H. A. N., & Weidong, B. A. O. (2021). A Pilot Study on Home Range and Habitat Use of Chinese Goral (*Naemorhedus Griseus*): Exploring GPS Tracking Data in Cliff Landscape by Three Estimation Methods.

Treep, J., Bohrer, G., Shamoun-Baranes, J., Duriez, O., de Moraes Frasson, R. P., & Bouten, W. (2016). Using high-resolution GPS tracking data of bird flight for meteorological observations. *Bulletin of the American Meteorological Society*, 97(6), 951-961.

Urbano, F., Cagnacci, F., Calenge, C., Dettki, H., Cameron, A., & Neteler, M. (2010). Wildlife tracking data management: a new vision. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1550), 2177-2185.

Urios, V., López-López, P., Limiñana, R., & Godino, A. (2010). Ranging behaviour of a juvenile Bearded Vulture (*Gypaetus barbatus meridionalis*) in South Africa revealed by GPS satellite telemetry. *Ornis Fennica*, 87(3), 114.

Van Der Walt, S., Colbert, S. C., & Varoquaux, G. (2011). The NumPy array: a structure for efficient numerical computation. *Computing in science & engineering*, 13(2), 22-30.

Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace.

Van Rossum, G., Warsaw, B., & Coghlann, N. (2001). PEP 8—style guide for python code. *Python.org*, 1565.

Vieira, W. F., Kerry, C., & Hockings, K. J. (2019). A comparison of methods to determine chimpanzee home-range size in a forest–farm mosaic at Madina in Cantanhez National Park, Guinea-Bissau. *Primates*, 60(4), 355-365.

Walter, W. D., Fischer, J. W., Baruch-Mordo, S., & VerCauteren, K. C. (2011). What is the proper method to delineate home range of an animal using today's advanced GPS telemetry systems: the initial step. *Modern telemetry*, 68.

Walton, Z., Samelius, G., Odden, M., & Willebrand, T. (2018). Long-distance dispersal in red foxes *Vulpes vulpes* revealed by GPS tracking. *European Journal of Wildlife Research*, 64(6), 1-6.

Webb, S. L., Dzialak, M. R., Mudd, J. P., & Winstead, J. B. (2013). Developing spatially-explicit weighting factors to account for bias associated with missed GPS fixes in resource selection studies. *Wildlife Biology*, 19(3), 257-273.

Wolfson, D., Fieberg, J., Lawrence, J. S., Cooper, T. R., & Andersen, D. E. (2017). Range overlap between mid-continent and eastern sandhill cranes revealed by GPS-tracking. *Wildlife Society Bulletin*, 41(3), 489-498.

Woolnough, A. P., Lowe, T. J., & Rose, K. (2006). Can the Judas technique be applied to pest birds?. *Wildlife Research*, 33(6), 449-455.

Yiu, S. W., Parrini, F., Karczmarski, L., & Keith, M. (2016). Home range establishment and utilization by reintroduced lions (*Panthera leo*) in a small South African wildlife reserve.

Zengetya, F., Murwira, A., & de Garine-Wichatitsky, M. (2011). An IKONOS-based comparison of methods to estimate cattle home ranges in a semi-arid landscape of southern Africa. *International Journal of Remote Sensing*, 32(22), 7805-7826.

Zhang, J., Rayner, M., Vickers, S., Landers, T., Sagar, R., Stewart, J., & Dunphy, B. (2019). GPS telemetry for small seabirds: using hidden Markov models to infer foraging behaviour of common diving petrels (*Pelecanoides urinatrix urinatrix*). *Emu-Austral Ornithology*, 119(2), 126-137.

Annexes

Annexe 1 : Lien vers le dépôt GitHub pour l'accès au code source et aux vidéos explicatives

<https://github.com/QuentinWATTHEZ/TracksMining>

Annexe 2 : Formulaire d'évaluation de l'enquête sur les cartes interactives

1. Quel est votre secteur d'activité *

- Bureau d'études
- ONG / Association
- Service Public : DNF, DEMNA,...
- Recherche scientifique
- Autre

2. Si vous réalisez des analyses de déplacements d'animaux, quels logiciels informatiques utilisez-vous?

3. Quel(s) usage(s) faites-vous des données de télémétrie animale?

- Utilisation pour des études d'impact d'aménagement sur la faune (ex: routes, barrages, éoliennes, etc.)
- Amélioration de la connaissance générale du comportement des espèces
- Mise au point de plans de gestion/régulation de l'espèce
- Mise en place de mesures de protection de l'espèce
- Modélisations épidémiologiques
- Autre

4. Seriez-vous intéressé(e) par l'accès à une interface en ligne, hébergée et maintenue par le DEMNA, vous permettant automatiquement de générer des cartes personnalisées (style+contenu)?

- Oui
- Non

5. Avez-vous pu visualiser la carte interactive sans souci technique? (bug, temps de chargement, problème d'affichage, etc)

- Oui
- Non

6. Si vous avez coché "Non", quel était le problème?

7. Parmi ces perspectives à l'étude, quelles fonctionnalités de la carte interactive seraient les plus pertinentes pour votre utilisation?

- Détection automatique de géolocalisations erronées
- Suivi simultané de plusieurs individus
- Transposition à d'autres méthodes de suivis (Radio, RFID, ...)
- Personnalisation du style: Choix des couleurs, ajout d'un logo, etc.
- Autre

8. Avez-vous des commentaires particuliers sur ce type de carte? (Suggestions, remarques, etc.)

Sur une échelle de 1 à 5, comment évalueriez-vous:

- 1= Médiocre
- 2= Mauvais
- 3= Moyen
- 4= Bon
- 5= Excellent

9. Le caractère intuitif général de la carte

- 1
- 2
- 3
- 4
- 5

10. La pertinence des informations contenues sur la carte

1	2	3	4	5
<input type="radio"/>				

11. L'aspect visuel et esthétique

1	2	3	4	5
<input type="radio"/>				

12. La facilité d'utilisation de la barre de défilement de l'animation

1	2	3	4	5
<input type="radio"/>				

13. La plus-value de la carte interactive par rapport à une carte statique (type pdf).

1	2	3	4	5
<input type="radio"/>				