



HAL
open science

Définition, conception et évaluation d'un système d'information spatial pour le pastoralisme en zones périméditerranéennes françaises

Urcel Kalenga Tshingomba

► To cite this version:

Urcel Kalenga Tshingomba. Définition, conception et évaluation d'un système d'information spatial pour le pastoralisme en zones périméditerranéennes françaises. Sciences de la Terre. AgroParisTech, 2023. Français. NNT : 2023AGPT0004 . tel-04105363

HAL Id: tel-04105363

<https://pastel.hal.science/tel-04105363>

Submitted on 24 May 2023

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

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

**THÈSE POUR OBTENIR LE GRADE DE DOCTEUR
DE L'INSTITUT DES SCIENCES ET INDUSTRIES DU VIVANT ET DE
L'ENVIRONNEMENT - AGROPARISTECH**

N° : 2023AGPT0004

En Géomatique

**École doctorale GAIA – Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau – n°584
Portée par l'Université de Montpellier**

Unité de recherche TETIS - Territoires, Environnement, Télédétection et Information Spatiale

**Définition, conception et évaluation d'un système
d'information spatial pour le pastoralisme en zones
périméditerranéennes françaises**

**Présentée par Urcel KALENGA TSHINGOMBA
Le 28 mars 2023**

**Sous la direction de Magali JOUVEN et Maguelonne TEISSEIRE
et Lucile SAUTOT, Co-encadrante**

Devant le jury composé de

François PINET, Directeur de Recherche, INRAE (centre Clermont-Auvergne)	Rapporteur
Gilles BRUNSCHWIG, Professeur, VetAgro Sup Marcy-L'Etoile	Rapporteur
Elvira SALES-BAPTISTA, Maîtresse de Conférences , Université d'Evora (Portugal)	Examinatrice
Eliel GONZALEZ-GARCIA, Directeur de Recherche, INRAE (centre Occitanie-Montpellier)	Examineur
Carmen GERVET, Professeure, Université de Montpellier	Présidente
Pierre-Yves VION, ICPEF, AgroParisTech	Invité

Résumé

"Définition, conception et évaluation d'un système d'information spatial pour le pastoralisme en zones périméditerranéennes françaises"

Le pastoralisme est un mode d'élevage très répandu en régions péri-méditerranéennes françaises, cohérent avec les attentes sociétales et pourvoyeur de multiples services. Sa pérennité dépend notamment de sa capacité à valoriser efficacement et durablement les végétations de parcours. Le comportement alimentaire et spatial des troupeaux est un élément clé à prendre en compte pour ajuster la conduite du pâturage ; aujourd'hui, il peut être documenté et analysé à travers des outils numériques. L'application de l'élevage de précision au pastoralisme est contrainte par l'hétérogénéité des milieux et la diversité des systèmes agro-sylvo-pastoraux. Cependant, on dispose aujourd'hui d'une diversité de données pour documenter le fonctionnement de l'écosystème pastoral. L'intégration et la combinaison de données hétérogènes au sein d'un système d'information (SI) pourrait permettre de proposer une analyse approfondie du comportement des troupeaux sur parcours. Pour y parvenir, ce SI doit prendre en charge l'hétérogénéité des données et fournir les fonctionnalités d'analyse en vue de produire des indicateurs d'aide à la conduite du pâturage sur parcours.

Le principal objectif de cette thèse était de proposer et tester des approches techniques et méthodologiques permettant de définir un SI qui répond à ces considérations. Pour atteindre cet objectif, nos travaux ont été réalisés en trois étapes : i) identification des besoins des éleveurs et traduction en caractéristiques techniques et fonctionnelles du SI ; ii) définition des critères de qualité des données spatiales : précision, échelle de représentation et sémantique, granularité temporelle ; iii) exploration de différentes approches d'analyse de données spatiales, et en particulier analyse de séquences fréquentes et extraction de motifs de points. A l'issue de cette thèse, nous retenons le lac de données Hadoop d'Apache comme SI pour ses nombreux avantages pour la gestion des données hétérogènes et l'intégration de modules d'analyse et de diffusion d'information adaptées au contexte de notre étude. Nous proposons de gérer le stockage distribué et le traitement parallélisé via le système de fichiers HDFS (Hadoop Distributed File System). Les critères de qualité des données spatiales que nous retenons sont : i) 10m de résolution spatiale pour bien décrire l'hétérogénéité du milieu, ii) une précision suffisamment élevée des classes d'occupation des sols en vue de mieux caractériser les classes d'intérêt pastoral, et iii) une fréquence d'acquisition de 5 minutes pour les données de localisation des troupeaux afin de faciliter l'extraction de connaissances à différentes

échelles temporelles (horaire, journalière, phase de pâturage). Les prétraitements effectués n'ont pas permis d'améliorer la précision d'identification des surfaces de pelouse, lande et bois pâturés par rapport à la base OSO ; nous suggérons plutôt de mobiliser des données complémentaires telles que le NDVI et l'OSAVI pour décrire l'état et la dynamique des végétations. Concernant les traces GPS des troupeaux, les méthodes d'analyse de motifs de points ont permis de produire des indicateurs d'étendue explorée, de dispersion et de densité de présence au sein du parc. Malheureusement, les conditions expérimentales et notamment la taille des parcs (environ 1ha), n'ont pas favorisé une variabilité des facteurs du milieu et ainsi, n'ont pas permis de caractériser leur influence sur le comportement des troupeaux.

Cette thèse apporte une contribution opérationnelle et scientifique pour le pastoralisme en zones méditerranéennes et montagneuses françaises. La contribution opérationnelle concerne la définition et la conception d'un SI en adéquation avec les besoins des éleveurs pastoraux et les enjeux liés aux données hétérogènes, et pouvant accompagner la conduite du pâturage sur parcours. La contribution scientifique concerne la définition des méthodes d'analyse permettant de produire des indicateurs pouvant traduire l'utilisation spatiale et temporelle des milieux pastoraux variés par les troupeaux et de hiérarchiser l'influence des facteurs du milieu sur le comportement des troupeaux.

Mots clefs : *Élevage de précision, données hétérogènes, systèmes pastoraux, comportement spatial, indicateurs.*

Abstract

"Definition, design and assessment of a spatial information system for pastoralism in French Mediterranean areas"

Pastoral farming is a very widespread grazing practice in the French peri-Mediterranean regions, consistent with societal expectations and a provider of multiple services. Its sustainability depends in particular on its ability to effectively and sustainably develop rangeland vegetation. The grazing and spatial behavior of the herds is a key element to take into account in order to adjust the grazing management ; today, it can be documented and analyzed through digital tools. The application of precision livestock farming to pastoral farming is constrained by the heterogeneity of environments and the diversity of agro-sylvo-pastoral systems. However, a diversity of data is available today to document the functioning of the pastoral ecosystem. The integration and combination of heterogeneous data within an information system (IS) could allow for an in-depth analysis of the behavior of herds on pasture. To achieve this, the IS must be able to handle the heterogeneity of the data and provide analysis functionalities to produce indicators to assist in the management of grazing on rangelands.

The main objective of this thesis was to propose and test technical and methodological approaches to define an IS that meets these considerations. To achieve this objective, our work was carried out in three steps : i) identification of the needs of the breeders and translation into technical and functional characteristics of the IS ; ii) definition of spatial data quality criteria : accuracy, representation scale and semantics, temporal granularity ; iii) exploration of different spatial data analysis approaches, in particular frequent sequence and point pattern analysis. At the end of this thesis, we select the Apache Hadoop data lake as the IS for its numerous advantages in managing heterogeneous data and integrating analysis and diffusion modules adapted to the context of our study. We propose to manage distributed storage and parallelized processing via the HDFS (Hadoop Distributed File System). The spatial data quality criteria we use are : i) 10m spatial resolution to describe the heterogeneity of the environment, ii) a sufficiently high precision of the land use in order to better characterize the classes of pastoral interest, and iii) an acquisition frequency of 5 minutes for the herd location data in order to facilitate the extraction of knowledge at different temporal scales (hourly, daily, grazing phase). The pre-processing carried out did not improve the accuracy of grassland, moorland and woodland identification compared to the OSO database ; we sug-

gest instead to mobilize complementary data such as NDVI and OSAVI to describe the state and dynamic of vegetation. Regarding the GPS tracks of the herds, point pattern analysis methods were able to produce indicators of the explored area, dispersion and presence density within the park. Unfortunately, the experimental conditions and in particular the size of the parks (about 1ha), did not favor a variability of environmental factors and thus did not allow characterizing their influence on herd behavior.

This thesis brings an operational and scientific contribution for pastoral farming in French Mediterranean and mountainous areas. The operational contribution concerns the definition and the design of an IS in adequacy with the needs of the pastoral farmers and the challenges related to the heterogeneous data, which can be used as reference for to support grazing management on rangelands. The scientific contribution concerns the definition of analysis methods to produce indicators allowing to describe the spatial and temporal use of rangelands by herds and to hierarchize the influence of environmental factors on they grazing and spatial behaviour.

Key words : *Precision livestock farming, Heterogeneous data, Pastoral systems, Spatial behavior, Indicators.*

Remerciements

Je ne puis me retenir d'avouer mon émotion inattendue à l'écriture de ces dernières lettres de ma thèse retrouvées dans cette partie. Des moments d'enthousiasme, de joie et d'encouragement, mais aussi de stress, d'anxiété et de dépression foisonnent en moi comme pour dresser un bilan de mes trois années et quelques mois de thèse passés. Si durant tout mon cursus scolaire et académique j'ai pu me lancer avec assurance et certitude, j'avoue que le terrain de cette thèse m'a dévoilé une autre face du défi, l'iceberg du doute et de l'ignorance. Mais finalement j'y arrive, et cela n'aurait pas été possible sans la participation de différentes personnes qui ont apporté chacune une pierre à l'édifice de ma thèse et que j'aimerais sincèrement remercier.

Je commencerai par le labo CarTeS, où en suivant les pas de la professeure Catherine NSIAMI et l'engagement des collègues Scott Naweji, John Tshibangu, Marc Mashau, William Delonga, Junior Muyumba, Olivier Ngoie et Herman Banza, sans oublier l'encadrement du professeur Joseph Hosana et de l'ingénieur John Katanga (paix à son âme), que je découvre mon goût pour la recherche et l'enseignement. Cela m'a emmené à poursuivre la formation de MS-SILAT en France où j'ai été sélectionné après un entretien avec Frédérique Portet et Pierre-Yves Vion et m'a ouvert la voie à la thèse. Je remercie toutes ces personnes pour ce chemin tracé.

A vous mes encadrantes, qui avez cru en moi et donné le meilleur de vous-même pour me mener vers cette réussite. Merci pour votre accompagnement et votre soutien durant ces moments inoubliables de ma vie. La suite de ma carrière restera incontestablement marquée par vos empreintes respectives.

Merci Lucile, par qui tout a commencé pour cette thèse. Alors que je doutais d'y candidater, tu m'as encouragé à le faire après t'en avoir parlé dans le couloir du MS-SILAT. Pour ta présence, ton encadrement et ton rôle si important que tu as joué durant ma thèse, merci.

Merci Magali. Ta capacité et ta profondeur d'analyse m'as toujours motivé à fournir des efforts pour m'améliorer. Les réunions et les travaux passés ensemble, m'ont permis d'éclairer ma compréhension de la recherche et de prendre de la hauteur dans la réalisation de mes travaux. Pour tes conseils et tes orientations scientifiques qui resteront toujours une référence pour moi, merci.

Merci Maguelonne. Tes compétences scientifiques et organisationnelles seront constamment un repère pour moi. Pour tes orientations techniques et méthodologiques, pour ta présence régulière dans l'encadrement, pour ta rigueur et pour ta franchise, merci.

Mes remerciements s'adressent également aux membres de mon comité de thèse, Mathieu Roche, Charles-Henri Moulin, Éric Leclercq et Carmen Gervet, pour avoir consacré leur temps à mon écoute et pour leurs orientations scientifiques.

A mes collègues de la maison de la Télédétection et de l'UMR SELMET. Tous les instants passés ensemble lors des balades, des soirées, etc., ont été des séances de thérapie sans égales pendant ces moments si particulier. Merci Marie S., Nuria P., Mehtab S., Bahdja B., Kamel L., Anouk S., Ellen L., Charlotte E., Etienne S., Momo H., Johann D., Sami, Léo G., Clovis G., Frédérique M., Vicent D., Jean Eudes G., Béthynna N., ... Merci également à Rémy Découpes d'avoir consacré des moments pour me débloquent dans la mise en place du Data Lake et à Imad Shaqura pour les échanges techniques sur la caractérisation des parcours.

Merci aux stagiaires Bassira, Daria, Ines et Zoe, aux éleveurs pastoraux, aux chercheurs et enseignants qui ont rendu possible l'avancement de mes travaux grâce à leur participation au projet P@astor_All.

Un très grand merci aux amis, ceux avec qui j'ai partagé des moments agréables et que j'ai dû sans relâche indépendamment de ma volonté casser les pieds avec mes "problèmes de thèse". Ceux qui ont servi d'exutoire à mes angoisses, stress et inquiétudes. Oriane Dubreuil, Christela Badjina, Ruth Muakana, Armila Ilunga, Dav Ebegno, Junior Muyumba, Herman Banza, Hamza Debabi, Yannick Dogne, Pascal Kahazi, Lewis Luhembwe, je vous remercie d'être restés près de moi.

Enfin, ma famille qui m'a beaucoup manquée depuis mon arrivée en France ça fait plus de 4 ans déjà. Merci d'avoir endurer la peine de me voir partir loin de vous et de m'accorder toujours cet amour inconditionnel. Papa, maman, Chillya, Gloria, Amélia, Yoshua, Hans, Moïse, Esli, trouvez ici l'expression de toute ma gratitude. Merci également à mon cher oncle Othis Nyembo qui a toujours été avec moi depuis mon arrivée en France, et à mes tantes Chillya et Michou Kalenga pour nos échanges téléphoniques si apaisantes.

Urcel

Table des matières

Résumé	ii
Abstract	iv
Remerciements	vi
Table des figures	xii
Liste des tableaux	xviii
1 Le pastoralisme en zones méditerranéennes et montagneuses françaises et ses enjeux	1
1.1 Le pastoralisme en zones méditerranéennes françaises	2
1.1.1 Définition	2
1.1.2 Milieux pastoraux en zones méditerranéennes françaises	3
1.1.3 Les savoirs pastoraux en zones méditerranéennes françaises	3
1.2 Les systèmes agro-sylvo-pastoraux en zones méditerranéennes françaises	5
1.3 Enjeux autour du pastoralisme en zones Méditerranéennes françaises	8
1.3.1 Viabilité économique du pastoralisme	8
1.3.2 Adaptation au changement climatique	9
1.3.3 Gestion et conservation des ressources naturelles et de la biodiversité	9
1.3.4 Accès au foncier	10
1.4 Potentiel des technologies numériques pour la documentation du fonctionnement de l'écosystème pastoral	11

1.4.1	Quelles informations clés pour contribuer à la documentation de l'écosystème pastoral ?	11
1.4.2	Projet P@stor_All	12
1.5	Objectifs et intérêts de la thèse	13
2	Conduite du pâturage sur parcours : panorama des contributions d'outils numériques	15
2.1	Élevage et pâturage de précision : concept et définition	16
2.2	Contribution des technologies numériques dans le cadre l'élevage de précision et opportunités pour la conduite du pâturage en régions Occitanie et PACA .	18
2.2.1	Applications en élevage de précision	18
2.2.2	Enjeux associés au développement d'un outil de pâturage de précision dans le cadre du pastoralisme en zones méditerranéennes françaises .	21
2.3	Émergence des données spatiales et opportunités pour le développement d'un système d'information pour le pastoralisme en régions Occitanie et PACA . .	22
2.4	Problématique et questions de recherche	24
3	Approche de co-construction d'un système d'information pour l'intégration des données du pastoralisme	27
3.1	Introduction	28
3.2	Modélisation de l'écosystème agro-sylvo-pastoral	29
3.2.1	Approche méthodologique	29
3.2.2	Modèle du comportement alimentaire et spatial des animaux	32
3.3	Sélection des données pour la documentation de l'écosystème agro-sylvo-pastoral	33
3.3.1	Données publiques	34
3.3.2	Données locales	35
3.3.3	Association des données pour modéliser l'écosystème agro-sylvo-pastoral	36
3.4	Définition des fonctionnalités du système d'information	37
3.5	Choix du système d'information pour l'intégration des données du pastoralisme	39
3.5.1	État de l'art sur les systèmes de gestion de données	39
3.5.2	Critères de choix et de l'implémentation du SI	41

3.6	Le système d'information spatial pour la gestion des données du pastoralisme : P@stor_All	42
3.6.1	Modèle conceptuel des données	42
3.6.2	Architecture du système d'information	43
3.6.3	Prototype opérationnel du lac de données	49
3.7	Conclusions et perspectives	51
4	Préparation et prétraitement des données du pastoralisme	54
4.1	Introduction	55
4.2	Qualité des données spatiales du pastoralisme	57
4.2.1	Modélisation des ressources pastorales	57
4.2.2	Modélisation de l'environnement de pâturage	59
4.2.3	Modélisation du comportement des animaux	60
4.3	Cartographie des surfaces pastorales à l'aide des données Sentinel 2 et des données publiques	61
4.3.1	Données	62
4.3.2	Méthodologie	65
4.3.3	Résultats	72
4.4	Prétraitements des données de localisation des troupeaux	76
4.4.1	Caractéristiques des données GPS collectées	76
4.4.2	Préparation des données GPS	77
4.5	Densité de végétation et topographie du milieu	78
4.5.1	Densité de végétation	78
4.5.2	Pente et orientation	79
4.6	Conclusion	80
5	Croisement des informations spatiales pour la production d'indicateurs d'aide à la décision en élevage pastoral	82
5.1	Introduction	83
5.2	Approches d'analyse des données spatio-temporelles spatiales pour la produc- tion d'indicateurs du pastoralisme	84

5.2.1	Extraction de séquences fréquentes	85
5.2.2	Analyse de motifs de points	87
5.3	Croisement des informations spatiales pour modéliser et caractériser l'utilisation des parcours par les animaux	98
5.4	Application en conditions expérimentales	100
5.4.1	Contexte	100
5.4.2	Données utilisées	101
5.4.3	Production d'indicateurs du comportement des agnelles	104
5.4.4	Analyse des informations	105
5.4.5	Résultats	108
5.4.6	Discussion	117
5.4.7	Conclusion	120
6	Discussion générale et conclusion	121
6.1	Vue générale des travaux	122
6.2	Évaluation du lac de données spatiales dans le cadre de l'élevage pastoral	123
6.2.1	Intérêt des données pour la gestion de l'élevage pastoral en zones méditerranéennes et montagneuses françaises	123
6.2.2	Complémentarité par rapport à d'autres outils développés pour la gestion de l'élevage pastoral	124
6.3	Contributions techniques et méthodologiques du lac de données spatiales pour la gestion et l'analyse des données du pastoralisme	126
6.3.1	Considérations pour l'hétérogénéité et la variété des conditions pastorales	126
6.3.2	Pertinence des méthodes d'analyse d'indicateurs pour décrire le comportement spatial des troupeaux	127
6.4	Valeur et portée des méthodes proposées	129
6.4.1	Qualité des données	129
6.4.2	Validité des analyses et des indicateurs d'utilisation de l'espace par les troupeaux	130
6.4.3	Application des approches développées à d'autres contextes d'élevage pastoral	132

6.5 Perspectives de recherche	133
6.6 Conclusion générale	134
Annexes	I
A Figures	II
B Tableaux	XXVIII
Références	XXIX

Table des figures

1.1	Milieus utilisés par l'élevage pastoral dans les régions Occitanie et PACA. Cette distribution spatiale correspond à la déclaration de 2021 et ne représente pas de manière exhaustive la distribution de ces végétations.	7
3.1	Différentes étapes réalisées pour identifier les facteurs clés influençant le comportement des troupeaux au pâturage. Les facteurs en couleur verte n'ont pas été retenus à la fin du processus.	32
3.2	Le modèle conceptuel du comportement alimentaire et spatial des herbivores domestiques des systèmes agro-sylvo-pastoraux en zones méditerranéennes et montagneuses françaises. La ligne continue représente l'impact d'un facteur sur le comportement des Troupeaux et la ligne discontinue représente l'impact entre les facteurs.	33
3.3	Modèle conceptuel de données du système d'information spatiale pour le système agro-sylvo-pastoral dans les zones méditerranéennes et montagneuses françaises.	43
3.4	Architecture globale du lac de données spatiales pour la gestion des données du pastoralisme	45
3.5	Infrastructure du prototype du lac de données spatiales implémentée pour l'intégration des données du pastoralisme.	50
3.6	Les interfaces des différents composants du lac de données. 1) Structure de fichiers dans HDFS listant les dossiers des données spatiales des exploitations pastorales. 2) Catalogue de métadonnées dans GeoNetwork donnant accès aux données via un lien web vers HDFS. 3) Visualisation dans Kibana (sous forme de carte de chaleur) des informations de localisation GPS (d'une journée) d'un animal, ingérées dans Elasticsearch.	51

4.1	F-score de 23 classes de l'occupation des sols OSO de 2019 et des classes regroupées (en rouge) de l'occupation des sols OSO de 2017. Les F-scores des classes regroupées d'OSO 2017 ont été utilisés à défaut de la disponibilité de l'information pour la donnée de 2019. Source : http://osr-cesbio.ups-tlse.fr/~oso/	62
4.2	Composition colorée infrarouge fausses couleurs de la mosaïque de 21 tuiles de Sentinel 2 L3A couvrant les régions Occitanie et PACA. Source de données : www.theia-land.fr	64
4.3	Distribution des variables issues des premiers composants de l'ADL appliquée pour analyser la séparabilité des 5 classes à partir des variables de réflectance, spectrales et texturales, pour le département des Alpes-de-Haute-Provence (04).	67
4.4	Classement des variables en fonction de leurs scores d'importance après 20 itérations sur le jeu de données avec 50 000 objets par classe. Les lignes verticale et horizontale marquent respectivement les limites des variables importantes et du score minimum choisi. La barre bleue représente la moyenne et la noire l'écart-type associé pour les 20 itérations.	70
4.5	Schéma général de traitement d'amélioration de la précision de la classification des classes pastorales et des cultures pérennes. Avec comme données en entrée : le RPG 2019 et la BD-Forêt (4.A), l'occupation des sols OSO 2019 (4.C) et les données Sentinel 2 L3A (4.B).	71
4.6	Comparaison entre l'occupation des sols produite, OSO 2019 et la composition colorée infrarouge fausses couleurs.	75
4.7	Aperçu des prétraitements réalisés pour la production des données de densité de végétation, des variables topographiques et des trajectoires de troupeaux.	80

5.1	Exemple d'indicateurs de centrographie : moyenne et médiane centrales du motif de points (en bleu) représentant l'utilisation des parcs par les deux groupes d'agnelles allaitées par les mères (MAT) et allaitées artificiellement (ART) dans le site expérimental de la Fage. L'ellipse standard représente la tendance du motif de points de localisation des animaux. Les graphes A et B représentent l'utilisation de deux différents parcs par les deux groupes d'agnelles en été, le 09-07-2021 entre 00h et 03h. La moyenne et médiane centrales y sont confondues, les deux motifs étant à peu près groupés. Les graphes C et D représentent l'utilisation des mêmes groupes d'agnelles, le même jour entre 15h et 18h. La moyenne et médiane centrales y sont bien séparées. Dans le graphe C, les agnelles MAT se dispersent plus dans la direction Ouest-Est et occupent une plus grande étendue que les agnelles ART qui se dispersent plus dans la direction Sud-Nord	90
5.2	Exemple d'indicateurs d'étendue. Ici, l'enveloppe convexe est confondue à l'Alpha shape qui est une de ses variantes.	92
5.3	Exemple d'indicateurs de densité. A gauche : densités locales calculées au sein de chaque cellule de 10m X 10m. A droite : Densité locale par noyau.	95
5.4	Localisation de l'unité expérimentale de la Fage dans le département de l'Aveyron. Le grand parc 9 comprend les différents sous parcs pâturés séparément par les animaux en été et en automne. Les contours jaunes représentent les périmètres des massifs de plantes ligneuses.	101
5.5	Distribution des valeurs d'étendues explorées des agnelles MAT (ayant reçu un allaitement maternel) et ART (artificiellement allaitée) pour les créneaux horaires des phases de pâturage d'été.	109
5.6	Distribution des valeurs de dispersion des agnelles MAT (ayant reçu un allaitement maternel) et ART (artificiellement allaitée) pour les créneaux horaires des phases de pâturage d'été.	109
5.7	Distribution des valeurs d'étendues explorées par les agnelles MAT (ayant reçu un allaitement maternel) et ART (artificiellement allaitée) pour les créneaux horaires des phases de pâturage d'automne.	110
5.8	Distribution des valeurs de dispersion par les agnelles MAT (ayant reçu un allaitement maternel) et ART (artificiellement allaitée) pour les créneaux horaires des phases de pâturage d'automne.	111
5.9	Distribution spatiale de la densité d'utilisation de l'espace dans les zones de 10 m x 10 m pour le pâturage d'été par les agnelles MAT. Les contours (en jaune) représentent les limites des massifs de ligneux. L'étoile verte bordée de jaune représente la localisation de l'auge.	112

5.10	Distribution spatiale de la densité d'utilisation de l'espace dans les zones de 10 m x 10 m pour le pâturage d'été par les agnelles ART. Les contours (en jaune) représentent les limites des massifs de ligneux. L'étoile verte bordée de jaune représente la localisation de l'auge.	112
5.11	Distribution spatiale de la densité d'utilisation de l'espace dans les zones de 10 m x 10 m pour le pâturage d'automne par les agnelles MAT. Les contours (en jaune) représentent les limites des massifs de ligneux. L'étoile verte bordée de jaune représente la localisation de l'auge.	113
5.12	Distribution spatiale de la densité d'utilisation de l'espace dans les zones de 10 m x 10 m pour le pâturage d'automne par les agnelles ART. Les contours (en jaune) représentent les limites des massifs de ligneux. L'étoile verte bordée de jaune représente la localisation de l'auge.	113
5.13	Nombre de pixels par catégorie de densité de présence des agnelles sous les massifs de ligneux pour le pâturage d'été.	114
5.14	Nombre de pixels par catégorie de densité de présence des agnelles sous les massifs de ligneux pour le pâturage d'automne.	115
6.1	Variation des distances parcourues par un troupeau en fonction de la fréquence d'acquisition. Données issues d'un élevage privé. Source : Candela (2022)	130
A.1	Distribution des variables issues des premiers composants de l'ADL appliquée pour analyser la séparabilité des 5 classes à partir des variables de réflectance, spectrales et texturales, pour le département de l'Aude (11).	II
A.2	Distribution des variables issues des premiers composants de l'ADL appliquée pour analyser la séparabilité des 5 classes à partir des variables de réflectance, spectrales et texturales, pour le département du Var (83).	III
A.3	Résultats de la régression linéaire entre l'indicateur d'étendue explorée(Alpha shape) et les variables spatiales du milieu pour les agnelles MAT pâturant l'été.	IV
A.4	Résultats de la régression linéaire entre l'indicateur d'étendue explorée(Alpha shape) et les variables spatiales du milieu pour les agnelles ART pâturant l'été.	V
A.5	Résultats de la régression linéaire entre l'indicateur d'étendue explorée(Alpha shape) et les variables spatiales du milieu pour les agnelles MAT pâturant l'automne.	V
A.6	Résultats de la régression linéaire entre l'indicateur d'étendue explorée(Alpha shape) et les variables spatiales du milieu pour les agnelles ART pâturant l'automne.	VI

A.7	Résultats de la régression linéaire entre l'indicateur de dispersion (AND) et les variables spatiaux du milieu pour les agnelles MAT pâturent l'été.	VI
A.8	Résultats de la régression linéaire entre l'indicateur de dispersion (AND) et les variables spatiaux du milieu pour les agnelles ART pâturent l'été.	VII
A.9	Résultats de la régression linéaire entre l'indicateur de dispersion (AND) et les variables spatiaux du milieu pour les agnelles MAT pâturent l'automne.	VII
A.10	Résultats de la régression linéaire entre l'indicateur de dispersion (AND) et les variables spatiaux du milieu pour les agnelles ART pâturent l'automne.	VIII
A.11	Résultats de la régression linéaire entre l'indicateur de densité et les variables spatiaux du milieu pour les agnelles MAT pâturent l'été.	IX
A.12	Résultats de la régression linéaire entre l'indicateur de densité et les variables spatiaux du milieu pour les agnelles ART pâturent l'été.	X
A.13	Résultats de la régression linéaire entre l'indicateur de densité et les variables spatiaux du milieu pour les agnelles MAT pâturent l'automne.	X
A.14	Résultats de la régression linéaire entre l'indicateur de densité et les variables spatiaux du milieu pour les agnelles ART pâturent l'automne.	XI
A.15	Résultats de la régression linéaire entre l'indicateur d'étendue explorée(Alpha shape) et les paramètres météorologiques pour les agnelles MAT pâturent l'été.	XII
A.16	Résultats de la régression linéaire entre l'indicateur d'étendue explorée(Alpha shape) et les paramètres météorologiques pour les agnelles ART pâturent l'été.	XIII
A.17	Résultats de la régression linéaire entre l'indicateur d'étendue explorée(Alpha shape) et les paramètres météorologiques pour les agnelles MAT pâturent l'automne.	XIII
A.18	Résultats de la régression linéaire entre l'indicateur d'étendue explorée(Alpha shape) et les paramètres météorologiques pour les agnelles ART pâturent l'automne.	XIV
A.19	Résultats de la régression linéaire entre l'indicateur de dispersion (AND) et les paramètres météorologiques pour les agnelles MAT pâturent l'été.	XV
A.20	Résultats de la régression linéaire entre l'indicateur de dispersion (AND) et les paramètres météorologiques pour les agnelles ART pâturent l'été.	XVI
A.21	Résultats de la régression linéaire entre l'indicateur de dispersion (AND) et les paramètres météorologiques pour les agnelles MAT pâturent l'automne.	XVI
A.22	Résultats de la régression linéaire entre l'indicateur de dispersion (AND) et les paramètres météorologiques pour les agnelles ART pâturent l'automne.	XVII

A.23	Cartes de pente des parcs utilisés par les agnelles. En haut à gauche : parc utilisé en été par les agnelles allaitées naturellement. En haut à droite : parc utilisé en été par les agnelles allaitées artificiellement. En bas à gauche : parc utilisé en été par les agnelles allaitées artificiellement. En bas à droite : parc utilisé en été par les agnelles allaitées naturellement	XVIII
A.24	Cartes de NDVI de trois dates correspondant au début des phases de pâturage d'été au sein du parc utilisé par les agnelles naturellement allaitées. L'étoile verte bordée de jaune représente la localisation de l'auge.	XIX
A.25	Cartes de NDVI de trois dates correspondant au début des phases de pâturage d'été au sein du parc utilisé par les agnelles artificiellement allaitées. L'étoile verte bordée de jaune représente la localisation de l'auge.	XIX
A.26	Cartes de NDVI de deux dates correspondant au début des phases de pâturage d'automne au sein du parc utilisé par les agnelles naturellement allaitées. L'étoile verte bordée de jaune représente la localisation de l'auge.	XX
A.27	Cartes de NDVI de deux dates correspondant au début des phases de pâturage d'automne au sein du parc utilisé par les agnelles artificiellement allaitées. L'étoile verte bordée de jaune représente la localisation de l'auge.	XXI
A.28	Distances journalières parcourues par cinq brebis équipées au domaine du Merle. Les GPS des animaux 1 et 3 présentent des anomalies en début d'estive.	XXII
A.29	Surfaces journalières parcourues par cinq brebis équipées au domaine du Merle. Cet indicateur est beaucoup moins sensible aux erreurs de mesure du GPS.	XXIII
A.30	Distances journalières parcourues par 3 chèvres équipées dans une exploitation privée d'avril à octobre 2022.	XXIV
A.31	Surfaces journalières parcourues par 3 chèvres équipées dans une exploitation privée d'avril à octobre 2022.	XXV
A.32	Distribution des distances journalières parcourues par cinq brebis équipées au domaine du Merle.	XXVI
A.33	Distribution des surfaces journalières parcourues par cinq brebis équipées au domaine du Merle.	XXVI
A.34	Distribution des distances journalières parcourues par trois chèvres équipées dans une exploitation privée.	XXVII
A.35	Distribution des surfaces journalières parcourues par trois chèvres équipées dans une exploitation privée.	XXVII

Liste des tableaux

2.1	Exemples d'applications faisant appel aux technologies d'élevage de précision développées dans le cadre des élevages intensifs. Source : Banhazi et al. (2012)	19
3.1	Résumé des facteurs qui influencent le comportement alimentaire et spatial des animaux au pâturage.	31
3.2	Résumé des caractéristiques et des usages des données externes pouvant être utilisées pour modéliser l'utilisation des parcours par les troupeaux	34
3.3	Les facteurs du comportement alimentaire et spatial des troupeaux, leurs composants et les données disponibles permettant de les caractériser.	37
3.4	Les 7 exploitations privées et 2 sites expérimentaux agro-sylvo-pastorales enquêtés pour la constitution des besoins	38
3.5	Besoins des éleveurs exprimés sous forme de questions, considérés comme une base pour les fonctionnalités du SI.	39
3.6	Comparaison entre lac de données et d'un entrepôt de données	41
3.7	Aperçu de certaines analyses au sein du SI qui peuvent aider à répondre aux besoins des éleveurs	48
4.1	Extrait de la matrice de confusion d'OSO 2018 sur les classes à faibles F-scores et celles qui influencent le plus leurs imprécisions. Ver. = Verger, Vig. = Vigne, F. F. = Forêt de feuillus, F.C. = Forêt de conifères, Pel. = Pelouse, L. L. = Landes ligneuses, Pré. = Précision, Rap. = Rappel, et F.-S. = F-score.	58
4.2	Résumé des critères de qualité des données spatiales permettant de caractériser le fonctionnement des exploitations agro-sylvo-pastorales.	61
4.3	Correspondance entre les classes du RPG 2019 et les classes pastorales (hormis les bois pâturés) maintenues pour l'apprentissage des modèle de classification.	65

4.4	Bandes spectrales et indices de végétation et de sol dérivés de Sentinel 2 L3A, pris comme variables pour l'entraînement du modèle.	66
4.5	Nombre d'objets extraits pour chaque classe retenue dans le RPG 2019 et la BD-Forêt V2.0 et 3 jeux de données utilisés pour l'entraînement et l'évaluation de l'algorithme de classification.	68
4.6	Matrice de confusion et qualité de la prédiction produite à partir du modèle utilisant le jeu de données test 1.	72
4.7	Matrice de confusion et qualité de la prédiction produite à partir du modèle utilisant le jeu de données test 2.	73
4.8	Matrice de confusion et qualité de la prédiction produite à partir du modèle utilisant le jeu de données test 3.	74
4.9	Types d'utilisation des balises GPS au sein des exploitations pastorales privées et expérimentales.	77
4.10	Exemple de tableau spatio-temporel des données de localisation des animaux.	78
5.1	Base de données de séquences	86
5.2	Récapitulatif des indicateurs permettant de représenter et de caractériser l'utilisation des parcours par les animaux. Les indicateurs en gras sont ceux qui ont été retenus dans la suite du travail. Ils permettent de répondre aux deux premières questions faisant l'objet de ce chapitre.	98
5.3	Fréquences d'analyse des informations pour caractériser l'utilisation des parcours par les animaux	99
5.4	Aperçu des données GPS utilisées dans les analyses. Le total des points représente la somme des données GPS avec plus de 220 points de localisation par jour.	102
5.5	Données utilisées et applications relatives aux facteurs influençant l'utilisation de l'espace par les animaux. Ces données peuvent également être combinées pour analyser leur interaction sur l'utilisation de l'espace par les animaux.	104
5.6	Intervalles de classes de valeurs de densité de présence au sein des zones des sous parcs par les agnelles. Découpage en classe réalisé à partir des déciles sur les valeurs de densités de présence des agnelles pour l'ensemble des phases d'été et d'automne.	106

5.7	Constitution de deux types de données pour l'analyse de l'impact des facteurs du milieu sur l'utilisation de l'espace par les agnelles : i) association entre indicateurs et variables spatiaux et ii) association entre indicateurs et variables temporels. Les paramètres météorologiques horaires ont été moyennés par période de 3 heures pour la correspondance avec les indicateurs d'étendue et de dispersion.	107
5.8	P valeurs obtenus à partir de l'ANOVA entre les indicateurs de l'utilisation journalière de l'espace pour les mêmes créneaux horaires entre les deux groupes d'agnelles pour chaque phase de pâturage. P = Phase, NS = non significatif.	116
5.9	Coefficients de détermination de la régression linéaire multiple entre les indicateurs sélectionnés et les variables spatiales du milieu : <i>pente, NDVI, distance à l'auge, distance au ligneux le plus proche</i>	116
5.10	Coefficients de détermination de la régression linéaire multiple entre les indicateurs sélectionnés et les variables temporelles du milieu (paramètres météorologiques) : <i>humidité instantanée, rayonnement global, température horaire, vitesse du vent et précipitations²</i>	116
6.1	Comparaison des fonctionnalités proposées par les différents outils d'aide à la décision pour la conduite du pâturage sur parcours. X = forte intégration de la fonction, (x) = faible intégration de la fonction.	126
B.1	Résultats des P-valeurs des ANOVAs sur les paires de créneaux horaires entre les phases de pâturage	XXVIII

Chapitre 1

Le pastoralisme en zones méditerranéennes et montagneuses françaises et ses enjeux

Dans ce premier chapitre nous présentons un état des lieux sur le pastoralisme en zones méditerranéennes et montagneuses françaises et les principaux enjeux autour de cette forme d'élevage. Dans le but de répondre à ces enjeux à travers l'usage du numérique, nous présentons également les potentialités qu'offrent la diversité des données numériques pour contribuer au pastoralisme. Nous présentons en fin, les objectifs de la thèse qui se fondent sur l'intégration des données hétérogènes pour la documentation du pastoralisme en zones méditerranéennes et montagneuses françaises.

1.1 Le pastoralisme en zones méditerranéennes françaises

1.1.1 Définition

Avec l'évolution des pratiques agricoles depuis le milieu du 19^{ème} siècle, plusieurs systèmes d'élevage ont émergé allant des systèmes extensifs à des systèmes plus intensifs établis sur des espaces réduits. Le pastoralisme fait partie du premier type de système et se base essentiellement sur l'exploitation des milieux aux ressources spontanées pour l'alimentation des troupeaux.

Plusieurs définitions ont émergé pour décrire ce mode d'élevage. Dans le rapport de la World Initiative for Sustainable Pastoralism (WISP), les auteurs le définissent comme étant une forme d'élevage extensif par nature intégrant la mobilité des animaux (Wane et al., 2006). La mobilité est due à la recherche de ressources pastorales pour assurer les besoins alimentaires des animaux. Cette définition généraliste a été abordée par la WISP au cours des travaux visant de présenter les aspects économiques du pastoralisme. Davies et al. (2017) soulignent quant à eux la responsabilité des éleveurs à gérer le bétail et les terres à des fins non seulement économiques mais aussi de durabilité environnementale. Ils mentionnent la capacité des éleveurs à adapter la conduite du pâturage face aux conditions diverses des milieux et notamment aux variations saisonnières et climatiques. Ainsi, il nous paraît indispensable de situer cette pratique dans un contexte faisant apparaître l'ensemble des composantes qui le caractérisent, intégrant le berger ou l'éleveur, les animaux ainsi que les milieux exploités. En effet, il existe une relation interdépendante entre ces trois composantes (Bourbouze, 2018). Les éleveurs utilisent leurs connaissances des milieux pour y conduire spatialement et temporellement leurs troupeaux en vue d'y prélever les ressources alimentaires nécessaires. La consommation des ressources alimentaires par les troupeaux influence à son tour l'état des milieux pâturés. Ces systèmes d'élevage peuvent également bénéficier d'autres sources alimentaires que les fourrages des milieux spontanés afin de répondre aux objectifs de production. Dans ce sens, Tchakerian (2008) décrit le pastoralisme comme un type d'élevage qui utilise des parcours pour une part plus ou moins importante de l'alimentation des troupeaux. Cette définition renvoie à l'importance des parcours dans un élevage pastoral sans mettre de côté les apports pouvant provenir d'autres sources de production en vue de combler les besoins des troupeaux. Une définition assez concise que nous adoptons également dans ce travail est celle de l'Association Française de Pastoralisme (AFP) : "*Le Pastoralisme regroupe l'ensemble des activités d'élevage valorisant par un pâturage extensif les ressources fourragères spontanées des espaces naturels, pour assurer tout ou partie de l'alimentation des animaux*".

Dans la suite de notre travail, nous emploierons également le terme "d'élevage pastoral" pour désigner les exploitations d'élevage qui pratiquent le pastoralisme.

1.1.2 Milieux pastoraux en zones méditerranéennes françaises

Les milieux pastoraux des zones méditerranéennes françaises sont très variés en termes de végétation, de pédologie, de topographie et de conditions climatiques. Ces milieux sont principalement occupés par la végétation spontanée et se situent essentiellement à l'interface entre les espaces forestiers et les cultures. Ils sont le produit d'interventions passées et présentes (défrichement des forêts originelles à l'action de la dent de l'animal en passant par les gyrobroyages d'entretien) qui, combinées à leurs contextes géographique, topographique et écologique, expliquent leur grande variabilité (Tchakerian, 2008). Ils sont en fait considérés comme des terres indigènes (climaciques ou sous-climaciques), constitués essentiellement d'herbes, de plantes herbacées, d'arbustes et de graminées, et forment des écosystèmes pastoraux utilisés ou utilisables pour l'élevage de bétail en pâturage (Allen et al., 2011).

A cause de ces variations de conditions climatiques, pédologiques et topographiques, les ressources végétales des parcours sont très variables et hétérogènes entre les saisons et les années, en terme de quantité et de qualité (Dumont et al., 2015; Dono et al., 2016). Ces milieux également appelés parcours, sont souvent moins productifs à cause des conditions physiques contraignantes et non utilisés pour la production agricole. Les conditions climatiques extrêmes et incertaines, les sols souvent peu productifs (superficiels, secs ou trop humides et fragiles), les difficultés d'accès dues à la topographie et l'absence d'autres ressources telle que l'eau, sont la cause principale de la faible valorisation de ces terres.

Cette grande diversité des parcours oblige les éleveurs à s'adapter aux situations spécifiques de chaque milieu et conditionne l'existence d'une diversité de modes d'exploitation pastoraux. Ainsi, les éleveurs utilisent leurs savoirs faire pour ajuster à court et à moyen terme l'exploitation de ces milieux en fonction des objectifs de production tout en assurant leur maintien pour une utilisation à long terme.

1.1.3 Les savoirs pastoraux en zones méditerranéennes françaises

La diversité des milieux pastoraux des zones méditerranéennes françaises façonne les pratiques des éleveurs pastoraux. Ils se retrouvent au cœur de la prise de décision entre objectifs de production et disponibilité des ressources pastorales qu'ils composent pour constituer la ration des animaux. Ces pratiques confèrent aux éleveurs un sens du métier particulier, structuré par un rapport très fort à la nature (Nozieres et al., 2021). Pour les éleveurs, le milieu pastoral est un espace d'autonomie, d'imagination et de création dans lequel ils apprennent principalement à travers l'observation des comportements au pâturage de leurs animaux. Ils appréhendent les parcours à la fois en tant qu'espace de ressources et espace de vie pour ces derniers (Blanc, 2009).

Les pratiques de l'élevage pastoral s'amorcent à partir de la mobilisation d'espèces animales particulières, rustiques (Petit et al., 1994) et aptes à valoriser la végétation des par-

cours. Étant donné la nature hétérogène des parcours conditionnant la disponibilité des fourrages en quantité et en qualité, dans l'espace et dans le temps, les éleveurs doivent être capables de gérer cette complexité pour élever leurs troupeaux. Ils adoptent des stratégies variées pour ajuster la conduite des troupeaux au fil des jours et des saisons, suivant des modalités variables d'un éleveur à un autre, d'un territoire à un autre. Dans ce sens, la mobilité est un élément structurant leur travail. Deux visions peuvent être dressées, distinguant ainsi les éleveurs "agropastoraux" qui privilégient des ressources de proximité des "transhumants" qui élargissent les cadres du territoire traditionnel de l'élevage. Cette mobilité qui implique la gestion du foncier, active à certains moments des liens entre éleveurs et dans d'autres des liens avec d'autres acteurs (Lebaudy, 2004). Ces interactions nécessitent un cadre foncier propice et une connaissance adéquate du milieu et des pratiques pastorales pour une bonne coexistence de différentes parties.

Le savoir-faire et les compétences que les éleveurs développent concernent principalement la prise d'information sur les parcours et l'application des pratiques adaptées pour la conduite du pâturage. La prise d'information se réfère à la reconnaissance et l'identification des plantes consommables par les animaux et des caractéristiques des parcours, localement et largement dans le territoire (Guérin et al., 2007), ainsi que la connaissance du climat et des dynamiques spatiales et temporelles des plantes fourragères. À l'issue de cela, les éleveurs pastoraux planifient et ajustent l'utilisation des ressources à court et à long terme selon qu'ils soient agropastoraux ou transhumants.

Une considération importante des savoirs pastoraux concerne leur consolidation, et leur transmission (Varga and Molnár, 2014) aux personnes intéressées par le métier ou encore aux éleveurs eux-mêmes en vue d'améliorer les connaissances autour de l'élevage pastoral. Cette consolidation est possible à travers l'hybridation de connaissances scientifiques et de savoirs locaux autour du pastoralisme (Javelle and Jouven, 2015). Actuellement, ces connaissances scientifiques et savoirs locaux sont fragmentairement disponibles auprès des institutions scientifiques et des éleveurs. En effet, d'une part, les connaissances scientifiques autour du pastoralisme, produites par des institutions variées, sont rarement accessibles et utilisées par les éleveurs. D'autre part les savoirs locaux qui sont également multiples, restent partiellement détenus par les éleveurs et souvent inaccessibles à la recherche. Une centralisation de ces deux sources principales de savoirs, contribuera à l'amélioration et à la promotion du pastoralisme en zones méditerranéennes françaises, tant à travers la simplification des activités des éleveurs que par l'amélioration d'un cadre propice à l'utilisation concurrente des espaces et à l'application des dispositifs d'aide.

1.2 Les systèmes agro-sylvo-pastoraux en zones méditerranéennes françaises

Les élevages pastoraux en France sont plus représentés dans la partie méditerranéenne, principalement dans les régions Occitanie, Provence-Alpes-Côte d'Azur (PACA) et la Corse, ainsi que dans les massifs montagneux du sud de la France. Ils sont caractérisés par l'exploitation et la valorisation des milieux aux caractéristiques botaniques et topographiques particulières. L'élevage pastoral réalisé dans ces milieux utilise une variété d'espèces animales herbivores. En plus de cette diversité de milieux et d'espèces animales, s'ajoute une diversité de modes de gestion pastorale qui, ensemble créent une diversité des systèmes agro-sylvo-pastoraux. Dans la suite de ce travail, nous nous focaliseront sur les régions Occitanie et PACA pour lesquelles nous disposons plus de données ayant servi à la réalisation de ce travail.

Les estimations réalisées par (Agrete, 2015) en 2014 sur la base d'analyses de l'occupation et de l'utilisation des sols montrent l'importance des terres de parcours dans ces milieux. En 2018, la répartition spatiale des terres utilisées par le pastoralisme dans le Registre Parcellaire Graphique (RGP, 1.1) a montré leur prépondérance dans ces trois zones méditerranéennes françaises (Nozieres et al., 2021). La disponibilité et la répartition spatiale de ces espaces favorisent l'existence d'une agriculture agro-sylvo-pastorale, caractérisée par la persistance de pratiques de gestion traditionnelles dans ces régions (Bernués et al., 2011).

Une diversité de formations végétales caractérisent ces milieux (CEN-L-R, 2017). La garrigue occupe la grande partie du territoire de l'arrière-pays constitué d'une succession de collines et plateaux calcaires. Le Coussoul, une végétation rase, recouvre le milieu dans la plaine de la Crau caractérisée par une steppe semi-aride. A cheval sur les départements de l'Aveyron, de la Lozère, de l'Hérault et du Gard, on trouve les parcours de pelouses sèches parsemées de buis et de genévrier, les landes à buis qui y croissent lentement mais résistantes à la sécheresse, et les forêts de chênes et de pins. Dans les Préalpes méridionales on trouve de vastes surfaces de pelouses pastorales. Les forêts composées de chênes verts et de châtaigniers occupent les vallées cévenoles en dessous de 900m, et les forêts de Bouleau et de Pin sylvestre au-dessus. En fin, dans les montagnes du pourtour méditerranéen, on trouve un paysage en mosaïque de pelouses, de landes arbustives et d'îlots de forêts. Les pelouses y sont dominés par les fétuques, le Nard raide, la Séslerie bleue, des laïches et le Trèfle alpin.

Trois types d'élevages pastoraux caractérisent ces milieux : i) les transhumants purs qui utilisent significativement les surfaces pastorales d'estives, ii) les sédentaires qui exploitent majoritairement des surfaces pastorales individuelles pour alimenter leurs troupeaux, et iii) les pastoraux transhumants qui combinent les parcours et la transhumance. En région Occitanie sur environ 4767 élevages pastoraux (représentant 23 des élevages herbivores), 26% (soit 1264) sont transhumants purs, 10% (soit 463) sont pastoraux transhumants et 64% (soit

3040) sont sédentaires (Micola and Genevet, 2022). Sur l'ensemble des élevages pastoraux, les proportions en termes d'orientation se présentent comme suit : 36% bovin viande, 20% ovin viande, 10% mixte viande, 10% mixte autres, 8% équin, 5% ovin lait, 3% bovin lait, et 3% caprin lait, et 5% mixte lait (Micola and Genevet, 2022).

Cette diversité de systèmes agro-sylvo-pastoraux soulève un certain nombre de préoccupations, notamment pour les éleveurs, les services de recherche et de vulgarisation. En effet, le fonctionnement des systèmes agro-sylvo-pastoraux est différent des systèmes conventionnels, et leur reconnaissance reste difficile. La difficulté de caractériser précisément les surfaces utilisées par les troupeaux peut rendre difficile l'accès des éleveurs à certaines subventions, notamment la politique agricole commune (PAC) européenne. Une autre préoccupation liée à cette diversité est la difficulté de proposer des solutions technologiques génériques dans une vision d'élevage de précision. Les conditions d'acquisition des informations sont très différentes entre les exploitations pastorales excluant ainsi les différentes offres technologiques disponibles sur le marché (Bocquier et al., 2014). Cette situation engendre le besoin développer des approches qui privilégient l'intégration de la diversité des situations pastorales pour diagnostiquer et documenter l'utilisation des parcours par les troupeaux.

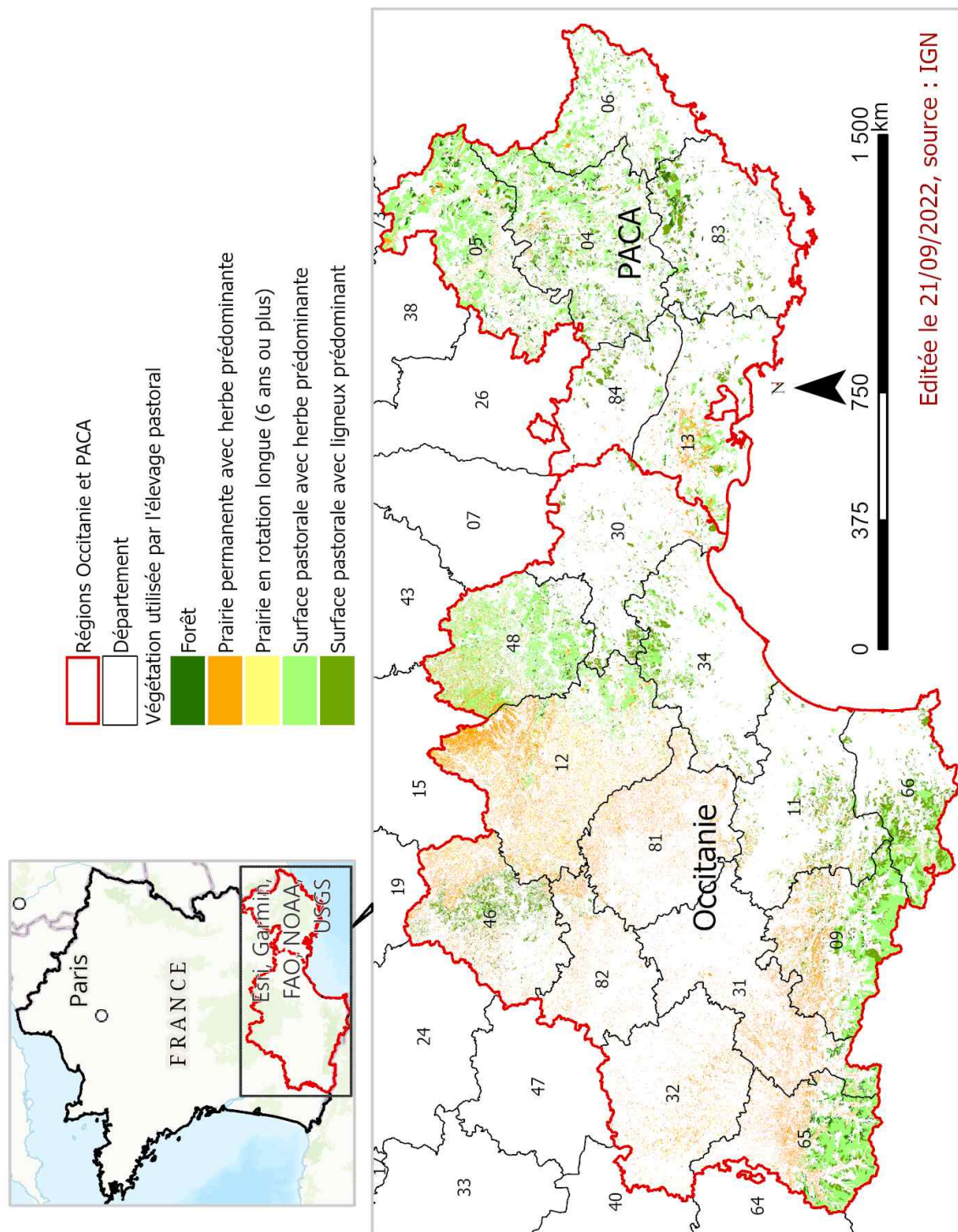


Figure 1.1 – Milieux utilisés par l'élevage pastoral dans les régions Occitanie et PACA. Cette distribution spatiale correspond à la déclaration de 2021 et ne représente pas de manière exhaustive la distribution de ces végétations.

1.3 Enjeux autour du pastoralisme en zones Méditerranéennes françaises

Le pastoralisme en zones méditerranéennes françaises est confronté à divers enjeux. Parmi eux nous pouvons citer : la viabilité économique de l'élevage, l'adaptation au changement et aux variations climatiques, la gestion et la conservation des ressources naturelles et de la biodiversité, et l'accès au foncier.

1.3.1 Viabilité économique du pastoralisme

En effet, à travers la diversité des ressources pastorales des parcours, l'éleveur peut constituer une meilleure ration alimentaire pouvant considérablement améliorer l'ingestion des animaux (Meuret, 2007; Jouven et al., 2019). Cette source d'alimentation disponible à moindre coût pour l'élevage pastoral permet aux éleveurs pastoraux de réaliser des économies d'intrants et d'aliments comparé aux élevages basés sur la production des cultures. En revanche, le pâturage contribue à la valorisation des parcours par leur entretien, notamment des zones embroussaillées (Agreil et al., 2004; Gautier et al., 2011). Cela permet à la fois d'éviter l'emploi des pratiques mécaniques qui peuvent s'avérer inappropriées pour l'environnement et des coûts engagés à leur utilisation.

Par ailleurs, la rentabilité de l'élevage pastoral se trouve affecté' notamment par la crise que traverse l'agriculture française confrontée à la mondialisation des échanges commerciaux (Hubert et al., 2010; Bourdoncle, 2022). Pour pallier à cette pression économique, quelques initiatives ont été établies en vue de susciter l'intérêt du pastoralisme et d'améliorer sa rentabilité. Parmi elles, l'on peut mentionner les démarches qualité concernant les produits animaux, avec pour objectif d'obtenir une plus-value par rapport aux produits standards via l'élaboration de produits de qualité spécifique (Aubron et al., 2014). A cela peuvent s'ajouter des revenus complémentaires issus des activités de diversification. Au-delà de la vente des produits et de la diversification, les primes se rapportant aux aides du premier pilier de la PAC et les prestations relatives à l'ensemble des aides contractualisées du deuxième pilier de la PAC, établies pour rémunérer les impacts positifs de l'élevage pour l'environnement, permettent également d'améliorer les conditions économiques du pastoralisme pour des éleveurs orientant leur système d'élevage conformément à ces dispositions. Cependant, cette dépendance à un dispositif d'aides dont l'avenir est incertain constitue un élément fort de fragilisation de l'activité d'élevage pastoral (Aubron et al., 2019). Il est donc d'intérêt de se pencher vers un système économe qui met plus l'accent sur la valorisation des ressources pastorale afin d'assurer l'autonomie de l'élevage pastoral de manière durable.

1.3.2 Adaptation au changement climatique

L'impact du changement climatique est perceptible en écorégion méditerranéenne et influence de manière considérable l'état des parcours qui subissent les effets liés à des variations de températures et de précipitations. Les simulations réalisées par [Dono et al. \(2016\)](#) révèlent de fortes hétérogénéités saisonnières de températures et de précipitations, avec des températures moyennes maximales et minimales croissantes sur toutes les saisons avec de fortes augmentations observées en été, des précipitations moyennes saisonnières qui diminuent, très faiblement en hiver, et plus fortement au printemps. Cette importante variabilité des paramètres climatiques impacte sur la productivité de la végétation des parcours et donc sur la disponibilité spatiale et temporelle des ressources fourragères. De ce fait, l'adaptation à ces conditions climatiques en plus de l'augmentation de la persistance et de la survie des espèces annuelles et pérennes, ainsi que le maintien de la biodiversité, constitue un défi auquel les acteurs du pastoralisme doivent faire face ([Porqueddu et al., 2016](#)).

Dans ces conditions, les éleveurs pastoraux ont besoin d'ajuster constamment, la conduite du pâturage en fonction des variations dans le temps et dans l'espace à la fois des conditions climatiques et de la disponibilité des ressources pastorales. Cet ajustement requiert un suivi de l'état des ressources pastorales sur parcours et des conditions du milieu ([Guérin et al., 2007](#)), en vue d'entreprendre des stratégies appropriées pour assurer le pâturage. Les éleveurs peuvent ainsi orienter leur troupeaux vers des endroits où la ressource est disponible pour y pâturer. Dans le cas où les ressources fourragères deviennent insuffisantes, ils mobilisent d'autres espaces tels que les prairies, les cultures (vignes et les vergers) à proximité des exploitations pastorales en accord avec les propriétaires. Les apports du foin et des aliments concentrés sont également donnés aux animaux pour assurer leurs besoins en fonction des objectifs de production et de leurs stades physiologiques.

La composition de toutes ces stratégies qui visent à court terme d'adapter la conduite des animaux sous contrainte du climat et de la disponibilité des ressources pastorales, et à long terme le maintien et la préservation des milieux, n'est pas une tâche aisée pour les éleveurs pastoraux. En effet, l'appréhension de toutes ces variabilités sur la végétation et aléas climatiques dont ils dépendent reste un défi. Une approche permettant de rassembler et de simplifier l'ensemble de ces informations permettra de lever cette contrainte au profit du pastoralisme.

1.3.3 Gestion et conservation des ressources naturelles et de la biodiversité

Les parcours des zones méditerranéennes françaises, comme européennes, constituent un important réservoir de biodiversité ([Cowling et al., 1996](#)) et sont caractérisés par une richesse élevée en végétaux ([Lozano and Schwartz, 2005](#)). A cause de l'abandon des terres par les éleveurs, ces parcours se trouvent plus exposés aux risques d'embroussaillage excessif ([Bouju, 2000](#)). Les conditions de sécheresse et de températures élevées pendant l'été,

favorisent la prise des feux (Pereira et al., 2005), pouvant non seulement détruire la biodiversité, mais également provoquer l'érosion des sols mis à nu (Hubert et al., 2010; Shakesby, 2011), et l'atteinte aux animaux et aux humains. Cette situation s'avère catastrophique pour le maintien des mosaïques paysagères et de leurs habitats.

Ces observations ont conduit à encourager l'utilisation effective de l'espace au profit de l'environnement et de la biodiversité en s'orientant vers une réforme des modes de conduite d'élevage tournée vers la valorisation des espaces pastoraux. A ce titre, (Lasanta et al., 2018) ont réalisé une analyse comparative de l'impact du débroussaillage motorisé sur l'intensité des feux de brousse en Espagne et dans la communauté de La Rioja. Ils remarquent une diminution considérable des incendies suite à cette pratique. La conduite du pâturage dans ces milieux peut également permettre d'atteindre ces objectifs d'entretien (Jáuregui et al., 2009; Riedel et al., 2013; Bataille et al., 2016), à travers une exploitation bien planifiée, avec en plus un avantage économique (Jouven et al., 2019) et environnemental. Pour ce faire, une bonne connaissance de l'espace de pâture et particulièrement de la répartition spatiale de la ressource pastorale et du pâturage est idéale (Jouven et al., 2010). Cela permet de mettre au point des stratégies d'exploitation en vue d'optimiser l'utilisation de l'espace à travers le pâturage et par conséquent améliorer les rendements et les bienfaits vis-à-vis de l'environnement et de la biodiversité.

1.3.4 Accès au foncier

La mobilité du pastoralisme renvoie à la notion d'accès au foncier. Il implique les déplacements des troupeaux souvent sur de longues distances (Davies et al., 2017) en transitant temporellement dans des espaces sous différentes modalités d'accès. Plusieurs contraintes relatives à l'accès au foncier peuvent être évoquées (Barrière, 2015) : difficulté de contacter les propriétaires des zones rurales, coûts élevés des terres à proximité des villes, et concurrence avec les usages non agricoles. La nécessité d'élaborer des outils juridiques pour faciliter la collaboration entre propriétaires et éleveurs au profit du pastoralisme s'est faite ressentir depuis quelques décennies. La loi pastorale (Charbonier, 1972) est venue formaliser cette collaboration dans une vision de gestion collective et concertées des espaces pastoraux (Barrière and Bes, 2017). Des outils tels que l'association foncière pastorale (AFP), le groupement pastoral (GP) et la convention pluriannuelle de pâturage (CPP), ont été élaborés pour légaliser et encadrer la gestion collective du pâturage. D'autres outils tels que le pacte pastoral¹ et le plan d'orientation pastoral intercommunal (POPI)² viennent également matérialiser une volonté collective dans l'intérêt général du pastoralisme pour l'ensemble du territoire, et de la coexistence de leurs usages multiples.

La mise en pratique de ces dispositifs juridiques doit passer par leur appropriation par les différentes parties concernées, notamment les propriétaires en vue d'arriver à des accords

1. <https://caussesaignoualcevennes.fr/competences/pacte-pastoral/>

2. <https://cerpam.com/guides-plaquettes-etudes/le-popi/>

visant à promouvoir l'exploitation pastorale. En effet, pour s'engager dans ces actions, une part du défi reste celle de persuader les propriétaires de l'intérêt et des actions positives de ces dispositifs pour le pâturage en vue de les motiver à l'adhésion. Au-delà de l'engagement de différentes parties, il est également utile qu'un système d'échange et de suivi existe en vue de les coordonner à travers le partage de l'information. L'accomplissement de ces actions renforcera la collaboration entre les éleveurs, les propriétaires ainsi que les différents acteurs du territoire, et contribuera à accroître l'impact positif du pâturage par le développement de l'éco-pâturage ou de l'éco-pastoralisme tels que définis par (Eychenne, 2014). Pour y parvenir, il est utile pour les acteurs d'avoir accès à une documentation partagée avec des informations objectives.

1.4 Potentiel des technologies numériques pour la documentation du fonctionnement de l'écosystème pastoral

1.4.1 Quelles informations clés pour contribuer à la documentation de l'écosystème pastoral ?

Les enjeux présentés traduisent la nécessité pour les acteurs du pastoralisme de maîtriser un ensemble d'informations concernant : les espèces animales, la connaissance des milieux et des dynamiques de végétation en lien avec l'action des troupeaux, la gestion de l'interface avec les autres utilisateurs de l'espace, et une connaissance du contexte politique et juridique d'exercice de la pratique pastorale. Ces différentes informations sont très diversifiées de part leurs nature et les sources pouvant les fournir. En fonction de leurs sources, nous pouvons les catégoriser en deux : i) les sources publiques et ii) les sources locales au niveau des exploitations pastorales. D'une part, les informations obtenues des sources publiques sont souvent librement accessibles en France. Leur contribution aux dispositifs de prise de décision en élevage est prouvée (Capalbo et al., 2017; Bahlo et al., 2019). Une part importante de ces informations peut permettre de modéliser des milieux à l'échelle régionale. Nous pouvons citer les données spatiales d'occupation des sols, des parcelles agricoles, de topographie et de sols. Une autre part de cette catégorie d'information concerne les documents techniques produits par des institution de recherche et d'appui technique au pastoralisme. Ce sont des références qui fournissent des orientations pratiques pour la conduite du pâturage et des documents présentant des résultats de recherches autour du pastoralisme. D'autre part, les informations au niveau des exploitations pastorales, peuvent quant à elle être collectées à partir des outils ou des observations des éleveurs pour traduire les phénomènes locaux. Une diversité d'informations peut être collectée sur : l'activité des animaux, les types de fourrages utilisés, les conditions météorologiques, et les techniques de conduite employées.

L'utilisation partielle de ces informations ne permet pas d'aborder de façon intégrale les enjeux liés à l'utilisation des parcours et de constituer une source de référence pour

les acteurs afin de mieux coordonner les activités du pâturage. Leur intégration au sein d'un système d'information (SI) contribuerait à améliorer la compréhension du fonctionnement des systèmes agro-sylvo-pastoraux. En plus, leur combinaison permettrait d'apporter une plus-value (Capalbo et al., 2017) pour la gestion et le pilotage du pâturage sur parcours. En effet, l'utilisation des approches disciplinaires combinées comprenant les systèmes d'information géographiques, les réseaux de capteurs et les protocoles standards, favoriserait la production d'informations utiles pour la prise de décision (Laca, 2009; Schellberg et al., 2008; Antle et al., 2017). Dans di Virgilio et al. (2018), les auteurs présentent un outil qui permet de produire des indicateurs du comportement des animaux et d'évaluer l'influence de différents facteurs du milieu sur la production animale, suite à la combinaison des données spatiales. Dans Terrasson et al. (2017), les auteurs présentent une plateforme expérimentale intégrant les données spatiales de l'élevage dans un but de suivre le pâturage des animaux. Cependant, l'opérationnalisation de ces types d'outils reste à tester en les confrontant aux conditions réelles de production. En plus, le contexte sur le quel se base la plupart des travaux de recherche présente moins d'hétérogénéité par rapport aux conditions d'élevage pastoral en zones méditerranéennes françaises.

En outre, ces informations sont variées et hétérogènes de par leurs sources (publique ou privée), leurs natures (données spatiales et non spatiales, structurées et non structurées, texte, images, etc.) et leurs formats. Leur intégration pour le développement d'outils de prise de décision pour l'élevage extensif présente actuellement un défi pour la recherche. Le problème de leur intégration se réfère notamment à la gestion de l'interopérabilité et des métadonnées (Bahlo et al., 2019) en vue de faciliter l'utilisation et le partage par les parties prenantes. En plus de ce défis d'intégration, il est nécessaire de développer des méthodes adaptées de croisement de ces informations afin de produire de nouvelles informations utilisables par les éleveurs au profit de l'élevage pastoral en zones méditerranéennes françaises. Ceux-ci constituent les principales questions auxquelles notre travail vise d'apporter une contribution.

1.4.2 **Projet P@stor_All**

Le projet P@stor_All financé par la Région Occitanie et ayant débuté le 01/03/2021 a été initié dans but le de valoriser, les données, les connaissances, et les fiches techniques à travers leur intégration dans un système d'information utile pour les parties prenantes. Il permet de mettre en relation éleveurs, techniciens, ingénieurs et chercheurs au service du développement des technologies numériques en élevage pastoral. Le projet s'appuie sur les développements récents, notamment du Projet CLOChÈTE relatifs au numérique en élevage pastoral pour co-concevoir un système d'information capable de traiter des données numériques hétérogènes en vue d'apporter des réponses documentées, et mieux comprendre et gérer les écosystèmes pastoraux. Ce projet s'insère dans le cadre des activités de l'UMT Pasto qui est une structure de partenariat entre L'INRAE comme institut de recherche,

L'institut Agro-Montpellier et l'institut technique d'Élevage. L'UMT Pasto a pour objectifs de : i) produire des connaissances et des méthodes pour accompagner le maintien et le développement des élevages, ii) faciliter la concertation pour favoriser la coordination des acteurs du pastoralisme et des actions de recherche et développement à l'échelle nationale et internationale, iii) et contribuer à la formation des acteurs de l'élevage et de la gestion des milieux.

Le projet P@stor_All a réuni différentes parties prenantes : 7 éleveurs pastoraux, 2 fermes expérimentales de l'UMT Pasto, des ingénieurs et techniciens pastoraux et en informatique, ainsi que les enseignants chercheurs des unités mixtes de recherche TETIS et SELMET spécialisées respectivement dans les domaines de géomatique et du pastoralisme en région méditerranéennes. Plusieurs rencontres d'échange et de co-construction ont été réalisées dans le but de mener le travail sur la base des besoins réels et exprimés par les parties prenantes, notamment les éleveurs.

Notre thèse s'est adossée sur le volet d'intégration et de mise en valeur des informations hétérogènes du projet P@stor_All pour contribuer à la définition et conception d'un système d'information spatial pour le pastoralisme.

1.5 Objectifs et intérêts de la thèse

Dans cette thèse, l'objectif principal que nous poursuivons est de définir et de co-concevoir un système d'information spatial permettant l'intégration et les croisements des données hétérogènes du pastoralisme en zones méditerranéennes françaises, dans le but de produire des informations pertinentes pour la prise de décision en élevage pastoral et d'établir une base de connaissances autour du pastoralisme.

Plus spécifiquement il s'agit de :

- Définir une approche méthodologique permettant d'identifier les besoins des éleveurs pastoraux et les données hétérogènes du pastoralisme, et concevoir un système d'information (SI) adapté pour leur intégration et leur analyse ;
- Collecter et préparer les différentes données hétérogènes permettant de caractériser l'utilisation des parcours par les troupeaux pour leur intégration dans le SI ; et
- Définir les méthodes et techniques d'analyse croisée des informations issues des données hétérogènes identifiées, en vue de proposer des bouquets d'information utiles à la prise de décision pour la conduite du pâturage.

Notre thèse contribue à la proposition des méthodes d'analyse et des indicateurs pastoraux qui permettent de poser les bases pour alimenter le système d'information développé

dans le cadre du projet P@stor_All. Nous avons expérimenté les méthodes et les outils développés dans d'autres domaines d'étude, en vue de proposer des solutions techniques et des indicateurs pouvant contribuer à la prise de décision dans le cadre de l'élevage pastoral en régions Occitanie et PACA.

Chapitre 2

Conduite du pâturage sur parcours : panorama des contributions d'outils numériques

Après avoir situé le contexte et les enjeux du pastoralisme en zones méditerranéennes et montagneuses françaises, dans ce chapitre 2, nous présentons un aperçu sur les applications des technologies numériques en élevage de précision et les opportunités offertes par les initiatives en régions Occitanie et PACA autour des données numériques et de l'implication de différents acteurs. Ensuite nous présentons les éléments de la problématique de notre thèse autour de la question centrale du développement d'un outil de documentation et de prise de décision dans le cadre de l'élevage pastoral.

2.1 Élevage et pâturage de précision : concept et définition

Dans un contexte de demande croissante en produits alimentaires, notamment en protéines animales, la production animale s'accroît progressivement depuis quelques années. [Ritchie et al. \(2017\)](#) indiquent que la production animale a quadruplé depuis les 50 dernières années. Avec le changement des habitudes alimentaires des personnes cette augmentation s'élèverait à environ 70% en 2050 ([Berckmans, 2017](#)). Cet accroissement se réalise par l'augmentation de la taille des troupeaux au sein des exploitations alors que le nombre d'éleveurs décroît. On aboutit ainsi à des exploitations avec de troupeaux beaucoup plus importants par éleveur et une main d'œuvre réduite. La gestion de telles exploitations fait face à de nouveaux enjeux de production, environnementaux, socio-culturels, de santé et de bien-être animal auxquels les éleveurs doivent faire face. En plus de ces enjeux, l'adaptation au marché global représente un important défi. Avec cette augmentation et les enjeux associés, les éleveurs ne peuvent plus assurer de manière effective la gestion de gros troupeaux d'animaux. Le concept d'élevage de précision a été proposé pour répondre à ces enjeux de gestion d'élevage et assurer une production conforme à la demande, l'environnement et le bien-être animal ([Berckmans, 2017](#)).

L'élevage de précision est hérité de l'agriculture de précision où le fondement de base est l'intégration du numérique pour adapter les pratiques à l'hétérogénéité spatiale des sols, des plantes et de l'environnement. Le but étant d'ajuster les interventions en fonction des besoins spécifiques de chaque plante ([Pierce and Nowak, 1999](#)). Elle offre des possibilités de réaliser la bonne action, au bon endroit, au bon moment et de la bonne manière ([Pierpaoli et al., 2013](#)), réduisant ainsi les coûts engagés et l'impact sur l'environnement. L'élevage de précision se base sur l'intégration d'outils technologiques pour la collecte, la transmission, l'analyse et la diffusion des informations permettant aux éleveurs de décider des actions à mener ([Zhang et al., 2002](#); [Pierpaoli et al., 2013](#); [Serrano et al., 2013](#)). Dans le cadre de l'élevage de précision, l'objectif est d'intégrer le numérique en vue d'ajuster finement la conduite du pâturage en fonction de la diversité des conditions des milieux et des comportements des troupeaux dans le processus de production, afin d'accroître la productivité, réduire l'impact environnemental ([Laca, 2009](#)) et assurer la santé et le bien-être animal.

L'élevage de précision se fonde sur le fait que les animaux sont complexes, individuellement différents, variant et dynamiques. De ce fait, il s'appuie sur le suivi individuel des animaux quel que soit leur effectif. Les opérations faites sur les animaux sont différenciées par unité du troupeau (individu ou lot), sur la base de mesures automatisées qui ont été transmises puis intégrées dans un système d'information (SI) ([Bocquier et al., 2014](#)). Les traitements sont individualisés dans le but de maximiser les performances de chaque animal et d'optimiser la production. L'individualisation du suivi impose l'équipement en capteurs de tous les individus ou des lots du troupeau dans un environnement pouvant être bien contrôlé. C'est ainsi que l'élevage de précision s'est d'abord développé en élevage industriel de poules et de cochons, puis ceux de vaches et chèvres laitières à haut rendement, donc à forte valeur ajoutée et

bonne capacité d'investissement (Meuret et al., 2013). Deux exemples typiques de ces formes d'élevage de précision peuvent être le distributeur automatique d'aliments concentrés (DAC) pour des animaux laitiers identifiés individuellement par collier électronique, et le robot de traite qui mesure la composition du lait, le rythme cardiaque, la température corporelle et l'activité physique de chaque animal (Wathes et al., 2008; Jago et al., 2013). Désormais en France et dans plusieurs pays de la Méditerranée, l'équipement d'une boucle électronique RFID sur les ruminants d'élevage est obligatoire pour réaliser l'identification et la traçabilité des animaux (Bocquier and Jouven, 2017)

Contrairement aux systèmes de production confinés et intensifs, dans les systèmes extensifs associés à de grands troupeaux qui se dispersent sur de vastes surfaces pastorales souvent hétérogènes, le suivi des animaux individuellement est difficile voire impossible pour des raisons pratiques et de coûts. Le pâturage est dominant dans ces exploitations avec des objectifs de production modérés. L'autonomie alimentaire dans ces systèmes dépend directement de la capacité à ajuster l'utilisation d'une diversité des ressources pâturées (souvent spontanées) par un troupeau hétérogène. Cet ajustement permet d'assurer le besoin alimentaire des animaux tout en gérant l'utilisation de l'espace pour favoriser le renouvellement de la ressource pâturée. Ainsi, ces exploitations se basent beaucoup plus sur le pilotage du pâturage des troupeaux ou des lots d'animaux plutôt que sur des traitements individualisés.

Pour aborder les enjeux liés à la gestion du pâturage, Laca (2009) propose le pâturage de précision. Il s'inspire des concepts de l'élevage de précision, sans s'attarder au suivi individuel des animaux tout en considérant les difficultés d'évaluation des ressources pastorales. L'enjeu est d'ajuster, au fil du temps et en fonction des aléas, la diversité des ressources alimentaires offertes par les pâturages, à la diversité des besoins du troupeau (Bocquier et al., 2014). Dans le cadre de l'élevage pastoral sur les parcours méditerranéens et montagneux en France, le principal objectif serait de valoriser la grande hétérogénéité spatiale et temporelle des ressources spontanées des parcours afin de parvenir à des performances animales et environnementales élevées. Pour y parvenir, de nouvelles approches basées sur l'intégration et l'analyse de la diversité des données spatiales doivent être abordées. Pour l'élevage pastoral la donnée spatiale prend tout son sens du fait des espaces vastes et de la mobilité des troupeaux (même pour les troupeaux dits "sédentaires"). Ces approches permettront de contribuer à la documentation du fonctionnement des écosystèmes agro-sylvo-pastoraux et par conséquent à la gestion du pâturage. Une telle solution peut se construire sur base des outils numériques existants tout en les adaptant pour répondre aux principaux enjeux liés à la nature des données du pastoralisme (Bocquier et al., 2014).

2.2 Contribution des technologies numériques dans le cadre l'élevage de précision et opportunités pour la conduite du pâturage en régions Occitanie et PACA

2.2.1 Applications en élevage de précision

Les applications d'élevage de précision sont variées en fonction des systèmes d'élevage, des plus intensifs à ceux utilisant les milieux pâturés pour élever les troupeaux. Le principe central autour des systèmes d'élevage de précision repose sur : i) la collecte de l'information sur l'environnement et sur les animaux à partir des capteurs appropriés, ii) la transmission de l'information collectées vers un système de gestion des données grâce aux technologies de l'information et de la communication, iii) des applications de diagnostic sur l'environnement ou l'état des animaux en vue d'orienter la prise de décision. La prise de décision consiste en une adaptation des conditions de l'environnement de l'élevage et des interventions sur les animaux en vue d'optimiser la production. De nombreux équipements électroniques ont été développés dans le cadre de l'élevage de précision. Un exemple d'outil électronique le plus largement utilisé est le lecteur RFID (Radio Frequency Identification) qui enregistre l'identité de l'animal porteur d'une puce RFID et donne accès à plusieurs applications telles que la traçabilité des animaux et le tri sur terrain de gros troupeaux d'animaux lorsque le lecteur RFID est équipé à un dispositif de tri automatisé (Bocquier and Jouven, 2017). Le processus de tri permet en plus de différencier les interventions sur les animaux.

Les systèmes d'élevage intensifs ou en bâtiment se prêtent mieux aux objectifs d'élevage de précision, étant donné la facilité à maîtriser les variables de leur environnement (température, humidité, luminosité) et des animaux (poids, mouvement, bruit, etc.) (Bocquier et al., 2014). Une revue non exhaustive des applications autour des technologies d'élevage de précision en systèmes intensifs est présentée dans le Tableau 2.1. Dans les élevages peu intensifiés avec de grands troupeaux d'animaux herbivores utilisant des espaces pâturés, d'autres objectifs peuvent être visés. Nous pouvons citer entre autres : i) la gestion du cycle de reproduction dont dépend la productivité animale, ii) la contention et le tri de grands troupeaux d'animaux qui détermine la productivité du travail, iii) les traitements antiparasitaires qui s'imposent lorsque les animaux pâturent, mais dont la systématisation pose des problèmes sanitaires et environnementaux, iv) la géo-certification des modes d'élevage, et v) la gestion du pâturage (Bocquier et al., 2014). Ces objectifs variés imposent l'utilisation d'outils numériques adaptés à chaque contexte spécifique.

Table 2.1 – Exemples d'applications faisant appel aux technologies d'élevage de précision développées dans le cadre des élevages intensifs. Source : [Banhazi et al. \(2012\)](#)

N°	Technologie / outils d'élevage de précision
1	Amélioration des incubateurs d'œufs par synchronisation des éclosions
2	Contrôle intelligent de la ventilation dans les bâtiments d'élevage
3	Estimation du poids des porcs à l'aide d'outils de vision artificielle
4	Gestion des produits laitiers pour maximiser le profit
5	Amélioration de la rentabilité grâce à l'alimentation de précision pour les porcs
6	Robot de placement de capteurs pour porcs
7	Système de surveillance de bovins
8	Surveillance de la santé de la mamelle et de l'hygiène chez les bovins laitiers
9	Inspection des carcasses de volaille
10	Comptage et identification automatisés des œufs
11	Prédiction de la composition des carcasses de porcs
12	Calibrage et triage automatisés du poisson
13	Amélioration du contrôle thermique des porcs par vision artificielle
14	Reconnaissance de la toux chez les porcs

Dans le cadre de la gestion de la reproduction nous pouvons citer les travaux de [Bocquier \(2006\)](#); [Maton et al. \(2008\)](#) sur l'utilisation d'un détecteur électronique de chevauchements placé sur le mâle en vue d'identifier les femelles en chaleur. De tels dispositifs permettent d'identifier les moments propices à la reproduction animale. Les résultats produits peuvent ensuite être utilisés par les éleveurs pour réaliser les inséminations.

En ce qui concerne la contention et le tri, le lecteur RFID a déjà fait ses preuves ([Bocquier and Jouven, 2017](#); [Maton et al., 2008](#)). Il permet de réaliser l'inventaire et la différenciation des animaux. Le tri des animaux se réalise à travers un couloir de tri dynamique automatisé sur base des critères définis à l'avance par l'éleveur (poids vif, état corporel, âge, date de mise bas, appartenance à un lot précis) ou à travers les mesures au passage ([Maton et al., 2006](#)). Les interventions peuvent ensuite être différenciées en fonction des objectifs des éleveurs. Ces processus peuvent permettre d'appliquer diverses interventions sur les animaux en fonction des informations collectées : séparation des animaux dans parcelles disposant des ressources alimentaires différenciées pour améliorer l'état de certains animaux ([Maton et al., 2008](#)), ajustement des traitements antiparasitaires uniquement aux animaux en perte de poids, en évitant le surdosage ([Cabaret, 2004](#)). L'ajustement des traitements antiparasitaires se base sur le principe selon lequel les animaux parasités accusent une perte de poids. Plusieurs types d'application différenciées peuvent être facilités par les dispositifs de tri.

L'idée de la géo-certification serait de localiser les endroits valorisés par les animaux, surtout dans le cadre de l'élevage pastoral où les animaux se déplacent en fonction des périodes de l'année. Elle permettrait aux éleveurs d'être éligibles aux primes environnementales sur base des endroits valorisés. Pour ce faire, GPS et lecteur RFID peuvent être conjointement utilisés, embarqués sur les animaux. La certification se ferait alors par des inventaires géolocalisés selon les intervalles de temps établis par le cahier des charges ([Bocquier, 2006](#)).

Enfin, la gestion du pâturage représente un aspect important qui détermine la réussite

dans l'élevage des troupeaux en milieu extensif. En effet, les troupeaux se déploient de manière hétérogène dans les milieux qu'ils pâturent en fonction de l'état et la distribution des ressources alimentaires, de la topographie et des conditions météorologiques (Bailey and Provenza, 2008; Manteca i Vilanova and J. Smith, 2014; Bailey et al., 2015). En fonction de l'influence de ces facteurs environnementaux et de leur comportement, les troupeaux pâturent plus certains endroits que d'autres. Cela conduit au surpâturage et au sous pâturage localisés pouvant être néfastes tant pour les performances zootechniques que pour les communautés végétales (Jouven et al., 2010). Dans ce contexte, l'intérêt du pâturage de précision porte sur la capacité des outils numériques à contribuer à l'ajustement de la conduite du pâturage à travers une bonne valorisation de l'hétérogénéité spatiale et temporelle des ressources pastorales pour assurer l'alimentation des troupeaux. Ainsi, le succès dans l'application de ces outils dépendra de leur capacité à décrire les variables du milieu et du comportement des animaux, en vue de hiérarchiser les facteurs qui orientent la distribution spatiale et les choix alimentaires des troupeaux.

Avec les avancées de la technologie spatiale, les données issues de capteurs aéroportés et de stations météorologiques peuvent permettre de bien modéliser le milieu utilisé. Pour le suivi et l'analyse du comportement des animaux, les balises GPS et les accéléromètres ont été plus utilisés dans la recherche (Buerkert and Schlecht, 2009; Akasbi et al., 2012; Feldt and Schlecht, 2016). Les travaux de (Caja et al., 2020; Vaintrub et al., 2021) présentent et discutent une revue des solutions numériques et de leurs applications dans les systèmes agro-sylvo-pastoraux sud européens. Les accéléromètres triaxiaux ont été plusieurs fois intégrés dans des études pour déterminer les instants de repos, de pâturage et de rumination des animaux (Terrasson et al., 2016; Alvarenga et al., 2020; Fogarty et al., 2020). Ces comportements peuvent être associés aux conditions du milieu en vue d'analyser et de hiérarchiser leur influence (Rutter et al., 1997; Handcock et al., 2009). Les GPS ont également été expérimentés pour plusieurs applications : quantification de la durée passée par les ovins dans différentes zones du pâturage (Rutter et al., 1997), caractérisation de l'utilisation des végétations naturelles par différents types d'herbivores (Putfarken et al., 2008), identification de la trajectoire effectuée par les caprins pendant le pâturage (Buerkert and Schlecht, 2009), et évaluation de la capacité de stockage de différents pâturages. Certaines expérimentations telle que celle de Wang (2019) présentent les possibilités de développer des applications de prédiction de comportement des troupeaux. D'autres encore ont développé des applications de clôture virtuelle pour restreindre les troupeaux à pâturer au sein des zones bien définies (Monod et al., 2008; Anderson, 2007; Terrasson et al., 2017)

Dans les conditions des milieux très hétérogènes tels que les parcours des zones méditerranéennes et montagneuses françaises utilisées par l'élevage pastoral, la caractérisation de ces différents facteurs reste un défi. En effet, les outils et les moyens utilisés pour modéliser les petites zones ne peuvent pas être utilisés dans des milieux plus larges et hétérogènes. Dans ces milieux variables, le comportement des animaux est également susceptible de varier beaucoup. Dans ces conditions, les principales contributions devraient provenir des approches

permettant de lever les contraintes associées à l'hétérogénéité des milieux pastoraux et la variabilité des comportement des troupeaux en utilisant les données et les moyens les moins coûteux.

2.2.2 Enjeux associés au développement d'un outil de pâturage de précision dans le cadre du pastoralisme en zones méditerranéennes françaises

Les élevages pastoraux utilisent les ressources des végétations spontanées des milieux très variés et hétérogènes spatialement et temporellement du fait de la grande variabilité des facteurs environnementaux (topographie, sol, végétation et paramètres météorologiques). En plus de ces conditions de parcours variés, les formes d'élevages elles-mêmes varient également d'un éleveur à l'autre et d'un territoire à un autre (différentes espèces, pratiques pastorales et orientations de production). Le développement d'outils de prise de décision dans ce contexte devient de plus en plus problématique à cause des contraintes imposées par cette hétérogénéité et les réponses non linéaires qu'ils provoquent dans le comportement les animaux.

Dans une perspective d'élevage de précision, la disponibilité des données est une condition primordiale pour le développement d'un outil d'aide à la décision pour l'élevage pastoral. Dans le contexte de l'utilisation des parcours des zones méditerranéennes, les conditions d'acquisition des données sont très diversifiées conduisant à des données de qualité hétérogène pour la modélisation des écosystèmes agro-sylvo-pastoraux. En plus de l'hétérogénéité des données, la caractérisation du comportement des troupeaux à partir des données disponibles n'est pas aisée en élevage pastoral. En effet, la modélisation du comportement animal prend une importance fondamentale dans la compréhension et la conception de dispositifs permettant la maîtrise du comportement des animaux (Bocquier et al., 2014). Les travaux réalisés sur l'utilisation de données numériques pour caractériser le comportement des animaux sont plus expérimentales. L'application des méthodes développées en situations opérationnelles reste un défi. Compte tenu des situations particulières en élevage pastoral (milieux fermés, météorologie particulière), la précision des données GPS peut être impactée et rendre difficile la caractérisation du comportement des troupeaux. En plus, les éleveurs ne hiérarchisent pas de la même manière l'influence des facteurs limitant dans l'utilisation des parcours par les troupeaux. Ces principaux enjeux de modélisation des milieux pastoraux et du comportement des animaux doivent être abordés à partir de nouvelles approches permettant de palier aux contraintes issues de la diversité des conditions des parcours et des données.

L'hétérogénéité des milieux, des données et des savoirs d'éleveurs pastoraux, rend difficile la mise en place d'un SI permettant de répondre de manière générique aux différentes préoccupations des éleveurs. En effet, la diversité des situations pastorales complique l'interprétation des données, et ne permet pas d'automatiser la prise de décision car selon les cas les mêmes indicateurs pourront être interprétés différemment. L'intérêt pour la conduite du pâturage sur parcours serait de privilégier des approches simples et parcimonieuses, allant

moins loin dans l'analyse mais cumulant une plus grande diversité d'indicateurs. Selon les cas particuliers des élevages certains indicateurs seront plus pertinents que d'autres. Du point de vue de la gestion des données, il est important de que le SI soit en mesure d'assurer : i) le stockage et l'intégration des données hétérogènes et l'accroissement en volume, ii) l'interopérabilité des données et la gestion des métadonnées pour une meilleure gestion du SI, et iii) l'implémentation des méthodes appropriées pour l'analyse et le croisement des informations en vue fournir des indicateurs d'aide à l'élevage pastoral (Capalbo et al., 2017; Bahlo et al., 2019). Un SI qui réunit ces caractéristiques pourra servir de référentiel pour la documentation de l'élevage pastoral en zones méditerranéennes françaises à travers l'intégration d'informations sur les situation variées des d'élevage.

2.3 Émergence des données spatiales et opportunités pour le développement d'un système d'information pour le pastoralisme en régions Occitanie et PACA

Les régions Occitanie et PACA se trouvent favorisées par l'émergence d'actions autour du pastoralisme. Plusieurs initiatives ont été menées pour améliorer les pratiques pastorales dans ces deux régions, allant des aspects techniques à l'expérimentation d'outils technologiques pour améliorer l'utilisation des parcours. Ces initiatives ont donné lieu à la production de données intéressantes pour documenter l'écosystème pastoral, en plus des données publiques.

Tout d'abord en ce qui concerne les sources publiques, la disponibilité croissante de données satellitaires gratuites de meilleure résolution spatiale, spectrale et temporelle, telles que les données Sentinel 2 (10 m dans le visible et le proche infrarouge et 20 m et 60 m de résolution spatiale ; 13 bandes spectrales et une période de revisite de 5 jours), et les algorithmes de traitement associés, favorise la caractérisation de l'occupation des sols à de larges échelles territoriales (Inglada et al., 2017). D'autres types de données spatiales peuvent être associés, comme le Registre Parcellaire Graphique (RPG), la base de données forestières et la base de données altimétriques, produites par l'Institut National de l'Information Géographique et Forestière (IGN) pour mieux caractériser les parcours. Les travaux de Inglada et al. (2017); Shaqura and Lasseur (2019) se sont intéressés à la complémentarité de ces données pour cartographier l'occupation et l'utilisation des sols dans les territoires français de la métropole. Avec les avancées de la technologie les caractéristiques de ces données pourront davantage être améliorées et permettront de produire des informations encore plus précises. Depuis la loi pastorale de 1972 (Charbonier, 1972), différents services en France interviennent dans l'appui technique aux exploitations d'élevage : les chambres d'agriculture¹, les instituts techniques, les associations d'acteurs intéressés par le pâturage et le pastoralisme et les services pastoraux. Ils participent à l'élaboration des références techniques sous forme de manuels

1. <https://chambres-agriculture.fr/>

et plus récemment sous forme numérique (cf. IDELE²) et des outils d'aide à la gestion des pâturages. Ces références fournies par les différents services peuvent être mobilisées et alimenter un SI pour améliorer la gestion du pâturage sur parcours.

Les instituts techniques contribuent également à la production d'importantes références techniques sous forme de manuels et plus récemment sous forme numérique (cf. IDELE³). Ces références proposent des outils pratiques pour différents aspects de la conduite du pâturage.

L'UMT Pasto formée depuis 2015 contribue à la production d'informations à travers des projets de recherche au sein de trois sites expérimentaux : i) le domaine de la Fage situé sur le Causse du Larzac, comportant un élevage ovin viande sédentaire de 300 brebis Romanes élevées en plein air intégral sur pelouses sèches, landes à buis et prairies, ii) le domaine du Merle, dans la plaine de la Crau, comportant un élevage ovin viande transhumant de 1500 brebis Mérinos d'Arles conduites sur prairies et pelouses sèches de Crau, garrigues et alpages, et iii) la ferme expérimentale de Carmejane, dans les Alpes de Haute Provence, comportant un élevage ovin viande transhumant de 700 brebis Préalpes conduites sur prairies, parcours boisés et alpages. Différents travaux réalisés dans ces sites ont déployé des capteurs embarqués permettant de suivre l'activité des animaux et d'étudier leur comportement. C'est le cas du projet CasDAR CLOChèTE⁴ qui a utilisé les GPS et les accéléromètres pour localiser et déterminer le comportement des ovins et caprins. Les données produites par ces outils constituent une source d'information de grande importance pour caractériser l'activité des animaux. D'autres part, comme le reste de la population (INSEE⁵), les éleveurs se sont massivement équipés en smartphone, et quelques uns utilisent le GPS mobile pour suivre leurs troupeaux.

Suite au besoin de créer des cadres d'échange entre acteurs du pastoralisme, des réseaux de partage d'expériences se sont développés. Nous pouvons mentionner le réseau Pâ-tur'Ajuste⁶, créé en 2013 à l'initiative de la SCOP de formation, de conseil et d'accompagnement SCOPELA. Ce réseau réunit un collectif d'éleveurs, de conseillers agricoles, de chargés de mission environnemental, d'enseignants et de chercheurs. Son objectif est de constituer à l'échelle nationale, un lieu d'échange, de capitalisation et de diffusion des expériences sur la valorisation des végétations naturelles dans l'élevage. D'autres groupes informels sont aussi formés pour faciliter l'échanger entre éleveurs via les réseaux sociaux tels que Facebook ou WhatsApp.

La disponibilité de ces différentes sources de données offre aujourd'hui une opportunité pour documenter le fonctionnement de l'écosystème agro-sylvo-pastoral, et contribuer à améliorer la conduite du pâturage en zones méditerranéennes françaises. Cela peut être rendu

2. <https://idele.fr/outils>

3. <https://idele.fr/outils>

4. <https://idele.fr/clochete/>

5. <https://www.insee.fr/fr/statistiques/6036909>

6. <https://www.paturajuste.fr/>

possible grâce à une mobilisation et une meilleure intégration de ces données hétérogènes au sein d'un SI approprié. Ce SI devra être en mesure d'offrir la possibilité de stocker ces données hétérogènes, de les traiter et de les analyser pour fournir aux acteurs et particulièrement les éleveurs, les informations utiles en faveur de la conduite du pâturage.

2.4 Problématique et questions de recherche

La rentabilité économique et les services écosystémiques issus des activités pastorales constituent deux enjeux majeurs de l'élevage pastoral en zone méditerranéennes françaises. L'élevage pastoral est avant tout une activités génératrice de revenu pour les éleveurs. Cependant, la crise que traverse l'agriculture française associée à la mondialisation des échanges commerciaux (Hubert et al., 2010; Bourdoncle, 2022) constitue une contrainte à l'obtention des revenus significatifs pour les éleveurs pastoraux. Même si les dispositifs tels que les démarches qualité et la vente directe concernant les produits animaux et les activités de diversification peuvent contribuer à rehausser les revenus des éleveurs, ils ne permettent pas d'améliorer la productivité de l'élevage pastoral. En plus, les démarches qualité ne garantissent pas une place importante du pâturage dans l'alimentation (Aubron et al., 2014). La dépendance aux aides de la PAC dont l'avenir est incertain peut également contribuer à la fragilisation l'activité d'élevage pastoral (Aubron et al., 2019). Dans un contexte d'élevage basé sur l'utilisation des parcours variés et hétérogènes, les exploitations pastorales devraient se tourner vers une autonomie de la conduite du pâturage pour améliorer la productivité propre à l'élevage pastoral. Cette autonomie repose sur la prise en compte des aspects de productivité et de préservation des services écosystémiques de l'élevage pastoral. Les solutions techniquement efficaces devraient se tourner vers la valorisation de l'hétérogénéité spatiale et temporelle des ressources pastorales des parcours et la minimisation des coûts engagés dans la production.

La valorisation de l'hétérogénéité spatio-temporelle des parcours nécessite une double connaissance : sur les parcours et leurs dynamiques spatio-temporelles, et sur les besoins des animaux en fonction des objectifs de production. Cette connaissance peut être possible s'il existe des sources d'information utiles pouvant permettre de décrire les milieux pastoraux et le comportement spatial des troupeaux qui varie en fonction des conditions des milieux pâturés. Avec les conditions qui favorisent la disponibilité des données en régions Occitanie et PACA (données publiques, références techniques, données expertes d'éleveurs, données issues des capteurs animaux), il est possible d'explorer les potentialités de ces données à documenter la conduite du pâturage sur parcours à travers la mise en place d'un SI dans une vision d'élevage de précision.

Les avancées technologiques ayant rendu possible le développement de nombreux outils de collecte, de transmission et de gestion de données représentent aujourd'hui une opportunité pour l'élevage de précision, notamment l'élevage pastoral. Cette forte variabilité des

conditions pastorales doit être prise en compte dans le processus de développement d'un outil d'élevage de précision dans le cadre du pastoralisme. Les milieux pastoraux peuvent ainsi être décrits grâce aux données spatiales provenant des satellites et les comportements des troupeaux peuvent être décrits à travers les données de leur localisation. Plus encore les références techniques et les retours d'expérience d'éleveurs peuvent aussi être collectées. Ces informations peuvent ainsi être rassemblées dans un SI capable de les gérer, de fournir des fonctionnalités d'analyse, et de proposer des indicateurs qui vont aider la prise de décision en élevage pastoral. Cela pourrait également aider les chercheurs à identifier des thèmes et objets clés pour la recherche appliquée, voire à formuler des hypothèses à tester ensuite dans des projets de recherche.

Bien que les outils numériques et les données soient disponibles, l'un des principaux défis pour une transition vers l'élevage de précision dans le cadre du pastoralisme reste la diversité et la variabilité spatio-temporelle des milieux pastoraux. Elles rendent difficile par exemple le développement de modèles de classification générique pouvant permettre de prédire les classes pastorales. D'autre part, elles favorisent la variabilité des réponses chez les animaux et donc une variabilité des comportements des troupeaux. Ces conditions imposent des méthodes d'analyse spécifiques à chaque contexte d'exploitation pour l'élevage de précision. Les indicateurs produits par ces méthodes seront en fait variés et spécifiques à chaque contexte d'étude. D'où l'intérêt de définir les méthodes permettant de produire plusieurs indicateurs pouvant être plus ou moins représentatifs de la variété des conditions pastorales dans notre zone d'étude. Ces conditions limitent également l'automatisation de l'interprétation et de la décision. L'interprétation des indicateurs produits et la prise de décision devront être associées à l'expertise des éleveurs qui ont une connaissance des milieux et des conditions de pâturage de leurs troupeaux.

La diversité des données oblige également d'adopter de nouvelles approches de leur gestion. Les différentes sources de données identifiées sont de diverses natures et formats. En plus de l'utilité de conserver leur état pour des analyses ultérieures, ces données hétérogènes nécessitent un SI adéquat pour leur intégration. Une fois le SI défini, ces données disparates pourraient être rassemblées et analysées ultérieurement pour accompagner les éleveurs dans leur prise de décision. Les indicateurs produits à partir de l'analyse des données GPS contribueront au fur et à mesure à la compréhension du comportement spatial des troupeaux sur parcours. D'autres données complémentaires provenant des smartphones et des données libres spatialisées pourront conforter la compréhension du comportement des troupeaux. Enfin, les références techniques pourront être utilisées comme source d'information sur les pratiques de conduite du pâturage sur parcours.

Pour aborder les aspects développés autour de l'élevage de précision en zones méditerranéennes françaises, nous avons analysé trois principales questions de recherche en basant notre étude sur les régions Occitanie et PACA :

— Quelles méthodes et techniques permettent de définir les besoins des éleveurs pasto-

raux, d'identifier les données pastorales et de définir le SI adapté à leur intégration ?

- Quelles approches pertinentes doivent être abordées pour prétraiter et préparer les données pastorales en vue de faciliter la production d'indicateurs pastoraux ?
- Quelles méthodes d'analyse sont favorables à la production d'indicateurs d'aide à la prise de décision dans le cadre de l'élevage pastoral ?

Dans les chapitres 3, 4 et 5, nous traitons ces questions de recherche dans le but de poser une base d'analyse pour la production d'indicateurs d'aide à la prise de décision en élevage pastoral. Les approches développées dans ce travail seront par la suite implémentées dans le système d'information développé dans le cadre du projet P@stor_All dont l'objectif est de rassembler toutes les informations identifiées pour servir de référence et de source d'indicateurs pour l'élevage pastoral.

Chapitre 3

Approche de co-construction d'un système d'information pour l'intégration des données du pastoralisme

Dans le précédent chapitre nous avons présenté un état de l'art sur les outils d'élevage de précision et les perspectives pour leur intégration dans le cadre du pastoralisme en zones méditerranéennes françaises. Dans le présent chapitre 3, nous abordons l'approche de co-construction d'un système d'information adapté à l'élevage pastoral dans ces milieux. Nous intégrons l'ensemble des parties prenantes (éleveurs pastoraux, chercheurs et techniciens) dans ce processus.

3.1 Introduction

La gestion de l'élevage pastoral en zones méditerranéennes françaises est confrontée à différents enjeux : viabilité économique (di Virgilio et al., 2018), gestion et conservation des ressources naturelles et de la biodiversité (Bailey et al., 2019), adaptation au changement climatique (Dono et al., 2016), et accès au foncier (Barroso et al., 2000). Pour l'élevage pastoral, ces enjeux sont directement liés au pâturage et en particulier à l'utilisation de l'espace par les troupeaux. Pour faire face à ces enjeux, de nouvelles approches nécessitant l'intégration d'un ensemble d'informations permettant de caractériser les dynamiques au sein des écosystèmes agro-sylvo-pastoraux doivent être considérées. Ces informations se réfèrent à différents facteurs qui impactent le comportement alimentaire et spatial des animaux ainsi que la gestion du pâturage. Ces facteurs décrivent les espèces animales, les milieux explorés, et les pratiques de gestion. Une modélisation de l'écosystème agro-sylvo-pastoral permettra de mieux comprendre les liens existant entre ces facteurs et d'identifier les informations pouvant être intégrées dans un système de prise de décision.

Ces différents facteurs peuvent être décrits à partir d'une diversité de données provenant de sources publiques ou locales grâce au développement des technologies de l'information et de la communication. D'une part, les données spatiales publiques telles que les images satellites (par exemple Sentinel 2 ou SPOT 6/7), associées à des données de référence peuvent être utilisées pour décrire les milieux pastoraux utilisés (Shaqura and Lasseur, 2019; Castro et al., 2020). D'autre part, les données locales (par exemple GPS et retours d'éleveurs) collectées au niveau des exploitations pastorales, peuvent contribuer à la compréhension du comportement spatial des animaux sur parcours (Handcock et al., 2009; Akasbi et al., 2012; Feldt and Schlecht, 2016). Diverses données non spatiales (exemple documents numérisés) peuvent aussi être utilisées, pour fournir des informations pratiques pour la conduite de l'élevage.

Ces données publiques et locales sont hétérogènes en termes de nature (spatiales, non spatiales, structurées ou non structurées), de format et de résolution. Pour répondre aux enjeux du pastoralisme, il est nécessaire que les informations liées aux différents composants des systèmes agro-sylvo-pastoraux soient intégrées dans un référentiel adéquat. Cependant, l'hétérogénéité des données pose un défi pour leur intégration. Les systèmes de gestion de données actuels utilisent plus les données structurées et gèrent difficilement les données semi et non structurées. Par ailleurs, la diversité de formats et de métadonnées des données spatiales des sources publiques et locales, suscite en plus un défis d'interopérabilité (Bahlo et al., 2019). Un système d'information (SI) adapté devra donc être en mesure de lever ces contraintes en vue d'assurer le stockage continu de données hétérogènes, l'accès et l'utilisation par différents groupes d'utilisateurs.

Au-delà de ces considérations d'intégration, un regard doit également être porté sur la valeur des informations que procurent les données. Prises individuellement, ces informations

sont souvent peu adaptées pour la caractérisation des processus au sein des exploitations pastorales, et pour la prise de décision en élevage (O'Grady and O'Hare, 2017; Bahlo et al., 2019). La valeur ajoutée se trouve en effet dans leurs combinaisons. Afin de produire des informations pertinentes pour le suivi et l'ajustement à court et à long terme de la conduite du pâturage, des méthodes appropriées de croisement de ces informations doivent donc être intégrées (Bahlo et al., 2019) dans le SI. De plus, l'implication de différents acteurs du pastoralisme dans la mise en place du SI est nécessaire pour répondre aux enjeux du pastoralisme et de susciter l'engagement pour le développement d'un tel outil.

Dans ce chapitre nous présentons les différentes étapes de définition et de co-construction d'un système d'information spatial, capable d'intégrer les données hétérogènes du pastoralisme en zones méditerranéennes françaises ainsi que les composants de traitement, d'analyse et de visualisation des informations pour la gestion de l'élevage pastoral. Nous présentons dans un premier temps dans la section 3.2 la modélisation de d'un écosystème agro-sylvo-pastoral. Puis dans la section 3.3 nous identifions les différentes sources de données permettant de documenter cet écosystème, et présentons un aperçu sur l'association de données. Dans la section 3.4 nous présentons l'approche nous ayant permis de définir les fonctionnalités du système d'information ensemble avec les parties prenantes. Le choix de l'architecture du SI approprié est présentés dans la section 3.5. Enfin le système d'information à proprement parler est présenté dans la section 3.6.

3.2 Modélisation de l'écosystème agro-sylvo-pastoral

Le comportement alimentaire et spatial des animaux sur parcours est le résultat de l'effet combiné de plusieurs facteurs du milieu et propres aux animaux. Il détermine les choix opérés par les animaux pour explorer la variété des milieux pastoraux et influe directement sur leur nutrition, leurs performances zootechniques, leur santé et leur impact sur le couvert végétal (Ginane et al., 2008). La modélisation de l'écosystème agro-sylvo-pastoral permet de représenter et de comprendre les différents facteurs qui agissent sur le comportement alimentaire et spatial des animaux au pâturage.

Dans cette partie du travail, nous montrons l'approche ayant permis d'identifier et de caractériser les différents facteurs du comportement alimentaire et spatial des animaux sur parcours, et présentons un modèle de l'écosystème agro-sylvo-pastoral représentatif des zones méditerranéennes françaises.

3.2.1 Approche méthodologique

La modélisation de l'écosystème pâturé a été réalisée à travers un processus d'identification des facteurs clés du comportement alimentaire et spatial des animaux (cf. tableau 3.1). Nous nous sommes basés sur les résultats du travail réalisé par Djibo Saibou (2020), dans

le cadre du projet P@stor_All dans le but de caractériser les facteurs qui impactent le comportement alimentaire et spatial des animaux sur parcours. Ce comportement est caractérisé par : le déplacement, la distribution spatiale et le choix alimentaire des animaux. Trois étapes ont été réalisées durant ce travail : i) tout d'abord, une analyse de la littérature scientifique a permis d'identifier globalement les facteurs du comportement des animaux au pâturage, ii) ensuite une validation puis une complémentation de ces facteurs a été réalisée par le biais d'enquêtes individuelles auprès des éleveurs, et iii) une analyse finale a permis de dresser la liste des facteurs influençant spécifiquement le comportement des troupeaux d'animaux au pâturage. Nous présentons une synthèse de ce processus à la Figure 3.1.

La grande partie des références utilisées correspond à des conditions différentes du contexte de notre étude (en termes d'environnements, parfois de types d'animaux, et dans des conditions expérimentales) en raison de la disponibilité limitée d'études scientifiques dans les zones montagneuses et méditerranéennes du sud de l'Europe sur le thème du comportement alimentaire des animaux. La vision des éleveurs a été intégrée par Djibo Saibou (2020) pour se rapprocher du contexte de l'élevage en zones méditerranéennes françaises (tant géographique qu'en termes d'usages). Ce processus a permis de valider, compléter et établir une liste de facteurs qui valorisent les connaissances scientifiques abordées tout en les restituant dans un contexte (en termes de vocabulaire et de représentation) habituel pour les éleveurs pastoraux dans notre zone d'étude.

L'analyse de la littérature scientifique a permis d'identifier différents facteurs regroupés en 4 catégories concernant : les ressources pastorales, les animaux du troupeau, l'environnement du pâturage, et le mode de gestion du pâturage (Figure 3.1). Les entretiens avec les éleveurs pastoraux ont montré que certains facteurs identifiés dans la littérature n'étaient pas suffisamment cohérents avec leur vision du système d'élevage et les paramètres utilisés pour le gérer.

Table 3.1 – Résumé des facteurs qui influencent le comportement alimentaire et spatial des animaux au pâturage.

Facteurs issus de la littérature scientifique	Influence sur le comportement des animaux : Déplacement, occupation de l'espace et choix alimentaire
Ressources pastorales	
Disponibilité	Les animaux se déplacent et occupent les endroits où la ressource est disponible (Chebli et al., 2020).
Accessibilité	Les animaux explorent les ressources végétales accessibles. Les endroits non accessibles (embroussaillés) sont évités (Gautier et al., 2011).
Distribution spatiale	Les animaux occupent l'espace en fonction de la distribution des ressources pastorales accessibles (Wang et al., 2009a).
Environnement	
Topographie	La pente est le principal facteur limitant. Les zones à fortes pentes sont évitées par les animaux, en fonction des espèces (Bailey et al., 2015).
Météorologie	Les fortes températures d'été modifient le mode de pâturage des animaux. Ils développent le pâturage de nuit et se reposent le jour. En cas de fortes pluies, et d'absence de visibilité du aux brouillards, les animaux diminuent le déplacement et se regroupent (Manteca i Vilanova and J. Smith, 2014).
Points d'attraction	Les animaux se regroupent souvent autour des points d'attraction (point d'eau, mangeoires, ...) (Poilecot et al., 2013).
Animaux / Troupeau	
Physiologie	Les choix alimentaires (qualité et quantité ingérées) des animaux sont déterminés par leur état physiologique (lactation, gestation). Ex : l'ingestion est élevé en période de lactation (Manteca i Vilanova and J. Smith, 2014).
Morphologie	Détermine l'accessibilité des animaux aux ressources pastorales. Les caprins peuvent explorer les endroits non accessibles pour les ovins. La qualité et la quantité des fourrages consommée peut également varier selon l'espèce (Shipley, 1999).
Mémoire spatiale	Les animaux mémorisent les endroits visités et savent y retourner pour pâturer. Ils gardent également en mémoire les endroits non intéressants qu'ils évitent. Ils varient donc l'occupation de l'espace (Laca, 2009).
Relations sociales	Les interactions (dominance, leadership, affinité) entre les animaux impactent sur les trois composants du comportement : réduction du déplacement, regroupement ou dispersion pendant le pâturage, ... (Barroso et al., 2000).
Conduite du pâturage	
Mode de conduite	La complémentation influe sur l'appétit et modifie le choix alimentaire des animaux. L'utilisation des parcs clôturés contraint les animaux à pâturer la végétation des espaces réduits. La présence du berger oriente le déplacement, les choix alimentaire, et l'occupation de l'espace par les animaux (Rochon and Goby, 2009).

La plupart des facteurs de la littérature retenus pour le modèle conceptuel de l'écosystème agro-sylvo-pastoral ont été sélectionnés en fonction du nombre de fois où ils ont été validés par les éleveurs. Les facteurs validés par l'ensemble des éleveurs (en termes de leur vocabulaire) sont *l'abondance, la qualité et la diversité des ressources pastorales, la présence de l'éleveur et la chaleur*. Ces facteurs correspondent à la disponibilité des ressources pastorales,

aux pratiques de gestion et aux conditions météorologiques dans la littérature scientifique. Certains facteurs concernant les troupeaux (*taille du troupeau, connaissance de l'environnement, race animale et espèces animales*) ont été principalement issus du point de vue des éleveurs. Bien que les éleveurs n'aient pas validé l'impact de l'accessibilité des ressources pastorales, nous avons également conservé ce facteur car il peut être dû à l'embroussaillement des massifs de ligneux qui caractérisent les parcours montagneux et méditerranéens.

L'unification des facteurs scientifiques validés et des facteurs complémentaires issus des éleveurs a permis de constituer la liste finale des facteurs clés du comportement alimentaire et spatial des troupeaux au pâturage dans un contexte adapté aux conditions de parcours montagneux et méditerranéens. Ces facteurs ont ensuite été utilisés pour modéliser l'écosystème agro-sylvo-pastoral et représenter leurs différents liens sans aborder l'effet de leurs interactions.

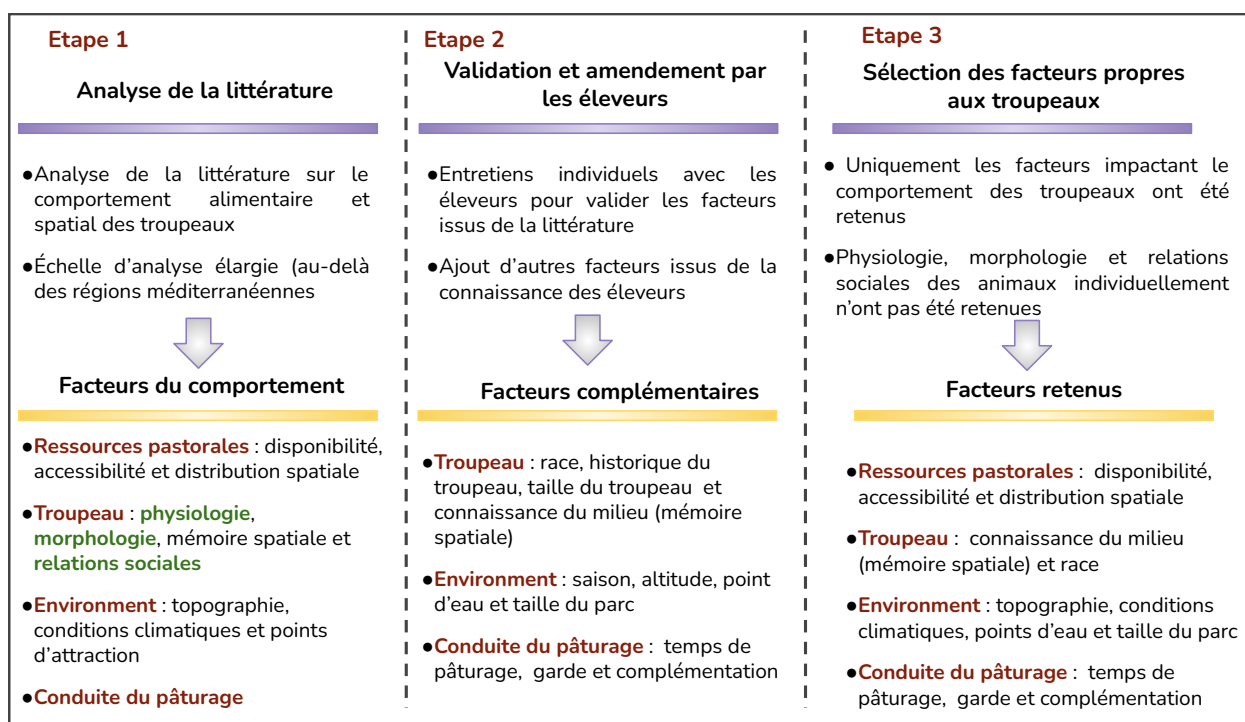


Figure 3.1 – Différentes étapes réalisées pour identifier les facteurs clés influençant le comportement des troupeaux au pâturage. Les facteurs en couleur verte n'ont pas été retenus à la fin du processus.

3.2.2 Modèle du comportement alimentaire et spatial des animaux

Nous présentons dans la figure 3.2 le modèle conceptuel de l'écosystème agro-sylvo-pastoral dans les conditions des parcours montagneux et méditerranéens français. L'échelle de notre étude concerne le troupeau. Ainsi, seuls les facteurs qui agissent directement sur le comportement alimentaire et spatial des troupeaux ont été pris en compte. Ceux qui ont des effets sur l'individu (animal) n'ont pas été retenus. C'est le cas pour la morphologie, la

physiologie et les relations sociales des animaux.

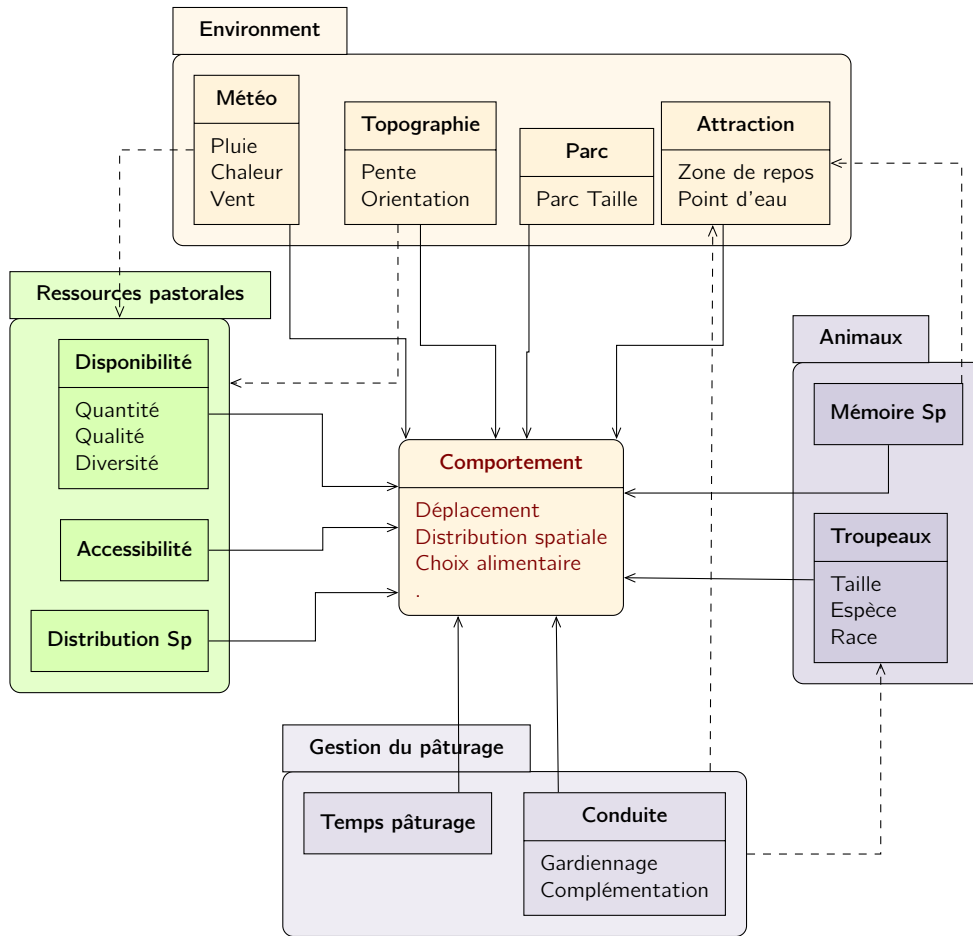


Figure 3.2 – Le modèle conceptuel du comportement alimentaire et spatial des herbivores domestiques des systèmes agro-sylvo-pastoraux en zones méditerranéennes et montagneuses françaises. La ligne continue représente l'impact d'un facteur sur le comportement des Troupeaux et la ligne discontinue représente l'impact entre les facteurs.

3.3 Sélection des données pour la documentation de l'écosystème agro-sylvo-pastoral

La modélisation de l'écosystème agro-sylvo-pastoral a permis de présenter les différents facteurs qui influencent le comportement alimentaire et spatial des animaux. Une représentation de ces facteurs est utile en vue de caractériser et de hiérarchiser leur influence sur le comportement des animaux. Différentes sources de données peuvent être utilisées à cet effet. Pour atteindre les objectifs de cette étude, nous avons visé la disponibilité de données de précision acceptable, individuellement ou en combinaison avec d'autres données. L'indisponibilité des données ou le manque de précision de certaines des données disponibles peut être surmonté par des informations complémentaires provenant des éleveurs.

Les différentes données identifiées peuvent être classées en deux types en fonction de leurs

sources : (i) les données publiques, fournies par les institutions publiques, et (ii) les données locales qui peuvent être collectées au niveau de chaque exploitation pastorale à partir d'outils appropriés.

3.3.1 Données publiques

Les données publiques identifiées concernent les images Sentinel 2, le Registre Parcellaire Graphique (RPG), l'occupation des sols OSO, les données altimétriques de l'IGN, et les références techniques sur le pâturage. A l'exception des références techniques, toutes les données externes sont fournies sous licence ouverte Etalab 2.0¹. Cette licence accorde un droit non exclusif et gratuit de réutilisation des informations à des fins commerciales ou non pour une période illimitée. Les caractéristiques de ces données et leurs usages sont présentées dans le tableau 3.2.

Table 3.2 – Résumé des caractéristiques et des usages des données externes pouvant être utilisées pour modéliser l'utilisation des parcours par les troupeaux

Données externes	Caractéristiques
Image Sentinel 2 ²	Description : image satellite au format TIFF Source : Theia Disponibilité : 5 jours pour les niveaux 1C et 2A, 1 mois pour le niveau 3A Domaine spectral et résolutions spatiales : 13 bandes de 10 m, 20 m et 60 m Fauchée : 290 km Contribution : Utilisation et couverture des sols
Registre parcellaire graphique (RPG) ³	Description : base de données des parcelles agricoles au format shp ou gpkg Source : Agence française de services et de paiement Disponibilité : annuelle Contribution : référence pour la cartographie des classes pastorales
Occupation des sols OSO ⁴	Description : Occupation des sols avec 23 classes Résolution spatiale / unité minimale de cartographie (vecteur) : 10 m/20 m Source : CNES / IGN Disponibilité : annuelle Contribution : description et cartographie des terres de parcours
MNT RGE ALTI ⁵	Description : Modèle de relief Source : IGN Résolution spatiale : 1 m à 5 m en X et Y Disponibilité : selon la mise à jour Contribution : Modélisation du relief du terrain
Références techniques ⁶	Description : documents décrivant les pratiques pastorales Source : instituts techniques Disponibilité : à chaque parution Contribution : gestion du pâturage

1. <https://www.etalab.gouv.fr/licence-ouverte-open-licence>

3.3.2 Données locales

Les données locales sont celles pouvant être collectées au niveau des exploitations pastorales. Elles concernent principalement les données de localisation des troupeaux à l'aide du GPS et le retour d'informations des éleveurs. Elles constituent une source importante de données car elles contribuent à la compréhension des dynamiques au sein des exploitations pastorales qui constituent l'échelle d'analyse de notre travail.

Données de localisation des troupeaux par GPS

Les données de localisation des troupeaux par GPS sont variées. Certains éleveurs, fermes expérimentales et chercheurs ont commencé à utiliser le GPS pour analyser le comportement de leurs troupeaux. Ces données se présentent sous différentes formes en fonction de la variété des équipements GPS utilisés et peuvent être utilisées pour différentes analyses. Elles peuvent être utilisées pour analyser les déplacements des troupeaux et leur comportement quand à utilisation des ressources des parcours. Plusieurs études ont montré le potentiel des données GPS pour modéliser le comportement des troupeaux au pâturage. L'activité des troupeau (Buerkert and Schlecht, 2009; Akasbi et al., 2012; Feldt and Schlecht, 2016; McGavin et al., 2018), le comportement des troupeaux et leur interaction avec les zones de pâturage (Putfarken et al., 2008; Handcock et al., 2009), et le temps de pâturage (Rutter et al., 1997), sont des exemples d'applications issues des données de localisation des troupeaux. Différentes méthodes analytiques peuvent être appliquées à ces données afin de produire de nouvelles informations précieuses pour modéliser l'utilisation des espaces pastoraux par les troupeaux, en combinaison avec d'autres sources de données.

Informations provenant des d'éleveurs

Le niveau de précision fourni par les données mentionnées peut présenter des limites pour décrire plus en détail l'écosystème agro-sylvo-pastoral. Pour fournir des informations complémentaires précises, les informations provenant des éleveurs peuvent être d'un grand intérêt. Outre la précision des informations, ces données peuvent également concerner d'autres informations concernant la gestion du pâturage et le comportement des animaux observés sur le terrain, que les données publiques ne peuvent pas procurer.

L'acquisition des données d'éleveurs peut se faire via des applications collaboratives sur smartphones pour collecter des informations sur les ressources pastorales, le sol, le climat, l'ingestion des animaux, ou toute autre information concernant la gestion du pâturage. Il existe plusieurs applications pour réaliser ces acquisitions de manière adaptée, permettant leur analyse dans un SI. Nous pouvons citer l'exemple d'applications smart phones de cartographie

et de navigation telle que OSMAnd⁷, Organic Maps⁸ qui proposent des services gratuits et hors ligne, pouvant permettre de collecter et d'enregistrer des informations spatialisées sur milieux explorés (végétation, milieu ou conditions particuliers) pouvant alimenter le SI. D'autres applications telles que Kobo collect issues de la plateforme Kobotoolbox⁹, peuvent permettre de collecter divers types d'information sur le milieu exploré, les animaux et la conduite du pâturage.

3.3.3 Association des données pour modéliser l'écosystème agro-sylvo-pastoral

Les données présentées (cf. Tableau 3.2) permettent d'identifier les analyses de base qui peuvent être réalisées pour caractériser les facteurs du comportement des animaux au pâturage. Dans le Tableau 3.2, nous présentons les différents facteurs du comportement alimentaire et spatial des troupeaux et les données identifiées qui peuvent être utilisées pour les caractériser. La caractérisation des facteurs peut être possible par une combinaison de différents types d'informations. Les composants du comportement des troupeaux que sont le déplacement, la distribution spatiale et les choix alimentaire peuvent également être modélisés à partir des données identifiées ou de leur combinaison. Cela peut permettre de produire des informations supplémentaires concernant les interactions entre les animaux, les ressources pastorales, l'environnement et la gestion du pâturage.

Toutes les données ne permettent pas de caractériser les facteurs du comportement des animaux au pâturage (voir les cases vides dans le tableau). Cependant, d'autres croisements entre les différents types d'informations sont susceptibles d'être identifiés a posteriori pour améliorer la caractérisation de certains facteurs. Les éleveurs peuvent également apporter des informations beaucoup plus précises (ressources pastorales consommées, complémentation, conditions de pâturage, particularités des milieux explorés) pour compléter les autres données et améliorer la caractérisation de l'utilisation spatiale des milieux par les troupeaux.

7. <https://osmand.net/>

8. <https://organicmaps.app/>

9. <https://www.kobotoolbox.org/>

Table 3.3 – Les facteurs du comportement alimentaire et spatial des troupeaux, leurs composants et les données disponibles permettant de les caractériser.

Facteurs	Données externes/publiques				Données de l'exploitation	
	Sentinel 2	RPG	OSO	MNT RGE	Herd location	Feedback
Ressources pastorales						
Disponibilité	X	X	X			X
Accessibilité	X	X	X	X		X
Répartition de l'espace	X	X	X			
Environnement						
Météorologie						X
Topographie				X		
Paddock						X
Point d'attraction						X
Animal/Troupeau						
Mémoire Sp			X		X	
Caractéristiques du troupeau					X	
Gestion du pâturage						
Temps de pâturage					X	X
Alimentation						X
Composantes du comportement alimentaire						
Mouvement				X	X	
Occupation de l'espace	X	X	X	X	X	

3.4 Définition des fonctionnalités du système d'information

La définition des fonctionnalités du système d'information a été réalisée à travers un processus de co-construction avec les différents acteurs du pastoralisme que sont les éleveurs pastoraux, deux sites expérimentaux en élevage pastoral, les chercheurs et enseignants et l'institut technique de l'élevage. La disponibilité de ces parties prenantes a permis d'orienter l'identification des besoins conformément aux réalités des systèmes agro-sylvo-pastoraux représentés par les 7 élevages privés et 2 sites expérimentaux (cf. tableau 3.4). Cela a permis également de les motiver pour l'engagement à l'élaboration du SI. En effet, l'engagement des parties prenantes est un facteur non négligeable qui conditionne le succès de la construction d'un outil de prise de décision (Wilson et al., 1997; Bekker and Long, 2000). Ce processus a été regroupé en 3 phases (cf. figure 3.1) : (i) entretiens individuels semi-directifs avec les éleveurs pour déterminer leurs besoins sous forme de questions, (ii) discussions collectives et validation des besoins des éleveurs avec les parties prenantes, et (iii) Sélection des questions finales à traiter dans le SI. La première phase a été réalisée à travers un stage de deuxième année dans le cadre du projet P@stor_All.

Table 3.4 – Les 7 exploitations privées et 2 sites expérimentaux agro-sylvo-pastorales enquêtés pour la constitution des besoins

Exploitation	Production	Nb mères	Espaces pâturés	Complémentation
1	Ovin et caprin viande	300	Parcours et vigne	Aucun
2	Caprin lait	40	Parcours	Foin de Crau
3	Ovin viande	140	Parcours et vigne	Aucun
4	Ovin viande	280	Parcours et prairie	Aucun
5	Ovin lait	550	Parcours et culture	Foin et grain
6	Ovin et caprin viande	123	Parcours et prairie	Foin et orge
7	Caprin lait	-	-	-
Le Merle	Ovin viande	1670	Parcours et prairie	Aucun
La Fage	Ovin viande	300	Parcours et vigne	Fourrage et concentré

Les entretiens individuels ont permis de recueillir des questions de 6 éleveurs par rapport à la conduite et la gestion du pâturage, les milieux et les ressources pastorales utilisés et la santé des animaux. Au terme de cette phase 51 questions ont été retenues et ont constitué la base pour les discussions avec l'ensemble des parties prenantes lors des rencontres collectives. A travers les rencontres collectives, les questions des éleveurs ont été passées en revue, reformulées et fusionnées. La fusion a concerné les questions similaires formulées différemment par les éleveurs lors des entretiens. Au terme de ces travaux, 38 questions ont été retenues (tableau 3.5), pour servir de base aux fonctionnalités du SI. Nous les avons réparties en 3 groupes selon leur nature : (i) celles qui s'intéressent aux repères permettant d'ajuster le pâturage, (ii) celles qui concernent les méthodes pour atteindre un objectif donné, et (iii) celles qui permettent de comprendre l'écologie du système pastoral. Le SI sera en mesure d'intégrer une série d'informations pouvant être utilisées pour fournir aux éleveurs de nouvelles informations répondant à leurs besoins formulés sous forme de question dans le tableau 3.5. Cependant, la possibilité de répondre aux questions retenues dépend à la fois de la disponibilité de données adéquates et des méthodes de traitement appropriées pour produire de nouvelles informations.

Enfin, nous avons sélectionné 9 questions finales pour constituer les besoins auxquels les SI permettra d'apporter des réponses à travers l'intégration et l'analyse des données. La sélection de ces questions tient à la fois sur les critères de priorité pour la conduite du pâturage et la possibilité d'y apporter des réponses à travers les moyens techniques et méthodologiques à disposition.

Table 3.5 – Besoins des éleveurs exprimés sous forme de questions, considérés comme une base pour les fonctionnalités du SI.

N°	Questions
Repères pour ajuster la conduite des pâturages	
1	Quelle est la surface pastorale utilisée actuellement ?
2	Quel trajet mon troupeau emprunte-t-il chaque jour ?
3	Combien de temps mon troupeau passe-t-il au pâturage ?
4	Quelle est la superficie occupée par le troupeau ?
5	Comment mes animaux utilisent-ils l'espace (par semaine, par mois, par saison, par an) ?
Méthodes pour atteindre un objectif donné	
6	Comment valoriser les zones embroussaillées ?
7	Comment valoriser la ressource pastorale pour répondre aux besoins de mon troupeau ?
8	Comment inciter les animaux à se déplacer dans les zones inexplorées ?
Mieux comprendre l'écologie de l'écosystème pastoral	
9	Comment le comportement des animaux évolue-t-il en fonction de la saison ?

3.5 Choix du système d'information pour l'intégration des données du pastoralisme

Les données présentées sont très variées en termes de nature, de structure, de format, et de précision. En fonction de leur nature, nous classons deux types de données : les données spatiales et les données non spatiales. Les données spatiales permettent de localiser les milieux ou les exploitations, ou les phénomènes qui s'y produisent. C'est le cas de l'occupation des sols OSO, du RPG, de la localisation des troupeaux, des images Sentinel 2 et du MNT. Les données non spatiales concernent principalement les documents des références techniques. En fonction de leurs structures, nous avons les données structurées et non structurées. Les données structurées concernent toutes les données spatiales vectorielles OSO (shp), RPG (shp) et GPS (GPX, shp, gpkg) et les retours d'éleveurs sous forme de formulaires. Les données non structurées concernent quant à elles les documents des références techniques (PDF) et les données rasters (TIFF ou GRID). On peut également identifier la variété de formats que portent ces données.

Dans cette sous section, nous présentons un état de l'art sur les systèmes de gestion des données et le choix porté sur le lac de données spatiales pour la gestion et les croisements des données pour l'utilisation par les acteurs du pastoralisme, notamment les éleveurs.

3.5.1 État de l'art sur les systèmes de gestion de données

Les systèmes de gestion des données ont évolué depuis le développement des systèmes de gestion de bases de données relationnelles (SGBDR) puis et des bases de données NoSQL (Not only SQL) aux systèmes décisionnels plus élaborés tels que les entrepôts de données

de plus en plus utilisés dans des entreprises. Les SGBDR sont utilisés par les entreprises depuis les années 1970 et sont actuellement une technologie mature pour stocker et gérer les données relationnelles. Avec le développement de l'informatique et du WEB on assiste à une production de gros volumes de données dans plusieurs domaines. Ces données provenant entre autres des réseaux sociaux, des institutions médicales et commerciales sont de natures variées et non essentiellement structurées dans des tables. Les SGBDR conçus pour gérer les données structurées ont trouvé leur limite face à cet avènement des données massives, notamment à cause du volume important et de la variété des données. Ces limites des SGBDR ont suscité au sein des entreprises le besoin de développer de nouvelles manières de gérer les données. On assiste ainsi à un développement des systèmes de gestion de données massives (McAfee and Brynjolfsson, 2012), les systèmes de gestion des données NoSQL (Corbellini et al., 2017; Davoudian et al., 2019) mis au point pour répondre aux enjeux de gestion des données massives non uniquement relationnelles.

Les entrepôts de données ont été conçus comme système de prise de décision utilisant des multiples sources de bases de données relationnelles (Favre et al., 2013). Leurs modèles conceptuels se basent sur des croisements de tableaux de données suivant les processus analytiques choisis. Ainsi, ils ingèrent les données selon un schéma prédéfini au préalable dit "Schema-on-Write", dans une démarche d'ETL (Extract-Transform-Load) où les données doivent d'abord subir une transformation avant d'être utilisées. La grande diversité des données non structurées actuellement disponibles, rend difficile l'utilisation des entrepôts de données idéalement conçus pour gérer les données structurées. Même si certains travaux ont abordé l'utilisation conjointe des systèmes NoSQL et des entrepôts de données (Chevalier et al., 2017; Boussahoua et al., 2017), l'intégration au sein d'un seul système d'information d'un ensemble de données hétérogènes reste un défi.

Le concept de lac de données a été introduit par Dixon (2010) pour répondre aux différents défis imposés par les données massives et dont la variété ou l'hétérogénéité est l'une des caractéristiques majeures, avec le volume et la vitesse. Les processus d'intégration dans un lac de données privilégient la variété des données et la conservation de leurs états en vue de favoriser leur exploration par différentes communautés d'utilisateurs. Dans Madera (2019), l'auteur indique que là où un entrepôt de données pousse les données de leur origine vers leurs consommateurs selon un chemin relativement fixe où chaque magasin de données est censé satisfaire un besoin, on a dans un lac de données une bien plus grande flexibilité. C'est en effet à chaque utilisateur de matérialiser son besoin et d'extraire les différentes données sources puis de les combiner pour produire de nouvelles informations.

En se référant à Hai et al. (2016); Mathis (2017) et Gorelik (2019), nous pouvons définir le lac de données comme un système évolutif permettant de stocker de grands volumes de données structurées, semi-structurées et non structurées sans subir des traitements de normalisation à l'entrée, disposant des fonctionnalités permettant leurs requêtes, leur gestion et leurs analyses pour produire des résultats ou services opérationnels, à travers son utilisation par différentes communautés selon leurs besoins. Ces caractéristiques du lac de données sont

intéressantes pour répondre aux enjeux d'hétérogénéité et de valeur des données identifiées. Cependant, les lacs de données sont en cours de maturation et nécessitent des implémentations aux niveaux de différents composants, notamment ceux qui concernent la gestion des métadonnées spatiales et les modules analytiques. Pour les besoins du pastoralisme il est donc important de définir les considérations architecturales et fonctionnelles adaptées en vue de répondre aux enjeux spécifiques à la conduite du pâturage en zones méditerranéennes françaises.

Dans le tableau 3.6 qui suit, nous présentons un résumé de comparaison entre le lac et l'entrepôt de données.

Table 3.6 – Comparaison entre lac de données et d'un entrepôt de données. Sources : Fang (2015); Khine and Wang (2018)

Dimension	Entrepôt de données	Lac de données
Données	Structurées et semi-structurées	Structurées, semi-structurées, non structurées, données brutes et non traitées
Stockage	Couteux	Moins coûteux
Agilité	Moins agile, configuration fixée	Très agile, configuration flexible
Sécurité	Mature	En cours de maturité
Charge de travail	Entre des centaines et des milliers d'utilisateurs concurrents. Interface analytique performée	Traitement en lot pour des données à grande échelle. En cours d'amélioration des capacités à supporter des interactions importantes des utilisateurs
Schéma	Définis avant le stockage des données (Schema On Write).	Définis après que les données ont été stockées (Schema On Read). Requiert du travail à la fin du processus
Echelle de données	Gros volumes de données à des coûts modérés	Volumes de données extrêmes à faibles coûts
Requête	SQL	Programmation
Complexité	Jointures complexes	Traitements complexes
Accès	Via les requêtes SQL et les outils de BI standardisés	Via des programmes créés par des développeurs

3.5.2 Critères de choix et de l'implémentation du SI

Quelques contraintes ont été prises en compte pour le choix du SI adapté aux besoins du pastoralisme en zones méditerranéennes françaises :

- La possibilité d'offrir les fonctionnalités d'intégration, de traitement et de croisement des données hétérogènes identifiées.
- La capacité de mieux gérer l'accès aux données par les différents utilisateurs à travers un système de gestion de métadonnées adapté.
- La capacité de mise à échelle pour gérer l'augmentation en volume des données et l'intégration de nouvelles fonctionnalités analytiques.

Pour répondre à ces considérations, nous avons opté pour une implémentation du lac de données Hadoop d'Apache. Nous portons notre choix sur Hadoop pour deux aspects essentiels : (i) il s'agit d'une implémentation centrale des technologies de big data, et (ii) la plupart de distributions qui reposent sur le fondement d'Hadoop sont libres d'accès et améliorées par de grandes communautés d'utilisateurs.

Hadoop est basée sur deux composants essentiels : HDFS (Hadoop Distributed File System) et MapReduce. HDFS (Honnutagi, 2014) assure le stockage distribué et parallélisé des données lorsqu'elles sont ingérées dans le lac de données. La structure d'HDFS offre plusieurs avantages, notamment (i) la facilité d'ingestion et de récupération des données à partir de leurs métadonnées et (ii) la parallélisation du stockage et du traitement, ce qui apporte plus de performance pour la sécurité des données et améliore significativement le coût de traitement. Outre le système de stockage, HDFS est flexible en ce qui concerne la configuration avec d'autres moteurs de traitement. Cela permet de créer un écosystème qui offre une gamme complète d'outils de stockage, d'analyse et d'exploration des données pour la gestion de l'élevage pastoral. Ces différents avantages techniques offerts par les lacs de données basés sur Hadoop HDFS répondent mieux aux contraintes imposées par les données hétérogènes.

3.6 Le système d'information spatial pour la gestion des données du pastoralisme : P@stor _ All

3.6.1 Modèle conceptuel des données

La figure 3.3 présente le modèle conceptuel de données que nous proposons. Les différents composants du système agro-sylvo-pastoral sont représentés par les différentes classes correspondant aux données (cf. tableau 3.3). Les informations sur les différents facteurs peuvent être obtenues à la suite d'une analyse appliquée aux informations des données en lien dans le SI. Différents niveaux d'analyse peuvent être réalisés afin de produire de nouvelles informations pour la prise de décision. D'autres informations résultant de l'analyse entre les troupeaux et la zone de pâturage peuvent être historicisées pour suivre leur dynamique dans l'espace et le temps. Ces nouvelles informations contribueront ensuite à d'autres types d'analyse impliquant l'interaction entre le comportement du troupeau, les ressources pastorales, l'environnement et les pratiques de gestion.

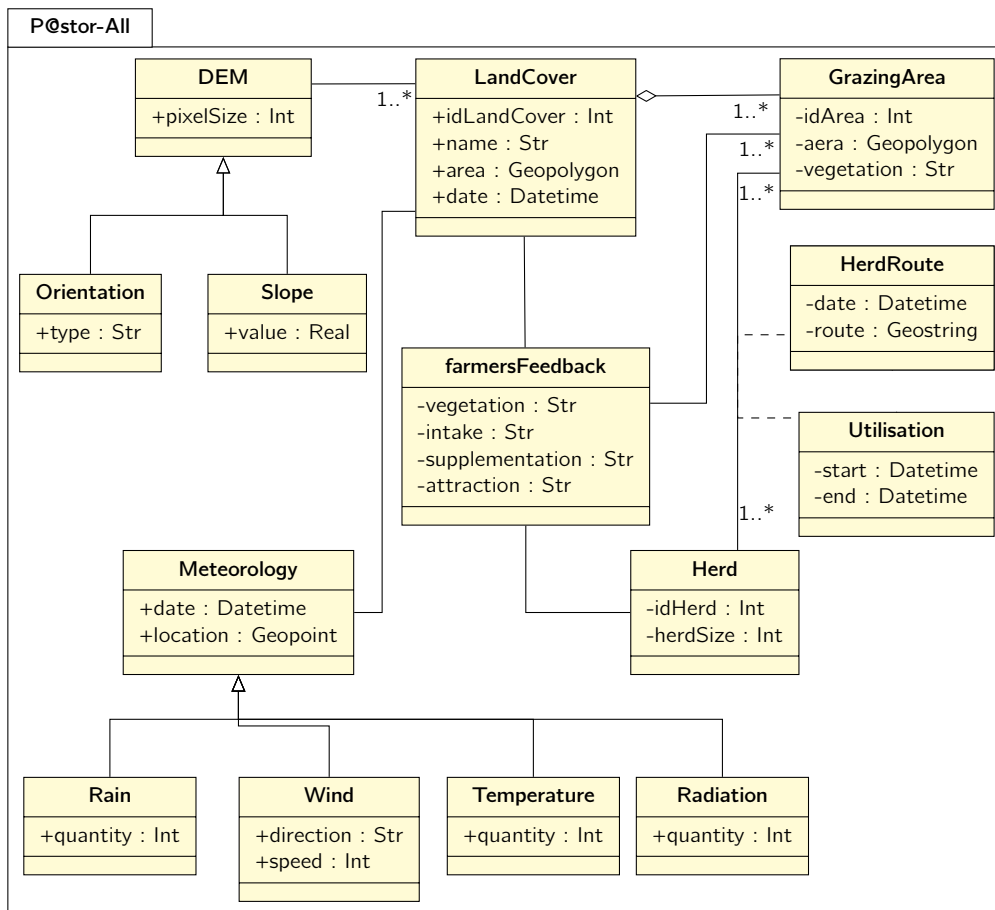


Figure 3.3 – Modèle conceptuel de données du système d'information spatiale pour le système agro-sylvo-pastoral dans les zones méditerranéennes et montagneuses françaises.

Ce modèle décrit les différentes relations entre les données permettant de modéliser l'utilisation des parcours par les troupeaux. Un environnement donné caractérisé par une ou plusieurs occupations de sol (LandCover) contient une ou plusieurs zones de pâturage (GrazingArea), visitées par un ou plusieurs groupes d'animaux (Herd). Les deux premières entités sont caractérisées par l'influence de deux phénomènes spatiaux qui impactent leur utilisation par les troupeaux : le relief (DEM) et la météo (Meteorology). L'utilisation dans le temps et l'espace des zones de pâturage (GrazingArea) par les troupeaux (Herd) permet de produire de nouvelles informations spatio-temporelles : les trajectoires de pâturage (HerdRoute) et l'utilisation des ressources pastorales d'une zone de pâturage (Utilisation), qui peuvent aider à comprendre comment les animaux utilisent les parcours. Enfin, le retour d'éleveurs (FarmersFeedback) portera à la fois sur l'environnement et les zones utilisées, ainsi que sur la gestion des animaux.

3.6.2 Architecture du système d'information

L'architecture du lac de données que nous proposons pour l'intégration et l'analyse des données du pastoralisme comporte 5 principaux composants : collecte des données, prétraitement, ingestion et stockage, traitement et analyse, et visualisation. La figure 3.4 présente l'architecture globale du lac de données spatiales pour l'intégration de données hétérogènes

du pastoralisme et l'analyse croisée des informations afin de produire des indicateurs pour la prise de décision.

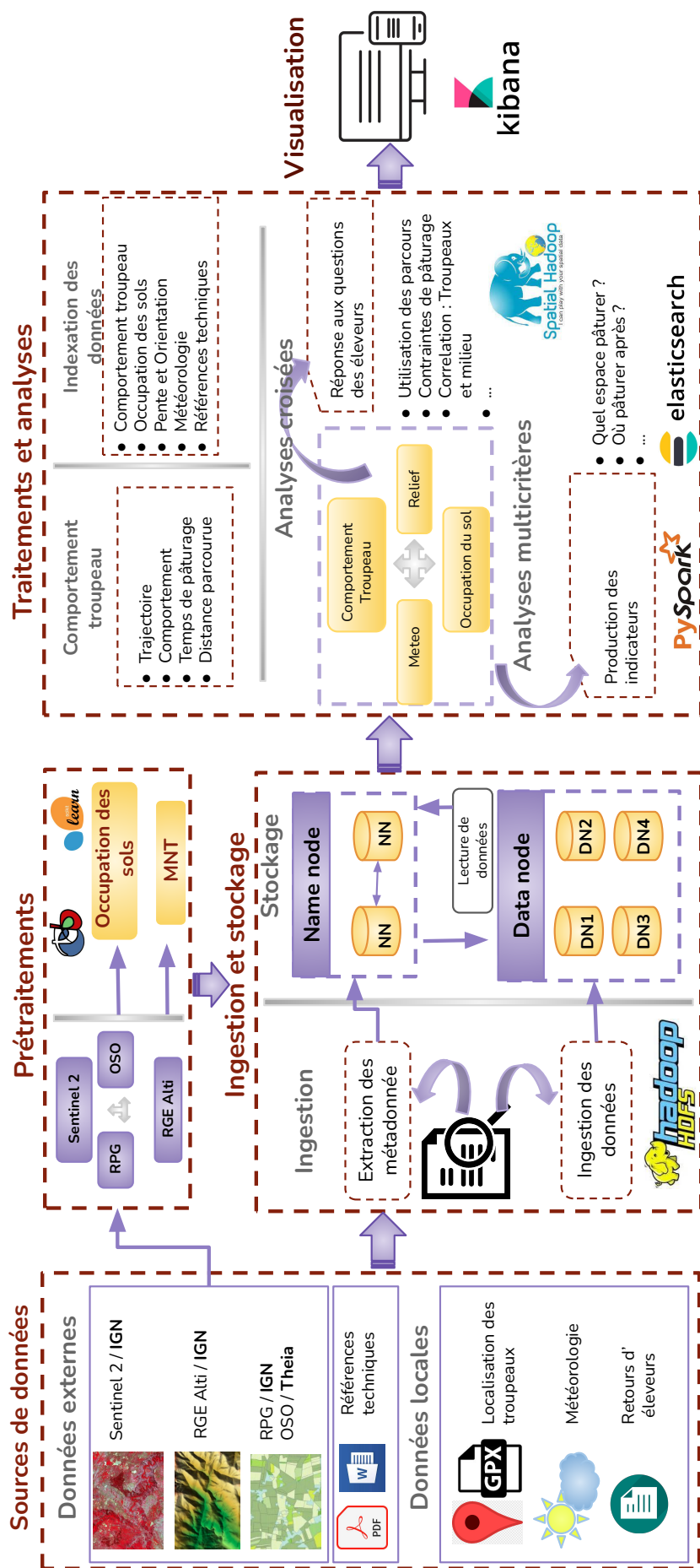


Figure 3.4 – Architecture globale du lac de données spatiales pour la gestion des données du pastoralisme

Prétraitement des données

L'objectif des prétraitements réalisés dans cette composante concerne la production de l'occupation des sols d'intérêt pastoral à partir des données Sentinel 2, du RPG et de l'occupation des sols OSO. Ces étapes de prétraitement ne font pas partie des différentes analyses au sein du lac de données. Les données sont traitées en externe, puis l'occupation des sols produites est injectée dans le lac de données. Ce prétraitement concerne les données identifiées au moment de la réalisation de cette étude. Les différentes étapes de ces traitement sont présentées dans le chapitre 4. A l'avenir, d'autres données et techniques pourront permettre de produire les données d'occupation des sols d'intérêt pastoral avec plus de précision.

Les prétraitements concernent également le MNT en vue de produire différentes variables de relief, à savoir la pente et l'orientation. Il est également possible d'envisager, si nécessaire, un prétraitement pour d'autres types de données afin de faciliter leur intégration dans le lac de données.

Ingestion et stockage des données

La phase d'ingestion et de stockage se réalise grâce aux deux principaux types de nœuds d'HDFS : le NameNode et le DataNode. Le rôle du NameNode étant d'assurer la gestion des données (stockage, création, suppression, duplication, accès) à travers la gestion des métadonnées . Les données sont quant à elles stockées au sein des DataNodes qui composent un cluster d'HDFS. Pendant le processus de stockage, chaque fichier est fragmenté en plusieurs fichiers de petites tailles qui seront par la suite répliqués autant de fois sur les DataNodes. Ce mécanisme permet au lac de données d'assurer la disponibilité des données en cas de défaillance d'un ou de plusieurs nœuds du système. Toutes les opérations appliquées aux données stockées se font de manière parallèle sur les DataNodes et les résultats sont renvoyés après agrégation des informations.

Dans un lac de données, la gestion des données est primordiale. Cela se fait par des requêtes au niveau des métadonnées. Cependant, Apache Hadoop ne dispose pas de toutes les fonctionnalités nécessaires pour capturer les métadonnées de la grande variété de données existantes. Face à la variété des données que nous avons identifiées dans cette étude, il est nécessaire de définir des moyens adaptés pour extraire les métadonnées de manière objective. Pour surmonter cette contrainte, différentes études ont proposé des solutions complémentaires (Hai et al., 2016; Quix and Hai, 2019; Sawadogo et al., 2019). Cependant, ces études n'ont pas présenté une implémentation capable de traiter la variété de données hétérogènes (notamment spatiales) que nous avons. L'implémentation réalisée par Kafando et al. (2020), intègre ces fonctionnalités. Le système de gestion des métadonnées mis en œuvre dans le lac de données est basé sur GeoNetwork¹⁰ qui utilise la norme ISO 19115 pour gérer les métadonnées des données spatiales. Geonetwork permet de stocker les métadonnées, ainsi

10. <https://geonetwork-opensource.org/>

que le chemin HDFS du lac de données, et facilite ainsi la récupération des données.

Pour cette phase, tout d'abord, les données subissent le processus d'ingestion, où leur analyse préliminaire est effectuée pour les différencier (par exemple, par rapport à leur format ou leur nature) et extraire leurs métadonnées grâce à GeoNetwork. Comme données d'entrée dans le lac de données, nous répertorions :

- l'occupation des sols issue du prétraitement,
- les données topographiques,
- les données météorologiques,
- les fiches techniques,
- les données de localisation des troupeaux, et
- les informations d'éleveurs

Une veille est nécessaire pour assurer l'alimentation du lac de données par de nouvelles données en fonction de leurs disponibilités.

Traitement et analyse des données

Ce composant permet au SI d'effectuer des analyses pour produire des indicateurs d'aide à la décision. Une fois les données stockées, la recherche devient plus facile grâce aux liens définis entre les données et les métadonnées via GeoNetwork. Ainsi, les données peuvent être utilisées en fonction des besoins des utilisateurs.

Les applications associées aux besoins des éleveurs pastoraux sont mises en œuvre dans le SI. Nous classons ces applications en 5 groupes (voir Tableau 3.7) : 1) requêtes spatiales, 2) métriques statistiques, 3) analyse spatiale, 4) analyse de corrélation, et 5) fouille de texte.

Les questions concernant la fonction "repères pour l'ajustement du pâturage" sont principalement traitées à l'aide de requêtes spatiales et d'analyses spatiales. D'autres questions de ce groupe nécessitent des métriques statistiques pour produire des informations agrégées. Les requêtes spatiales sont appliquées à des données spatiales (occupation du sol, trajectoire du troupeau, topographie et météo) au niveau de chaque exploitation. Elles permettent aux agriculteurs d'identifier les caractéristiques de l'environnement (ressources, topographie et climat si possible) et d'observer les caractéristiques du mouvement effectué par les troupeaux lors du pâturage. Toutes les questions concernant la fonction "méthodes pour atteindre un certain objectif" font référence à l'exploration de texte sur les références techniques pastorales et les retours d'éleveurs. Enfin, la seule question concernant la fonction "connaissance de l'écologie des écosystèmes agro-sylvo-pastoraux" fait appel à des analyses de corrélation entre les caractéristiques du milieu et les différents aspects de la gestion et de la production agro-sylvo-pastorale. Cette question sera principalement abordée par les scientifiques, en dehors du SI qui produiront des "fiches scientifiques" à injecter dans le SI par la suite comme

ressource documentaire pour les éleveurs. Nous présentons un résumé sur les différents types d'analyses pouvant être prises en charge par le SI dans le tableau 3.7 qui suit.

Au-delà des besoins des éleveurs, les chercheurs pourront creuser la compréhension nécessaire à répondre à des questions pratiques de l'élevage pastoral. Cependant, pour réaliser ces analyses selon les normes scientifiques suivant un protocole bien élaboré, ils devront réaliser une sélection des données suffisantes pour permettre de tester des hypothèses et déterminer leurs domaines de validité. Là où le SI fournit des résultats d'analyse pour la prise de décision des éleveurs, la recherche trouvera des informations pour fonder de nouvelles hypothèses et les vérifier par d'autres types d'analyse. Ces nouvelles connaissances issues de la recherche peuvent ensuite être intégrées dans le SI.

Table 3.7 – Aperçu de certaines analyses au sein du SI qui peuvent aider à répondre aux besoins des éleveurs

Besoins des éleveurs	Exemples d'analyses à intégrer dans le SI
Fonctions 1 : Méthodes pour ajuster le pâturage sur parcours	
Quelle est la surface de pâturage utilisée actuellement (1) ?	Données : - Occupation des sols + localisation GPS
Quel parcours mon troupeau emprunte-t-il chaque jour (2)	Analyses : - Requêtes spatiales permettant d'identifier le type d'occupation des sol utilisé par les animaux (1) - Transformation des données GPS en trajectoires parcourues par les troupeaux (2) - Intersection entre les trajectoires et l'occupation des sols pour caractériser l'utilisation des types de végétation (2)
Combien de temps mon troupeau passe-t-il au pâturage ? (3)	Données : - Localisation GPS des troupeaux + météo + occupation des sols + MNT
Quelle est la superficie occupée par le troupeau ? (4)	Analyses : - analyse de motifs de points - Caractérisation du comportement (repos-pâturage) et estimation du temps de pâturage (3)
Comment mes animaux utilisent-ils l'espace ? (5)	- Calcul de l'étendue représentant le motif de point à un instant T (heure, créneau ou jour) (4) - Analyse de la distribution des points dans le milieu exploré : regroupement ou dispersion des animaux ? (5)
Fonctions 2 : Méthodes pour atteindre un certain objectif	
Comment valoriser les zones embroussaiées ? (6)	Données : - Références techniques, fiches techniques et retours d'éleveurs
Comment valoriser la ressource pastorale pour répondre aux besoin de mon troupeaux ? (7)	Analyses : - Recherche par mots clés (6, 7 et 8)
Comment inciter les animaux à se déplacer vers les zones inexplorées ? (8)	
Fonctions 3 : Mieux comprendre l'écologie de l'écosystème agro-sylvo-pastoral	
Comment le comportement des troupeaux évolue-t-il en fonction de la saison ? (9)	Données : - Localisation GPS + Occupation des sols + MNT + Météo
	Analyses : - Analyse de motifs de point pour produire les indicateurs d'utilisation de l'espace (8) - Corrélation entre les indicateurs et les facteurs du milieu (8) - Analyses spatio-temporelle des indicateurs d'utilisation de l'espace (8)

Visualisation

Les informations produites au sein du lac de données sont accessibles aux utilisateurs via différentes interfaces de visualisation en fonction de différents niveaux d'accès. Trois rôles d'accès peuvent être établis : un dédié aux exploitations agro-sylvo-pastorales, un aux chercheurs, et un à tous les types d'utilisateurs. Le rôle spécifique aux exploitations agro-sylvo-pastorales permettra à chaque éleveur ou site expérimental d'accéder aux données décrivant son exploitation. Toutes les informations relatives au tableau 3.5 du lac de données sont utilisées pour caractériser chaque exploitation individuelle. Aucun éleveur n'a accès aux données spécifiques à une autre exploitation que la sienne. Le rôle pour les chercheurs leur donne accès aux informations et aux outils qui leur permettent de réaliser des analyses à grande échelle. Ils sont capables de combiner les différents types d'informations du SI pour produire de nouvelles connaissances sur le pastoralisme. Enfin, le rôle dédiée à tous les utilisateurs, concernera à la fois les éleveurs, les chercheurs et les autres utilisateurs intéressés par l'application. Elle permettra d'accéder à tous les types d'informations, à l'exception de celles collectées au niveau des exploitations agro-sylvo-pastorales (données GPS ou données sur les troupeaux). Les personnes qui veulent se lancer dans l'élevage pourront trouver les informations anonymisées nécessaires à la conduite à partir de cette interface en utilisant les références techniques. De nouvelles fonctionnalités pour les usages par les techniciens et les enseignants pourront également être implémentées au fur et à mesure.

3.6.3 Prototype opérationnel du lac de données

Un prototype opérationnel du lac de données destiné à fournir les fonctionnalités pour l'élevage pastoral est présenté dans la figure 3.5. Trois principaux clusters sont construits pour assurer : le stockage et la gestion des données, les requête et le catalogage des données, et la recherche et la représentation des informations.

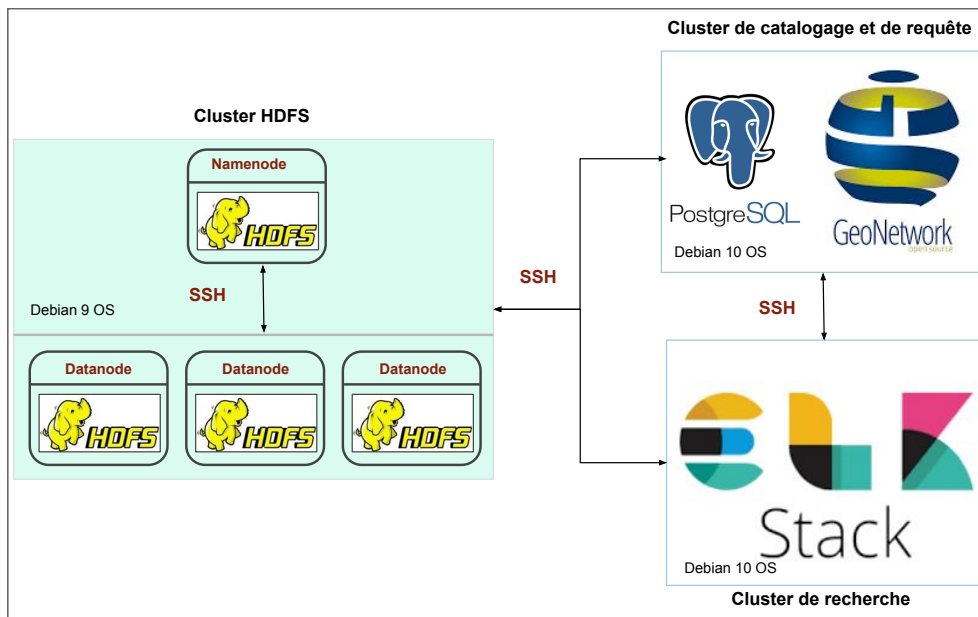


Figure 3.5 – Infrastructure du prototype du lac de données spatiales implémentée pour l’intégration des données du pastoralisme.

Le composant principal du lac de données consiste en un cluster HDFS¹¹ qui assure le stockage et la gestion distribués des données. Le cluster HDFS implémenté fonctionne sur 3 nœuds basés sur le système d’exploitation Linux Debian 9. Sa structure est constituée d’un NameNode assurant la gestion des données. De nombreux types de tâches peuvent y être exécutés (création, suppression, duplication et lecture de fichiers stockés dans les différents nœuds du cluster). Trois nœuds du cluster HDFS sont constitués de DataNodes où les données sont distribuées et stockées. Différents outils de traitement et d’analyse peuvent facilement avoir accès au cluster HDFS. De cette manière, les analyses décrites précédemment seront réalisées et implémentées afin de produire des informations précieuses pour l’élevage pastoral.

Le cluster de catalogage et de requête¹² comprend GeoNetwork et PostgreSQL pour la gestion des données spatiales. Ils fonctionnent sur le système d’exploitation Linux Debian 10. Les différents types de données spatiales et structurées, à savoir l’utilisation et la couverture des sols, la pente et l’orientation, la météorologie et la localisation des troupeaux, sont stockés dans PostgreSQL après le prétraitement des données brutes dans le cluster HDFS. GeoNetwork fournit les métadonnées liées à chaque donnée stockée dans la base de données structurée en utilisant la norme de métadonnées ISO 19115-3.2018. L’objectif principal de ce cluster est de fournir à l’utilisateur, un accès aux différentes données structurées (spatiales et non spatiales) disponibles dans le lac de données, et de proposer des fonctionnalités de requête sur ces données.

Enfin, le cluster de moteur de recherche se compose la suite d’Elasticsearch Logstash et de

11. <https://github.com/aidmoit/ansible-deployment>

12. https://forgemia.inra.fr/beyond/beyond_si

Kibana (ELK), installée sur le système d'exploitation Debian 10 Linux. Elasticsearch fournit le moteur de recherche sur les données stockées dans le lac de données après leur ingestion via Logstash. L'utilisateur peut ensuite réaliser différentes requêtes sur les données indexées via l'interface Kibana pour visualiser, agréger et résumer de nombreux types d'informations sur la gestion du pâturage.

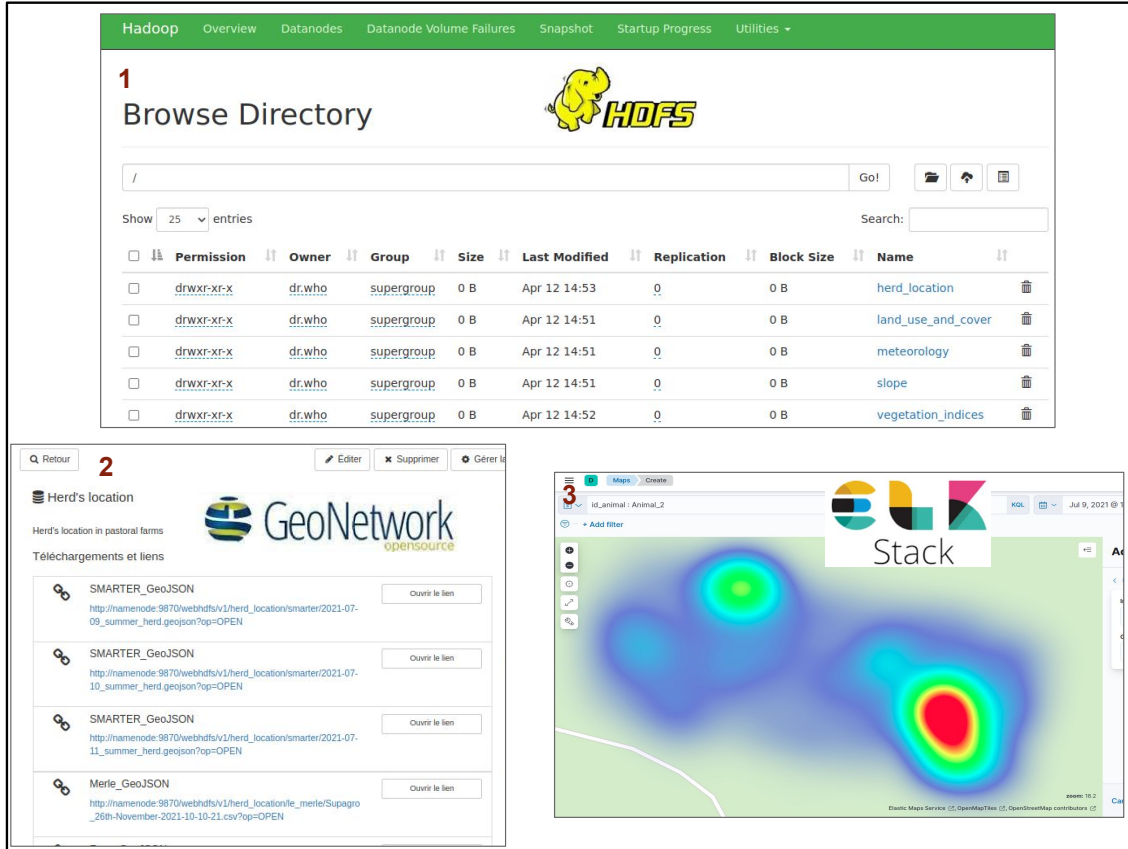


Figure 3.6 – Les interfaces des différents composants du lac de données. 1) Structure de fichiers dans HDFS listant les dossiers des données spatiales des exploitations pastorales. 2) Catalogue de métadonnées dans GeoNetwork donnant accès aux données via un lien web vers HDFS. 3) Visualisation dans Kibana (sous forme de carte de chaleur) des informations de localisation GPS (d'une journée) d'un animal, ingérées dans Elasticsearch.

3.7 Conclusions et perspectives

La diversité et les enjeux des systèmes agro-sylvo-pastoraux des zones méditerranéennes françaises favorisent une diversité d'informations hétérogènes et une complexité dans le processus de prise de décision. Dans ce contexte la mise en place d'un référentiel s'impose pour documenter le pastoralisme et améliorer le processus de prise de décision. Les lacs de données sont aujourd'hui une référence pour la gestion des données hétérogènes pour des applications dans de nombreux domaines. En agriculture, spécialement en élevage pastoral, la littérature scientifique abordée ne fait pas mention de l'usage de tels outils.

Dans ce chapitre, nous avons présenté les processus de co-construction (avec les éleveurs pastoraux, chercheurs, et techniciens de l'institut d'élevage) d'un système d'information spatial basé sur le lac de données en vue d'intégrer les données hétérogènes et de servir de référence pour le pastoralisme en zones méditerranéennes françaises. Ce processus a été réalisé en trois étapes essentielles : l'identification des données permettant de modéliser l'écosystème agro-sylvo-pastoral, la définition des fonctionnalités du système d'information et la conception d'un lac de données spatiales.

La principale contribution du travail consiste en la proposition d'un SI en adéquation avec les enjeux du pastoralisme en zones méditerranéennes françaises : (i) économiques, environnementaux et de biodiversité, et foncier, (ii) données hétérogènes. Tout d'abord, nous proposons un référentiel capable d'intégrer et de centraliser les données hétérogènes en adéquation avec les enjeux du pastoralisme et les besoins ciblés des éleveurs pastoraux en zones méditerranéennes françaises. Puis, nous établissons des approches de combinaison des données hétérogènes en vue de produire des indicateurs utiles pour la prise de décision en élevage pastoral.

Dans le contexte de variabilité et d'hétérogénéité des systèmes agro-sylvo-pastoraux, ce système de référence fournira des informations adaptées à chaque exploitation grâce aux données qu'elles produisent. Les éleveurs pastoraux disposeront ainsi d'un outil efficace pour les aider à prendre des décisions sur les actions à mener. Il constituera également une source d'information pour la recherche afin de réaliser des études à plus grande échelle et de documenter les connaissances sur ces systèmes variables et hétérogènes. De plus, ce SI pourrait déboucher sur un réseau d'échange d'une part entre les éleveurs dans le but d'améliorer les connaissances et les pratiques, et d'autre part entre les éleveurs et les autres usagers impliqués dans le pastoralisme pour faciliter la gestion du pâturage.

Pour contribuer à la gestion de l'élevage pastoral, l'implémentation des fonctionnalités d'analyse est nécessaire en vue de produire des indicateurs d'aide à la décision pour l'élevage pastoral. Ces indicateurs doivent contenir des informations qui font sens pour les éleveurs pastoraux en vue de leur adoption. Ces considérations constituent l'objet des analyses que nous présentons dans le chapitre 5.

Le processus de co-conception ne sera pas seulement limité aux applications définies dans cette étude. Tout au long du projet P@stor_All, des interactions régulières entre les acteurs permettront d'améliorer les applications établies et d'en proposer d'autres à intégrer dans le SI. Ces applications concerneront notamment le croisement et la visualisation des informations. Elles offriront aux utilisateurs des informations pertinentes et adaptées à leurs besoins.

Enfin, la disponibilité des données au niveau européen, et le développement de plus en plus d'outils d'acquisition des données aux niveaux des exploitations pastorales, est une opportunité pour la généralisation d'un tel outil. Ainsi, l'approche de conception présentée dans cette étude pourrait également être appliquée à d'autres régions de montagne et méditerranéennes

d'Europe afin de construire un SI adapté au contexte de chaque région.

Chapitre 4

Préparation et prétraitement des données du pastoralisme

Les données identifiées dans le chapitre précédent sont hétérogènes et possèdent des qualités variables. Dans le présent chapitre 4, nous abordons les aspects liés à la qualité des données spatiales pour modéliser l'écosystème agro-sylvo-pastoral et le comportement spatial des troupeaux. Les travaux de préparation et de prétraitement de certaines données spatiales sont également présentés.

4.1 Introduction

L'utilisation des parcours par les animaux dépend de plusieurs facteurs identifiés au sein des composants de l'écosystème agro-sylvo-pastoral (ressources pastorales, environnement, animaux et mode de gestion) présenté au chapitre 3. Elle est caractérisée à la fois par une dimension spatiale et temporelle sous l'action de ces différents facteurs. La dimension spatiale représente principalement le milieu physique statique, le sol avec sa topographie (Bailey et al., 2015) et les types de végétation, qui procurent les ressources pastorales (Wang et al., 2009b; Chebli et al., 2020) et varient spatialement et temporellement, et les paramètres météorologiques (Manteca i Vilanova and J. Smith, 2014). Pour parvenir à une bonne compréhension du comportement spatial des animaux sur parcours dans le temps et dans l'espace, une caractérisation des facteurs dans ces deux dimensions est utile.

La nature hétérogène des données identifiées suscite la question de leurs capacité à pouvoir documenter les différents facteurs de l'écosystème agro-sylvo-pastoral ainsi que le comportement spatial des animaux. Nous identifions trois enjeux différents liés cette hétérogénéité des données : la précision, l'échelle de représentation et la sémantique, et la granularité temporelle.

La précision concerne l'ensemble des données spatiales identifiées : OSO, RPG, BD-Forêt, localisation des troupeaux (GPS), Sentinel 2 et MNT RGE ALTI. Elle se réfère au niveau d'exactitude des données à représenter les différentes informations spatiales qu'elles portent (Li et al., 2018). Ce niveau d'exactitude spatiale varie pour les différentes données présentées. Le niveau de précision des classes pastorales dans l'occupation des sols OSO est faible. Cette faible précision est principalement liée à la difficulté de les discriminer avec les autres classes (Inglada et al., 2017). La précision des données RPG quant à elle est non évaluée et reste très variable s'il faut prendre en compte les erreurs systématiques liées à la production par un grand nombre d'opérateurs. Quant aux données de localisation des troupeaux, la qualité de la précision est liée aux caractéristiques des GPS et aux conditions variées des milieux pastoraux. La précision des données satellites (Sentinel 2) ou aéroportées (MNT RGE ALTI) est plutôt liée aux aspects propres aux capteurs et aux conditions de prise de vue. Cependant, dans cette étude, nous n'aborderons pas cet aspect de précision mais plutôt celle qui concerne leurs données dérivées, directement utilisables pour caractériser l'écosystème agro-sylvo-pastoral.

L'échelle de représentation détermine également la qualité des données géographique (MacEachren et al., 2005). Elle se réfère au niveau de détail sémantique que les données spatiales peuvent avoir. Dans notre étude, elle concerne principalement les données permettant de représenter la composante "ressources pastorales" végétales. Pour cela, les données OSO ne permettent pas de décrire avec plus de précision les zones explorées par les animaux, et particulièrement les surfaces comportant un embroussaillement ou un boisement, afin de caractériser leur utilisation par les troupeaux. Quant à la sémantique, il s'agit prin-

cipalement de données qui ne permettent pas de représenter l'information spatiale sans être transformées. Leur transformation est nécessaire afin de produire des informations utilisables.

Enfin, la granularité temporelle des données spatiales constitue un aspect important qu'il est utile de considérer pour caractériser les phénomènes dynamiques au sein des écosystèmes agro-sylvo-pastoraux. L'aspect dynamique des phénomènes au sein de l'écosystème agro-sylvo-pastoral est représenté par l'évolution de l'état de la végétation, les variations des paramètres météorologiques et les variations spatio-temporelles de la position des animaux. La fréquence de mesure des différents facteurs détermine les usages et informations qu'on pourra tirer de l'analyse des données. Par exemple des positions du troupeau toutes les 2h ne permettront pas de déterminer les zones explorées par les animaux. Par ailleurs, cette granularité temporelle doit être compatible entre types de données : une moyenne de température mensuelle ne peut servir à expliquer la répartition des activités au cours de la journée. Par contre, au vu de l'évolution lente de l'embroussaillage, on peut se contenter d'une image satellite annuelle pour analyser l'exploration des zones boisées par le troupeau. Les données météorologiques des sources publiques manquent de précisions pour caractériser les exploitations pastorales. Néanmoins, la présence dans certaines exploitations de station météorologique peut être un atout considérable. L'information sur l'activité des animaux est collectée à partir des GPS possédant des résolutions temporelles variées. Il est également utile de définir une fréquence d'acquisition temporelle qui permette de mieux caractériser l'utilisation de l'espace par les troupeaux.

Toutes ces considérations, liées à l'hétérogénéité de la qualité des données spatiales, impactent la caractérisation de l'écosystème agro-sylvo-pastoral. En effet, les imprécisions observées dans les données affectent directement la qualité des solutions qui en résultent (Redman, 1998; Saha and Srivastava, 2014). Il convient alors de trouver des compromis en termes de granularité spatiale et temporelle des différents types de données disponibles, pouvant permettre de réaliser des analyses (croisées ou pas) pour répondre aux besoins des éleveurs.

Dans ce chapitre, nous nous intéressons à l'approche globale de la qualité des données spatiales ainsi qu'à l'exploration des méthodes permettant d'améliorer, de préparer et de produire les données de qualité intéressante pour contribuer à la compréhension du fonctionnement d'un écosystème agro-sylvo-pastoral en zones méditerranéennes françaises. Dans un premier temps, nous abordons, dans la section 4.2, la question de la qualité des données spatiales du pastoralisme. Dans la section 4.3, nous présentons les approches méthodologiques d'amélioration de l'occupation des sols d'intérêt pastoral. Nous abordons par la suite la démarche suivie pour préparer les données de localisation des troupeaux dans la section 4.4. Puis dans la section 4.5, nous présentons les prétraitements des données rasters pour la production des données topographiques et de densité de végétation. Enfin, la section 4.6 conclut ce chapitre.

4.2 Qualité des données spatiales du pastoralisme

Pour définir la qualité des données spatiales, nous nous basons essentiellement sur leur capacité à décrire les différents facteurs de l'écosystème agro-sylvo-pastoral dans le but de caractériser l'utilisation spatiale des parcours par les troupeaux. La qualité des données spatiales s'évalue en se référant aux trois critères suivants : la précision, l'échelle de représentation et la sémantique, et la granularité temporelle. Les données se référant aux différents facteurs associés à chaque composant du comportement spatial des troupeaux sont concernées au moins par l'un des trois critères pour mieux les caractériser. Le composant "troupeau" n'a pas été présenté étant donné que les facteurs qui le décrivent (physiologie, morphologie, mémoire spatiale et relations sociales) sont intrinsèques aux animaux et ne peuvent pas être décrits à partir des données spatiales à disposition.

4.2.1 Modélisation des ressources pastorales

Deux types d'échelle de représentation peuvent être utilisés pour décrire les types de végétations pastorales. Premièrement une généralisation des classes de végétation à partir des données d'occupation des sols permet d'identifier les types de végétation utilisés par les éleveurs pastoraux. Ces classes d'occupation des sols sont utilisées par les animaux pour les ressources pastorales qu'elles contiennent. Dans le cadre du pastoralisme en zones méditerranéennes françaises, il s'agit des landes, des pelouses, des forêts ainsi que des espaces de cultures de vignes et de vergers. A cette échelle de représentation, on est en mesure de décrire plus l'aspect statique des espaces utilisés par les animaux. Deuxièmement, une échelle de représentation plus fine est essentielle pour caractériser les ressources fourragères consommées par les animaux. C'est en fonction de l'évolution spatio-temporelle de ces ressources que les éleveurs pastoraux et les animaux varient leur utilisation de l'espace.

Pour représenter les types de végétations pastorales, la donnée OSO constitue la principale référence. Elle contient 23 classes produites à partir des séries temporelles des données Sentinel 2 (Inglada et al., 2017), et permet de représenter l'information que nous considérons comme statique pour décrire les milieux utilisés par les animaux. Pour cette donnée nous nous intéresseront principalement à la précision des classes pastorales. En effet, les principales classes de parcours, les landes et les pelouses, n'y sont pas bien discriminées à cause de leur forte hétérogénéité spatiale avec des F-scores respectifs de 0,59 et 0,66 (cf. figure 4.1). La confusion entre ces classes de parcours et les autres classes avec lesquelles elles interagissent le plus est présentée dans le Tableau 4.1 qui montre un extrait de la matrice de confusion d'OSO 2018. Les landes ligneuses et les pelouses sont en confusion avec les vergers et les vignes et encore plus avec les forêts. La grande partie de pixels qui devrait appartenir à ces deux classes est attribuée aux autres classes (faux négatifs) et elles incluent également une part importante d'objets appartenant aux autres classes (faux positifs). Par ailleurs, même si les vergers possèdent un faible taux d'objets appartenant à d'autres classes

(faux positifs), une part importante de leurs objets se trouvent repartis dans toutes les autres classes, principalement les forêts de feuillus. Cette faiblesse de précision des classes pastorales peut impacter la gestion du pâturage. Un rehaussement de la précision des classes pastorales contribuera à améliorer la gestion pastorale. Pour y parvenir, il est utile de considérer l'ensemble des classes avec lesquelles elles sont le plus en interaction, que sont les vignes et les deux types de forêt.

Table 4.1 – Extrait de la matrice de confusion d'OSO 2018 sur les classes à faibles F-scores et celles qui influencent le plus leurs imprécisions. Ver. = Verger, Vig. = Vigne, F. F. = Forêt de feuillus, F.C. = Forêt de conifères, Pel. = Pelouse, L. L. = Landes ligneuses, Pré. = Précision, Rap. = Rappel, et F.-S. = F-score.

Classe	Ver.	Vig.	F. F	F. C.	Pel.	L. L	Rap.	F-S.
Ver.	2,53E+05	4,23E+04	7,11E+04	1,12E+04	3,57E+04	1,89E+04	0,38	0,53
Vig.	3,18E+03	2,73E+06	6,00E+03	3,88E+04	4,22E+04	2,55E+04	0,87	0,87
F. F	-	-	3,42E+07	1,14E+06	4,23E+05	3,60E+05	0,94	0,91
F. C.	-	-	1,41E+06	3,44E+07	6,17E+05	3,80E+05	0,92	0,91
Pel.	-	6,07E+04	1,04E+06	1,01E+06	1,24E+07	2,25E+06	0,64	0,65
L. L	-	4,39E+04	1,08E+06	8,63E+05	3,09E+06	7,36E+06	0,52	0,57
Pré.	0,84	0,88	0,89	0,9	0,66	0,64		

La description de l'état des ressources pastorales consommées par les animaux nécessite les données temporelles de meilleure précision spatiale. L'analyse de la littérature scientifique à ce sujet montre une utilisation des données de densité de végétation à partir d'indices spectraux, notamment le NDVI pour décrire l'état de la végétation au fil du temps et de l'utilisation par les animaux (Hill et al., 2004; Júnior et al., 2019; Serrano et al., 2021). Cheng et al. (2021) ont utilisé le NDVI pour caractériser l'intensité de pâturage de chèvres à partir des données mensuelles Landsat 8 OLI (30 m de résolution spatiale). Serrano et al. (2021) évaluent quant à eux les possibilité de suivre l'évolution des pâturages en utilisant le NDVI issus des données Sentinel 2 (10 m de résolution spatiale). Cependant, le NDVI peut subir l'influence des sols environnant qui auront tendance à saturer l'information des données satellites. Cela peut être le cas dans les parcours méditerranéens qui sont très hétérogènes. Pour pallier à ce type de problème, Fern et al. (2018) ont utilisé l'OSAVI qui prend en compte l'influence du sol environnant. Ils ont trouvé que l'OSAVI a été plus performant que le NDVI. Ces deux indices peuvent être conjointement utilisés pour caractériser l'état et l'évolution des ressources pastorales afin de suivre le pâturage en zones méditerranéennes françaises.

La disponibilité des données Sentinel 2 (10 m de résolution spatiale et 5 jours de passage effectués par les deux satellites) offre une meilleure résolution spatiale pour notre étude. Cependant, l'obtention des données tous les 5 jours peut être problématique du fait de la présence de nuages même sur les données corrigées (Sentinel 2 L2A). D'autres types de données gratuites à grande fréquence d'acquisition peuvent compenser ce manque. C'est le cas des données Planet qui proposent des données quotidiennes avec une résolution spatiale de 3 m. Le choix de la résolution temporelle est donc déterminé par les données disponibles. Nous estimons que les données de meilleure précision capturées à une fréquence entre 5 et 10 jours,

peuvent convenir pour le suivi de l'évolution des ressources pastorales au fil de leur utilisation par les animaux. Des données encore plus fréquentes comme Planet, pourront également être utiles surtout pour l'analyse des interactions multifacteurs avec le comportement spatial des troupeaux.

4.2.2 Modélisation de l'environnement de pâturage

Les facteurs qui caractérisent l'environnement sont la topographie (Bailey et al., 2015) et les paramètres météorologiques (Manteca i Vilanova and J. Smith, 2014). Ces deux types de facteurs peuvent être respectivement représentés à partir du MNT RGE ALTI et des données météorologiques enregistrées par des stations. La qualité des données MNT RGE ALTI¹ concerne principalement sa résolution spatiale et le type d'information qu'elle représente. Elle possède une taille de maille de 5 m et permet de représenter avec une très bonne précision spatiale la topographie des milieux pastoraux. Cependant, l'utilisation de cette donnée à l'état brute n'est pas de grand intérêt pour la conduite du pâturage étant donné qu'elle ne représente que les altitudes dans une zone donnée et pas la forme du terrain. Une transformation en donnée spatiale de pente et d'orientation est nécessaire.

Quant aux données météorologiques, nous distinguons les prévisions et les archives. La première catégorie de données concerne l'information de prévision météorologique fournie par Météo France à travers des applications utilisateurs. Elle peut être utile pour renseigner les conditions météorologiques journalières (pluie, température, humidité de l'air) en vue d'organiser le pâturage. Cependant, ce type de données présente un scénario météorologique probabiliste. Pour suivre l'influence de la météo sur le comportement spatial des animaux, nous avons besoin d'informations réelles ou des archives. Les archives fournies par Météo France se réfèrent aux stations implantées dans les villes et ne reflètent pas le microclimat des parcours et montagnes méditerranéens explorés par les animaux.

Les seules données faisant sens pour suivre le comportement des animaux sont celles qui peuvent être collectées par des stations au sein ou à proximité des zones d'exploitation pastorales. Cela n'est pas évident pour les élevages privés. L'utilisation des stations météorologiques est plus probable dans les fermes expérimentales car elles possèdent en général leur propre station météo. Nous pouvons nous référer aux données de la plateforme INRAE CLIMATIK² du site expérimental de la Fage gérée par le laboratoire AgroClim d'Avignon en France. Les enregistrements contiennent des informations horaires et journalières sur la température, la vitesse du vent, l'humidité de l'air et le rayonnement global. Cette fréquence d'acquisition est adaptée pour suivre les variations journalières et saisonnières dans le comportement des animaux.

1. <https://geoservices.ign.fr/rgealti>

2. <https://agroclim.inrae.fr/climatik/>

4.2.3 Modélisation du comportement des animaux

Le comportement des animaux est représenté à partir des données collectées par des GPS embarqués sur au moins un animal représentant chaque troupeau des élevages. Deux critères permettent d'évaluer les données GPS pour des applications pastorales : le niveau d'exactitude et la fréquence d'acquisition pouvant permettre de caractériser l'utilisation des parcours par les troupeaux.

Le niveau d'exactitude des GPS dépend principalement des facteurs liés au capteur et aux conditions du milieu. Lorsque les animaux fréquentent les milieux moins ouverts, l'exactitude des GPS peut être affectée. [Buerkert and Schlecht \(2009\)](#) ont analysé les données de trois différentes marques de GPS configurés pour enregistrer les données à chaque 15 secondes. Ils ont observé une différence significative des coordonnées en milieu obstrué alors qu'en milieu ouvert aucune différence n'a été remarquée. N'ayant pas le contrôle sur cet aspect de précision, la meilleure pratique serait d'orienter le choix plutôt vers la marque de GPS ayant montré la robustesse de sa précision face aux différents milieux et présentant la meilleure autonomie en vue de limiter les interventions régulières sur les animaux.

En dehors de l'exactitude, vu le nombre limité de GPS surtout dans les exploitations privées qui possèdent rarement plus d'un GPS, la question de représentativité de l'information pour le troupeau n'est pas à négliger. L'équipement en GPS de chaque animal du troupeau étant impossible pour des raisons tant pratiques, économiques qu'environnementales, on aura tendance à attribuer la localisation du groupe à travers celle de l'animal représentatif. La détermination de ce porteur de GPS reste une tâche difficile.

Les travaux réalisés dans le cadre du projet CLOChèTE ([Grisot et al., 2022](#)) ont permis de réaliser des expérimentations visant à définir les considérations techniques et stratégiques pour l'intégration des capteurs numériques dans la conduite de l'élevage. Les auteurs proposent d'identifier les animaux leaders et explorateurs qui sont les plus représentatifs des déplacements du troupeau. Ils initient les déplacements du troupeau, en recrutant un nombre d'individus significatif. Les éleveurs pastoraux peuvent se baser sur ces critères afin d'assurer la représentativité de l'information sur le déplacement des animaux.

En ce qui concerne la fréquence d'acquisition des informations GPS, les travaux consultés présentent une diversité de choix en fonction des objectifs des études et des contraintes d'autonomie des capteurs GPS. Selon l'ordre croissant du temps d'enregistrement d'informations GPS, [Evanglou et al. \(2014\)](#) utilisent une fréquence de 15 secondes pour déterminer 5 types de comportement d'ovins et de caprins (pâturage, déplacement, debout, rumination, assis). [Feldt and Schlecht \(2016\)](#) utilisent une fréquence de 30 secondes afin de caractériser les différences dans le mode d'usage des pâturages par les troupeaux de bovins et d'ovins. Pour déterminer les milieux les plus utilisés par les bovins et les ovins, [Putfarken et al. \(2008\)](#) utilisent une fréquence de 5 minutes. Et, [Akasbi et al. \(2012\)](#), expérimentent différentes intervalles (30, 45, 60, 75, 90, 105 et 120 minutes) pour évaluer la variation de la distance

parcourue par les caprins. Il ressort de ce bref aperçu sur les différents usages des données GPS que le choix de la fréquence d'acquisition peut être déterminé par les objectifs poursuivis dans chaque étude. En plus, le choix optimal de la fréquence d'enregistrement d'informations GPS dépend à la fois du compromis entre le temps suffisant permettant d'observer le comportement des animaux et l'autonomie d'énergie des capteurs. Dans le cadre de l'élevage en zones méditerranéennes françaises, le projet CLOCHÈTE a favorisé des travaux ayant permis de définir ce compromis à 1 point toutes les 5 minutes. Cette fréquence permet d'obtenir une vision exhaustive des territoires "utilisés" par les animaux (pour la pâture, la chôme, le repos), quand bien même elle induit une perte d'information sur les territoires simplement "explorés" (Grisot et al., 2022).

Table 4.2 – Résumé des critères de qualité des données spatiales permettant de caractériser le fonctionnement des exploitations agro-sylvo-pastorales.

Indicateur de qualité	Type de donnée	Observations
Résolution spatiale	- Images satellites : Sentinel 2 et Planet - MNT RGE ALTI	Imposée par les données disponibles. La résolution spatiale de 10 m de Sentinel 2 pourrait être la référence pour les données rasters.
Précision	- Occupation des sols OSO	Besoin d'appliquer des techniques d'amélioration des précisions des classes utilisées par les animaux.
Fréquence d'acquisition	- Images satellites : Sentinel 2 et Planet - Localisation GPS des animaux	Images satellites : 1 à 5 jours, caractérisant respectivement les données Planet et Sentinel 2. Cet intervalle de fréquence permet de mieux suivre les dynamiques au sein des exploitations. - GPS : 5 minutes d'après les résultats du projet CLOCHÈTE afin de permettre la caractérisation du comportement des animaux.
Echelle de représentation	- Occupation des sols	- Besoin d'apporter une profondeur d'information concernant l'état de la végétation à partir des données de densité de végétation.

4.3 Cartographie des surfaces pastorales à l'aide des données Sentinel 2 et des données publiques

La cartographie de l'occupation des sols d'intérêt pastoral consiste en une amélioration de la donnée OSO en utilisant les données Sentinel 2 L3A, le RPG et la BD-Forêt, prises respectivement comme variables prédictives et données de référence.

Dans la section 4.2.1 nous avons montré que les précisions des classes des milieux utilisées pour le pâturage des animaux sont faibles dans OSO 2019 et nécessitent d'être améliorées. Pour ce faire, nous avons distingué 2 groupes d'objets : (i) ceux aux faibles F-scores (notamment les classes de parcours) à améliorer et (ii) ceux aux F-scores élevés à maintenir.

Le regroupement des classes en ³ : urbain (bâti dense, bâti diffus, zones industrielles et commerciales et surfaces routières), cultures annuelles (colza, céréales, protéagineux, soja, tournesol, maïs, riz et tubercules), cultures pérennes (vergers et vignes), forêts (de feuillus et de conifères), et formations naturelles basses (pelouses et landes ligneuses) (cf. Figure 4.1) ; pour l'occupation des sols de 2017 a permis de réduire la confusion de la classification. Cela améliore les F-scores des zones artificialisées (urbain, 0,97), des cultures annuelles (0,98) et des forêts (0,97). Les groupes "cultures pérennes" et "formations naturelles basses" contenant les landes et les pelouses possèdent des F-scores qui nécessitent d'être améliorés (respectivement 0,82 et 0,79). Les classes "Surfaces minérales", "Glaciers ou neige" et "Eau" possèdent de grandes valeurs de F-score individuellement (respectivement 0,84 ; 0,89 et 0,99). La classe "Plages et Dunes" possède quant à elle un faible F-score (0,62). Ce regroupement de classes permet d'identifier les groupes d'occupation des sols de F-scores élevés ne nécessitant pas d'amélioration.

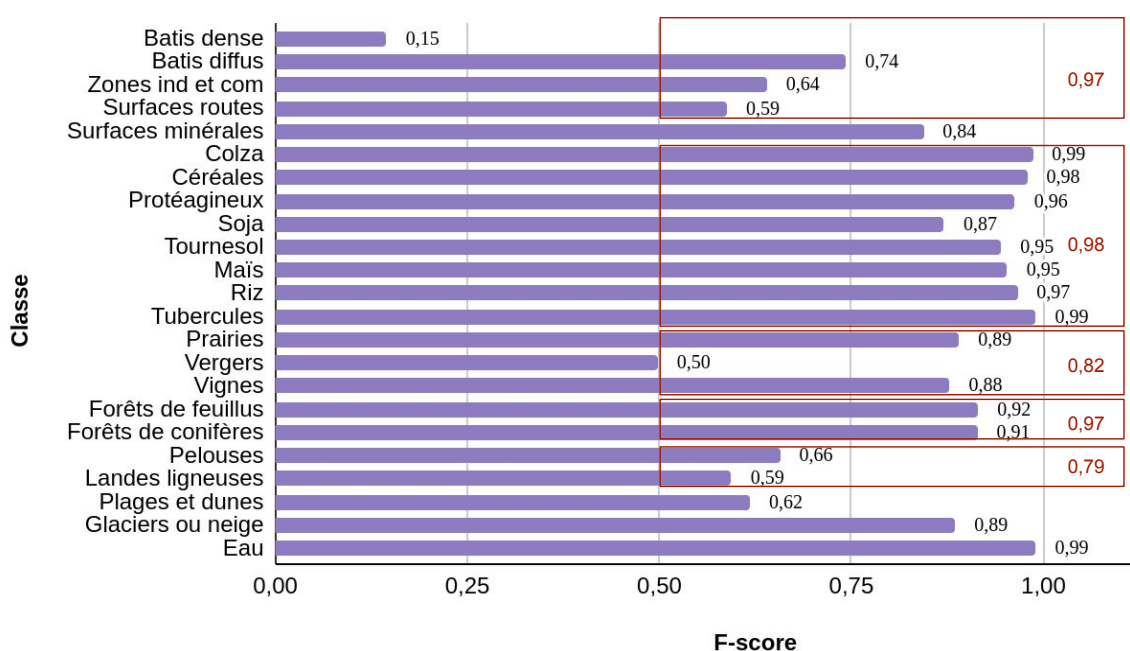


Figure 4.1 – F-score de 23 classes de l'occupation des sols OSO de 2019 et des classes regroupées (en rouge) de l'occupation des sols OSO de 2017. Les F-scores des classes regroupées d'OSO 2017 ont été utilisés à défaut de la disponibilité de l'information pour la donnée de 2019. Source : <http://osr-cesbio.ups-tlse.fr/~oso/>.

4.3.1 Données

Occupation des sols OSO 2019

L'occupation des sols OSO de 2019 a été obtenue en mode vectoriel avec une unité minimale de cartographie de 0,1 ha à partir du dépôt du CNES.

3. <http://osr-cesbio.ups-tlse.fr/oso/>

Les objets vectoriels d'OSO 2019 ont été scindés en deux (cf. Figure 4.5.C) : (i) les classes ou groupes de classes dont les F-scores sont supérieurs ou égal à 0,95 et n'ayant pas trop d'influence sur les classes moins bien discriminées, à maintenir, et (ii) les groupes de classes de faibles F-scores destinés à être améliorés par des traitements spécifiques en même temps que les autres classes bien discriminées avec lesquelles ils sont le plus confondus. Les classes que nous retenons pour l'amélioration sont celles des pelouses (F-score = 0,66), des landes ligneuses, des vergers, des vignes, des forêts de feuillus et de conifères. La classe "plages et dunes" n'a pas été retenue pour être classifiée du fait du manque de données de référence. Toutefois, le fait de ne pas intégrer cette classe dans le processus de classification n'affecte pas de manière significative les précisions des classes de végétation retenues car étant spectralement séparées.

Pour cette étude, nous avons tenté de distinguer les surfaces de parcours en fonction de la densité des ligneux. Ainsi, nous avons défini deux classes de landes ligneuses : les landes ouvertes et les landes fermées. Cette catégorisation des landes a été motivée par l'existence dans les données de référence du RGP 2019 des objets caractérisant les landes en fonction de leur densité de ligneux (Tableau 4.3).

Au total, 1 888 034 objets à classifier ont été extraits de l'occupation des sols OSO 2019 pour les régions Occitanie et PACA.

Sentinel 2 L3A

La période idéale pour effectuer le choix des données Sentinel 2 s'étend de mi-juin à mi-juillet. Ce choix tient à la fois à la faible proportion de couverture nuageuse mensuelle et à la possibilité de discriminer les classes pastorales, surtout les parcours de montagne utilisés pour le pâturage d'estive. Le choix de cette période d'acquisition des images a été également réalisé par Castro et al. (2020).

La couverture de l'ensemble de la zone d'étude a requis 21 tuiles de données Sentinel 2. Pour la période choisie, il n'a pas été possible d'obtenir toutes les tuiles d'images de niveau L2A dépourvues de couverture nuageuse. Nous avons donc opté pour les données de niveau L3A, sans couverture nuageuse, produites par la chaîne de traitement WASP⁴. Ce sont des données de synthèse mensuelle produites en choisissant parmi plusieurs images de niveau L2A sur une période de 45 jours, le meilleur pixel dépourvu de nuage. 21 tuiles ont ainsi été collectées.

Les données Sentinel L3A ne représentent pas fidèlement l'information temporelle car il est difficile de savoir la date exacte pour chaque pixel. Elles sont principalement utilisées pour la visualisation. Dans l'état actuel, la plupart des travaux se réfèrent plutôt à l'utilisation des données Sentinel 2 L2A. Cependant, dans cette étude, nous estimons que sur la période de 45 jours, la variabilité des propriétés spectrales des classes n'impacte pas significativement

4. <https://labo.obs-mip.fr/multitemp/theias-sentinel-2-l3a-monthly-cloud-free-syntheses/>

leur discrimination.

Les prétraitements réalisés sur les données Sentinel 2 L3A (cf. Figure 4.5.B) ont concerné le rééchantillonnage des bandes spectrales de 20 m de résolution spatiale à 10 m de résolution spatiale et l'empilement des 10 bandes spectrales pour toutes les 21 tuiles. Ensuite, un mosaïquage a été réalisé pour obtenir une couverture totale de la zone d'étude (cf. Figure 4.2).

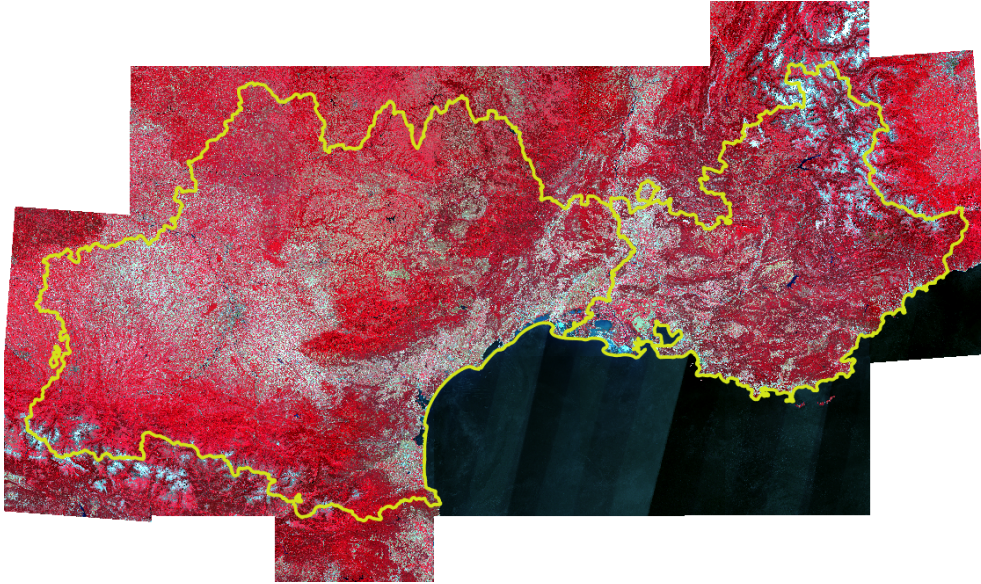


Figure 4.2 – Composition colorée infrarouge fausses couleurs de la mosaïque de 21 tuiles de Sentinel 2 L3A couvrant les régions Occitanie et PACA. Source de données : www.theia-land.fr.

Registre Parcellaire Graphique de 2019 et BD-Forêt version 2.0

Le RPG de 2019 a été utilisé comme source de référence pour les classes pastorales. Plusieurs classes de parcours (pelouses, landes ouvertes, landes fermées et forêts pâturées) y sont représentées, associées aux surfaces déclarées par les éleveurs dans le cadre de la Politique Agricole Commune (PAC). Ces objets représentent les classes OSO 2019 à cartographier. Le Tableau 4.3 montre la correspondance entre le champ "Code Culture" du RPG 2019 et les classes de végétation de parcours correspondantes. La classe "Bois pâturé" n'a pas été utilisée dans le processus de classification car elle se réfère aux classes forestières avec lesquelles il ne sera pas possible de la distinguer spectralement. Ces types de forêts sont utilisés par les troupeaux pour pâturer le sous-bois et s'abriter. Cependant, les objets de cette classe pourront être intégrés à la fin du processus de traitement pour représenter les espaces forestiers utilisés par les troupeaux, vu leur importance pour le pastoralisme.

Tous les objets du RPG 2019 correspondant aux classes de faible F-score dans OSO

2019 ont été extraits pour constituer la base de données de référence et d'évaluation des précisions. 4 classes ont été définies : "Pelouse et Prairie permanente", "Lande ouverte", "Lande fermée" et "Culture pérenne". La classe "Pelouse et prairie permanente" a été appelée "Pelouse" pour simplifier la nomenclature. Au total, 890 826 objets de référence ont été extraits du RPG 2019 pour la zone couverte par les régions Occitanie et PACA. Ce nombre d'objets a ensuite été réduit à 887 854 après élimination des géométries invalides.

Table 4.3 – Correspondance entre les classes du RPG 2019 et les classes pastorales (hormis les bois pâturés) maintenues pour l'apprentissage des modèle de classification.

Code culture	Libellé Culture	Classe correspondante
BOP	Bois pâturé	Bois pâturé
PPH	Prairie permanente - herbe prédominante (ressources fourragères ligneuses absentes ou peu présentes)	Pelouse et prairie
PRL	Prairie en rotation longue (6 ans ou plus)	Pelouse et prairie
SPH	Surface pastorale - herbe prédominante et ressources fourragères ligneuses présentes	Lande ouverte
SPL	Surface pastorale - ressources fourragères ligneuses prédominante	Lande fermée

BD-Forêt

La BD-Forêt de 2019 a été acquise à partir du dépôt de l'IGN. Elle possède une unité cartographique minimale de 5 000 m² (soit 0,5 hectare). Elle a été utilisée comme donnée de référence pour les forêts. Elle contient 5 types d'espaces naturels et semi-naturels : les forêts fermées, les forêts ouvertes, les peupleraies, les landes (maquis, garrigues), et les formations herbacées (pelouses alpines, pelouses pastorales). Une extraction des classes forestières (forêts fermées et forêts ouvertes) a permis d'obtenir 413 816 objets de référence pour la classe forestière. Ce nombre d'objets est passé à 412 705 après élimination des géométries invalides.

4.3.2 Méthodologie

Définition des variables prédictives

Trois groupes de variables ont été produits : les reflectances, les indices spectraux (cf. Tableau 4.4) et les indices de texture. Les statistiques de géométrie ne permettent pas de caractériser les classes à cause de la grande variabilité intra-spécifique de la forme de leurs objets.

Pour les objets OSO et ceux de référence (RPG et BD-Forêt combinés), les moyennes et les écart-types des 10 bandes spectrales de Sentinel 2 L3A (20 variables), ceux de 4 indices de végétation et 3 indices de sol (14 variables) calculés, ainsi que ceux des 8 indices de texture d'Haralick (Haralick et al., 1973) (moyenne, variance, homogénéité, contraste, dissimilarité, entropie, second moment, et corrélation) calculés pour les 4 bandes de 10

m de pixel (64 variables), ont été respectivement extraits. L'utilisation des reflectances et des indices spectraux a été motivé par le travail de [Castro et al. \(2020\)](#), qui ont pu mieux caractériser les classes de parcours à partir de ces variables. Ce processus a permis d'obtenir 98 variables prédictives pour l'entraînement et la validation des modèles, et pour la classification des objets OSO 2019. Tous les calculs d'indices spectraux ont été réalisés sous python 3.5 associé à la boîte à outils Orfeo Toolbox 7.2 (OTB 7.2)⁵. Les calculs de texture ont été réalisés avec logiciel ENVI 5.6.1⁶.

Table 4.4 – Bandes spectrales et indices de végétation et de sol dérivés de Sentinel 2 L3A, pris comme variables pour l'entraînement du modèle.

Bandes et indices spectraux	Définition	Formule
B ₂	Bleue (490 nm)	
B ₃	Verte (560 nm)	
B ₄	Rouge (665 nm)	
B ₅	Red edge (705 nm)	
B ₆	Red edge (749 nm)	
B ₇	Red edge (783 nm)	
B ₈	Proche infrarouge (842 nm)	
B _{8a}	Proche infrarouge (865 nm)	
B ₁₁	Moyen infrarouge (1610 nm)	
B ₁₂	Moyen infrarouge (2190 nm)	
NDVI	Indice différentiel de végétation normalisé (Rouse et al., 1974)	$\frac{B_8 - B_4}{B_8 + B_4}$
NDI ₄₅	Indice différentiel normalisé (Delegido et al., 2013)	$\frac{B_5 - B_4}{B_5 + B_4}$
IRECI	Indice de chlorophylle inversé sur le Red-edge (Delegido et al., 2013)	$\frac{B_7 - B_4}{B_5 / B_6}$
TNDVI	NDVI transformé (Senseman et al., 1996)	$(\frac{B_8 - B_4}{B_8 + B_4} + 0,5)^{0,5}$
CI	Indice de coloration (Pouget et al., 1990)	$\frac{B_4 - B_3}{B_4 + B_3}$
BI	Indice de brillance (Escadafal, 1989)	$(\frac{B_4^2 + B_3^2}{4})^{0,5}$
RI	Indice de rougeur (Pouget et al., 1990)	$\frac{B_4^2}{B_3^3}$

Définition des classes

La définition de classes a été réalisée à partir d'une l'observation de la séparabilité basée sur l'analyse discriminante linéaire (ADL). L'ADL permet d'effectuer une réduction supervisée de la dimensionnalité, en projetant les données d'entrée dans un sous-espace linéaire constitué des directions qui maximisent la séparation entre les classes. L'ADL a été appliqué sur les données de référence de 3 départements (Alpes-de-Haute-Provence (04), Aude (11) et Var (83)), en vue d'analyser les possibilités de maintenir ou de regrouper certaines classes ainsi que les variables sélectionnées.

Pour des raisons d'espace, nous présentons dans cette section uniquement les résultats

5. <https://www.orfeo-toolbox.org/>

6. <https://www.l3harrisgeospatial.com/Software-Technology/ENVI>

du département 04 (Figure 4.3). Les résultats des départements 11 (cf. Figure A.1) et 83 (Figure A.2) sont repris en annexe. On observe à travers la distribution des valeurs des premiers composants de l'ADL que les classes retenues sont difficilement séparables quel que soit le groupe de variables. Les cultures pérennes sont plus confondues avec les pelouses, et les landes fermées sont en confusion entre les forêts et les landes ouvertes. Dans les Figures A.1 et A.2 en annexe, les cultures pérennes sont bien séparées des pelouses et des autres classes. Cela peut être justifié par la variabilité spatiales des classes. Enfin, les boîtes à moustaches (cf. Figures 4.3, A.1 et 83 (Figure A.2) présentent les tendances similaires pour les variables sélectionnées, montrant leurs corrélations à ce point de vue. Cependant, l'utilisation d'un algorithme performant pourrait permettre de tirer profit de ces informations pour maximiser les possibilités de discrimination.

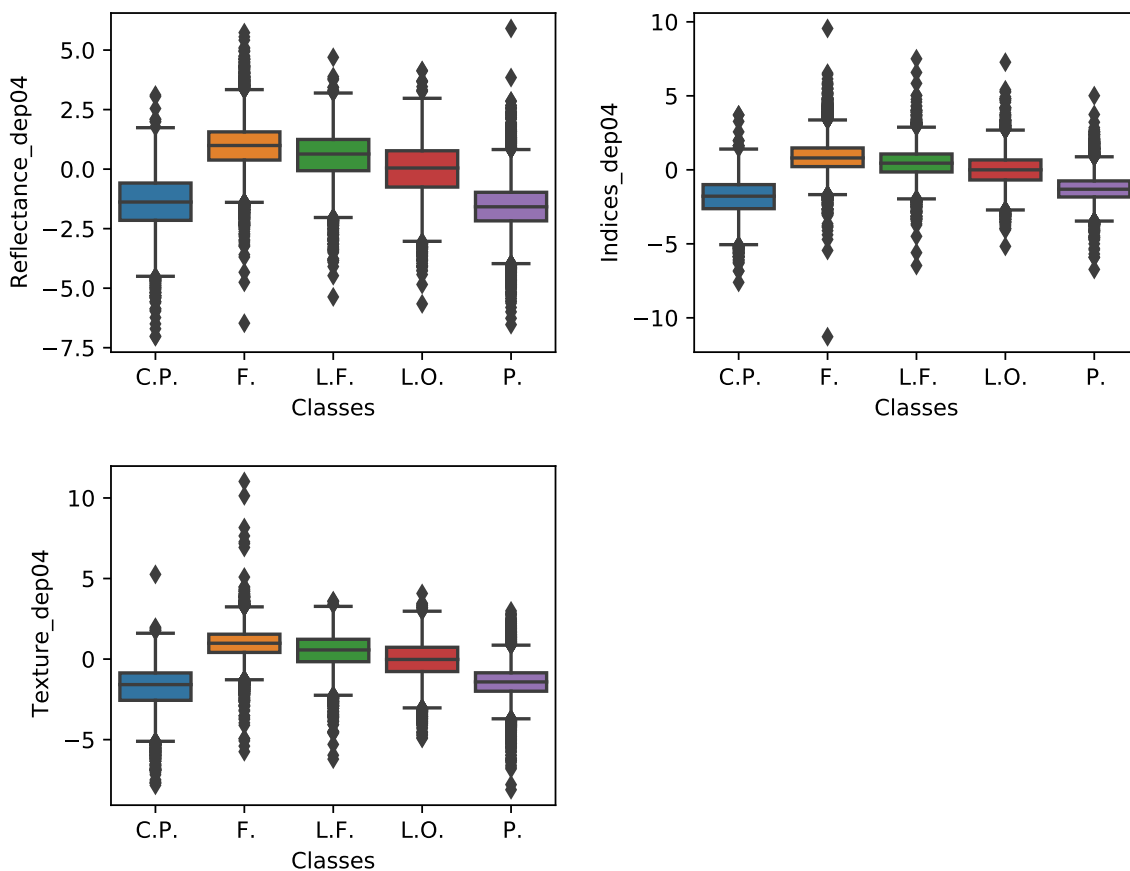


Figure 4.3 – Distribution des variables issues des premiers composants de l'ADL appliquée pour analyser la séparabilité des 5 classes à partir des variables de reflectance, spectrales et texturales, pour le département des Alpes-de-Haute-Provence (04).

A ce stade, nous avons estimé que les landes fermées, vu leur forte confusion avec les landes ouvertes et les forêts, risquent d'augmenter les imprécisions lors de la classification. De ce fait, nous avons opté pour le regroupement en une classe des landes fermées et ouvertes. Nous avons donc maintenu au final 4 classes : pelouse, lande, culture pérenne et forêt.

Sélection des jeux d'entraînement

Dans les données de référence, le nombre d'objets correspondant à chaque classe est très varié. Les landes sont minoritaires (cf. Tableau 4.5). Pour évaluer l'effet du poids des données d'entraînement pour chaque classe, nous avons constitué 3 jeux d'entraînement. Dans un premier temps, les données de référence ont été scindées en deux : 2/3 pour l'entraînement et 1/3 pour l'évaluation (le test) des modèles de classification. Les données d'entraînement ont par la suite été scindées en 3 jeux : (i) 50 000 objets pour chaque classe (Jeu 1), (ii) 200 000 objets pour les classes majoritaires et la totalité des objets pour les classes minoritaires (Jeu 2), et (iii) l'ensemble des objets de chaque classe (Jeu 3). Les résultats de ce processus sont présentés dans le Tableau 4.5.

Table 4.5 – Nombre d'objets extraits pour chaque classe retenue dans le RPG 2019 et la BD-Forêt V2.0 et 3 jeux de données utilisés pour l'entraînement et l'évaluation de l'algorithme de classification.

Classe	Effectif	Entraînement	Test	Jeu 1	Jeu 2	Jeu 3
Pelouse (RPG)	456991	305683	151308	50000	200000	305683
Lande (RPG)	170107	114156	55951	50000	114156	114156
Culture pérenne (RPG)	260756	174617	86139	50000	174617	174617
Forêt (BD-Forêt)	412705	276919	135786	50000	200000	276919
Total	1300559	871375	429184	200000	688773	871375

Choix de l'algorithme de classification

Les données acquises sont de sources différentes. En plus, les données RPG sont produites par un grand nombre d'exploitants. Plusieurs facteurs d'imprécision de mesure peuvent survenir pendant leur numérisation. Ces deux situations évoquées peuvent constituer une source de variabilité qui nécessite l'utilisation d'un algorithme performant. L'algorithme Random Forest (RF) (Breiman, 2001) est connu pour sa robustesse face aux données non équilibrées (Khalilia et al., 2011; Khoshgoftaar et al., 2007) et provenant des sources différentes (Frenay and Verleysen, 2014). C'est un algorithme de classification et de régression qui réalise un apprentissage basé sur un ensemble d'arbres de décision à partir d'une sélection aléatoire des échantillons au niveau de chaque arbre (Breiman, 2001; Fawagreh et al., 2014). Un tirage aléatoire avec remise des individus (bootstrap) est réalisé pour l'entraînement des arbres et les individus restants non tirés sont utilisés en interne en vue d'estimer la performance du modèle (Breiman, 2001; Belgiu and Drăguț, 2016). Le RF est également stable quant au choix des paramètres de traitement (Pelletier et al., 2016). Dans Lin et al. (2017), les auteurs ont également montré que le RF permet de mieux prédire les données massives, comparé à d'autres algorithmes de classification tel que le support-vector machine (SVM). L'ensemble de ces avantages du RF nous ont motivé à l'utiliser pour entraîner les modèles en vue de discriminer les classes retenues.

Sélection de variables prédictives

Le choix des variables prédictives a été réalisé pour maximiser les possibilités de discriminer les classes sélectionnées. Ne sachant pas à l'avance les variables déterminantes pour produire les meilleures performances des modèles sur les 98 sélectionnées, une méthode appropriée de réduction de dimensionnalité des variables a été nécessaire. Parmi les méthodes de réduction de dimensionnalité, celles qui sont intégrées dans les algorithmes de machine learning présentent un avantage. Elles prennent en compte l'interaction avec le modèle de classification pour définir la liste des variables importantes (Saeys et al., 2007). Nous avons utilisé les résultats de l'autoévaluation des performances du RF pour déterminer les variables importantes (Archer and Kimes, 2008; Breiman, 2001; Dewi and Chen, 2019).

Au cours de l'évaluation interne du modèle, l'algorithme attribue un score à chaque variable qui détermine son niveau d'importance dans l'apprentissage du modèle. Ce score d'importance est basé sur la décroissance moyenne de l'impureté lorsqu'une variable est choisie au niveau d'un noeud. De cette manière, les variables ayant des scores relativement élevés peuvent être choisies pour entraîner le modèle. A cause du caractère stochastique du RF, nous avons reproduit 20 fois l'entraînement du modèle sur le premier jeu de données (avec 50 000 objets par classe) pour définir les variables importantes à partir des moyennes et des écart-types de leurs scores. La Figure 4.4 présente le classement décroissant des variables en fonction de leurs scores d'importance. Nous avons sélectionné comme variables importantes celles dont les scores sont égaux ou supérieurs à 0,009. Cette limite inférieure a été fixée à partir du niveau présentant la faible décroissance du score des variables. Ainsi, 30 variables importantes sur les 98 ont été retenues.

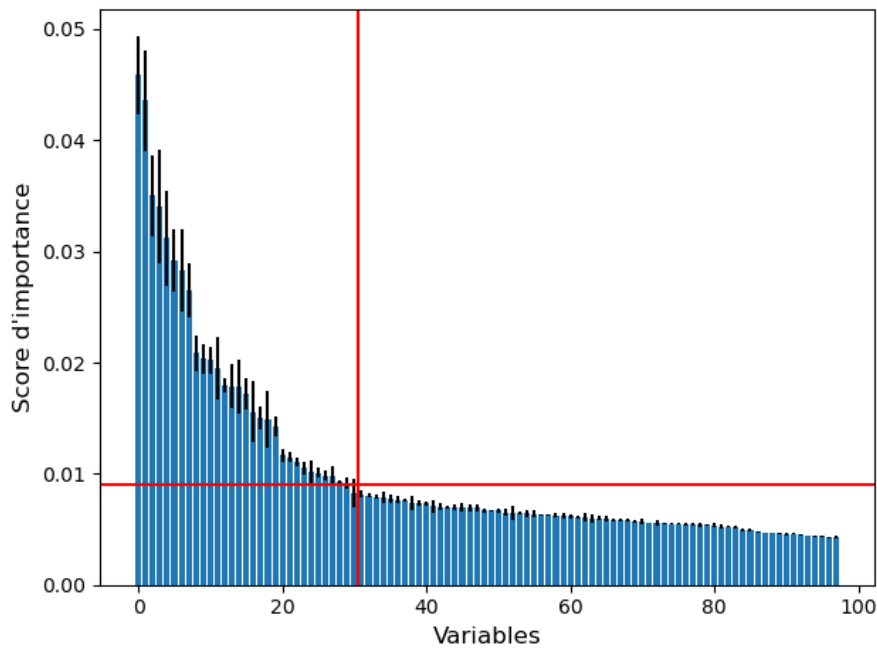


Figure 4.4 – Classement des variables en fonction de leurs scores d'importance après 20 itérations sur le jeu de données avec 50 000 objets par classe. Les lignes verticale et horizontale marquent respectivement les limites des variables importantes et du score minimum choisi. La barre bleue représente la moyenne et la noire l'écart-type associé pour les 20 itérations.

Entraînement et évaluation des modèles de classification

Pour évaluer les performances produites en utilisant l'ensemble des 98 variables de base et les 30 nouvelles variables réduites, nous avons entraîné l'algorithme sur les 3 jeux de données sans réduction de dimensionnalité et sur les 3 jeux de données avec réduction de dimensionnalité des variables prédictives (cf. Section 4.5).

Afin d'obtenir les meilleures performances de l'algorithme à entraîner sur les jeux de données préparés, une combinaison des valeurs de 2 paramètres importants du RF a été évaluée : le nombre d'arbres de décision à considérer (100, 250, 500, 750, 1000, 1250, 1500, 1750, 2000) et le nombre des variables à considérer à chaque noeud (4, 8, 12, racine carrée du nombre total). La meilleure valeur du nombre d'arbres a été de 500 pour les deux types de données. Et celle des variables au niveau des noeuds, de 10 et 5 respectivement pour les données à dimensionnalité non réduite et les données à dimensionnalité réduite. Nous avons par la suite réalisé une validation croisée à 10 blocs afin d'évaluer la stabilité des modèles pour l'entraînement des données utilisées.

Pour évaluer la performance des modèles de classification, nous avons utilisé la précision globale et le F-score comme indicateurs. Le F-score permet de réaliser un meilleur compromis entre la valeur de la précision et celle du rappel à partir de leur moyenne harmonique. Ces trois mesures sont présentées par les formules suivantes (Van Rijsbergen, 1979) :

$$Précision = \frac{VP}{VP + FP} \quad (4.1)$$

$$Rappel = \frac{VP}{VP + FN} \quad (4.2)$$

$$F - score = \frac{(\beta^2 + 1) * Précision * Rappel}{(\beta)^2 * Précision + Rappel} \quad (4.3)$$

où VP est le nombre des résultats positifs correctement classés comme positifs, FP le nombre des résultats négatifs incorrectement classés comme positifs, FN le nombre de résultats positifs incorrectement classés comme négatifs et β le paramètre qui contrôle la balance entre la précision et le rappel. Lorsque $\beta > 1$, le F-score devient plus orienté vers le rappel et si $\beta < 1$, il devient plus orienté vers la précision. Pour cette étude, la valeur choisie pour β est de 2, définie par défaut dans Scikit-Learn.

Une comparaison des précisions globales et des F-scores a été faite pour sélectionner le jeu de données permettant d'obtenir une meilleure discrimination des classes de parcours.

Tous les traitements décrits dans cette section ont été réalisés grâce aux modules de la librairie Scikit-learn⁷ sous python 3.8⁸.

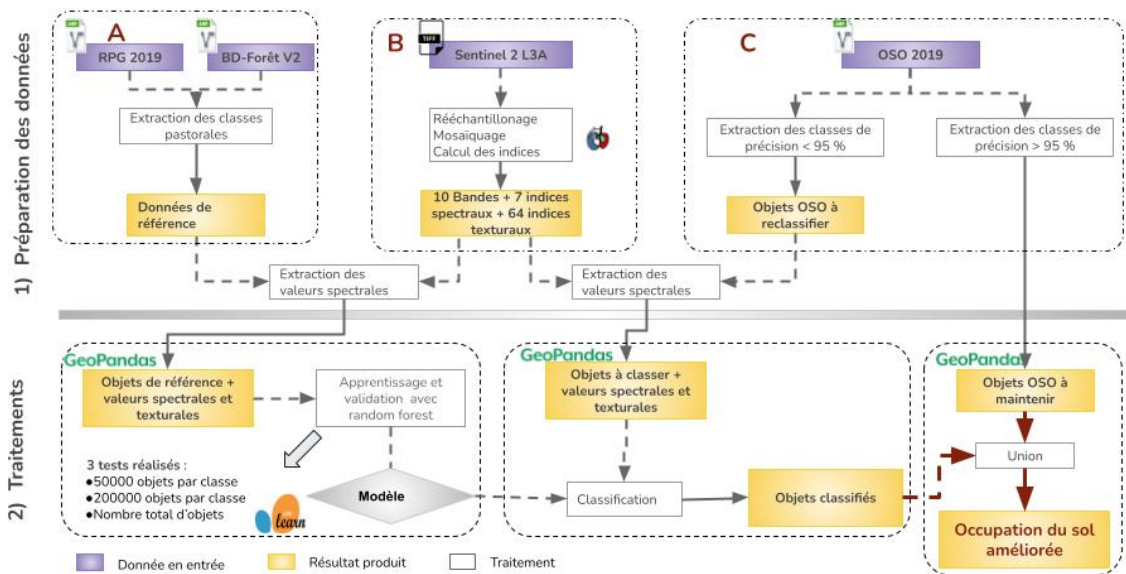


Figure 4.5 – Schéma général de traitement d'amélioration de la précision de la classification des classes pastorales et des cultures pérennes. Avec comme données en entrée : le RPG 2019 et la BD-Forêt (4.A), l'occupation des sols OSO 2019 (4.C) et les données Sentinel 2 L3A (4.B).

7. <https://scikit-learn.org/stable/>

8. <https://www.python.org/downloads/release/python-380/>

4.3.3 Résultats

Qualité de prédiction de classes

Les résultats de l'entraînement des modèles de classification sont présentés dans les Tableaux 4.6, 4.7 et 4.8 ci-dessous.

L'entraînement par le Random forest (RF) a produit des précisions globales stables pour les 3 jeux de données avec et sans réduction des variables, suite à la validation croisée, leurs écarts-types étant faibles. Les précisions globales produites pour les jeux de données avec réduction du nombre des variables sont respectivement $0,79 \pm 0,002$; $0,82 \pm 0,002$ et $0,83 \pm 0,003$. Celles des données sans réduction du nombre de variables sont respectivement $0,80 \pm 0,002$; $0,83 \pm 0,003$ et $0,83 \pm 0,003$. Les jeux de données à dimensionnalité non réduite ont produit de meilleures précisions globales comparés aux données à dimensionnalité réduite, sauf pour le jeu de données 3 (contenant l'ensemble d'objet des classes) où elles sont équivalentes.

Table 4.6 – Matrice de confusion et qualité de la prédiction produite à partir du modèle utilisant le jeu de données test 1.

Classes		Classification				
		Dimensionnalité réduite : précision globale de $0,79 \pm 0,002$				
		Pelouse	Lande	Culture P.	Forêt	Rappel
Référence	Pelouse	121458	20808	5553	3489	0,80
	Lande	9074	34822	1685	10370	0,62
	Culture P.	4104	3356	76896	1783	0,89
	Forêt	3832	23354	1415	107185	0,79
	Précision	0,88	0,42	0,90	0,87	
	F-score	0,84	0,50	0,90	0,83	
	Dimensionnalité non réduite : précision globale de $0,80 \pm 0,002$					
	Pelouse	Lande	Culture P.	Forêt	Rappel	
	Pelouse	122613	19531	6321	2843	0,81
	Lande	9004	34892	1725	10330	0,62
	Culture P.	4779	3791	75960	1609	0,88
	Forêt	3779	21175	1201	109631	0,81
	Précision	0,87	0,44	0,89	0,88	
	F-score	0,84	0,52	0,89	0,84	

Pour tous les jeux de données, les landes possèdent de faibles F-scores (de 0,42 à 0,50, et de 0,43 à 0,52 respectivement pour les jeux de données à dimensionnalité réduite et non réduite) comparé aux autres classes (pelouse, culture pérenne et forêt) qui possèdent des F-scores élevés (de 0,83 à 0,91, et de 0,84 à 0,90 respectivement pour les jeux de données à dimensionnalité réduite et non réduite). Les F-scores ont varié pour chaque classe selon les combinaisons de jeux de données d'entraînement et de variables. Aucune combinaison n'a été optimale, au sens d'une maximisation des F-scores pour toutes les classes. Il reste cependant difficile d'améliorer la discrimination de la classe des landes à cause de sa confusion avec toutes les autres classes.

En observant les matrices de confusion, on observe une forte confusion entre les landes et

les forêts, suivi des pelouses. Les landes sont plus en confusion avec les forêts que les cultures pérennes. Compte tenu des enjeux liés à la présence et à la répartition des ligneux dans les parcours, le choix du jeu de données et du nombre de variables serait orienté vers le modèle qui produit les meilleurs F-scores des landes. Il s'agit du jeu de donnée 1 (50 000 effectif d'objets d'entraînement pour chaque classe) sans réduction de dimensionnalité d'après les résultats obtenus. La qualité de discrimination des landes se trouve donc impactée par leur confusion avec les forêts, les cultures pérennes et les pelouses. Cependant, la discrimination des landes n'a pas été meilleure autant que les forêts, les cultures pérennes et les pelouses.

Table 4.7 – Matrice de confusion et qualité de la prédiction produite à partir du modèle utilisant le jeu de données test 2.

Classes		Classification					
		Dimensionnalité réduite : précision globale de 0,82 ±0,002					
		Pelouse	Lande	Culture P.	Forêt	Rappel	
Référence	Pelouse	131720	10234	4705	4649	0,87	
	Lande	13218	25357	1821	15555	0,45	
	Culture P.	4525	1669	77814	2131	0,90	
	Forêt	4922	11409	1406	118049	0,87	
	Précision	0,85	0,52	0,91	0,84		
	F-score	0,86	0,48	0,91	0,85		
			Dimensionnalité non réduite : précision globale de 0,83 ±0,003				
			Pelouse	Lande	Culture P.	Forêt	Rappel
	Pelouse	133196	9185	5054	3873	0,88	
	Lande	13338	25607	1832	15174	0,46	
	Culture P.	5278	1906	76962	1993	0,89	
	Forêt	4884	9789	1248	119865	0,88	
	Précision	0,85	0,55	0,90	0,85		
	F-score	0,86	0,50	0,90	0,87		

La réduction de dimensionnalité des variables a principalement affecté la classe des landes. Les F-scores de ces deux classes diminuent pour chaque jeu de données respectif quand la dimensionnalité des variables est réduite. Les classes des pelouses et des cultures pérennes ne sont pas impactées.

Table 4.8 – Matrice de confusion et qualité de la prédiction produite à partir du modèle utilisant le jeu de données test 3.

Classes		Classification					
		Dimensionnalité réduite : précision globale de 0,83 ±0,003					
		Pelouse	Lande	Culture P.	Forêt	Rappel	
Référence	Pelouse	137506	5356	3275	5171	0,91	
	Lande	17170	18406	1450	18925	0,33	
	Culture P.	6183	1024	76350	2582	0,89	
	Forêt	6289	6624	1046	121827	0,90	
	Précision	0,82	0,59	0,93	0,82		
	F-score	0,86	0,42	0,91	0,86		
			Dimensionnalité non réduite : précision globale de 0,83 ±0,003				
			Pelouse	Lande	Culture P.	Forêt	Rappel
	Pelouse	138773	4716	3456	4363	0,92	
	Lande	17343	18677	1453	18478	0,33	
	Culture P.	7368	1179	75196	2396	0,87	
	Forêt	6137	5465	948	123236	0,91	
	Précision	0,81	0,62	0,93	0,83		
	F-score	0,86	0,43	0,90	0,87		

Qualité de l'occupation des sols produite

Le modèle ayant produit le meilleur F-score des landes a été utilisé pour reclasser les objets OSO extraits (avec classes aux Faibles F-scores). Une visualisation des résultats sur deux différentes zones prises dans le département de l'Aude (11) est présentée à la Figure 4.6. Nous faisons une comparaison de la qualité de la nouvelle classification avec l'occupation des sols OSO originale et l'image Sentinel 2 L3A.

Pour la première zone, le résultat produit (la Figure 4.6 1a) montre une substitution des cultures pérennes dans le nord (la Figure 4.6 1b) par les forêts. De même, les cultures pérennes (1b) ont été substituées par les landes dans la nouvelle classification (1a) à l'est de la zone. Les forêts qui apparaissent au centre et au sud de la zone ont substitué les pelouses. On observe néanmoins une correspondance pour les cultures pérennes entre les deux produits, notamment vers le nord et le nord-est de la zone sélectionnée. En comparaison avec la composition colorée, le résultat produit présente des patches de forêts vers le sud de la zone tandis que vers le nord, il présente les forêts à la place d'une autre classe. Le produit OSO original présente plus les pelouses à la place des forêts notamment dans le sud et le centre de la zone. Une comparaison des deux produits avec la composition colorée n'a pas été faite pour les pelouses, les landes et les cultures pérennes étant donné qu'il est difficile de les différencier visuellement par photo-interprétation.

Les mêmes observations peuvent être faites pour la zone 2 où l'on remarque une occupation par les cultures pérennes et les landes dans le résultat produit (2a) vers l'est de la zone à la place des forêts dans OSO (2b). Vers le nord et le sud, les forêts ont remplacé les pelouses. En comparaison avec la composition colorée, le produit OSO d'origine (2b) présente mieux les forêts à l'est de la zone alors que le résultat obtenu (2a) montre une occupation par les landes et les cultures pérennes. Dans le nord de la zone, les deux produits présentent une

répartition variable des forêts comparé à la composition colorée.

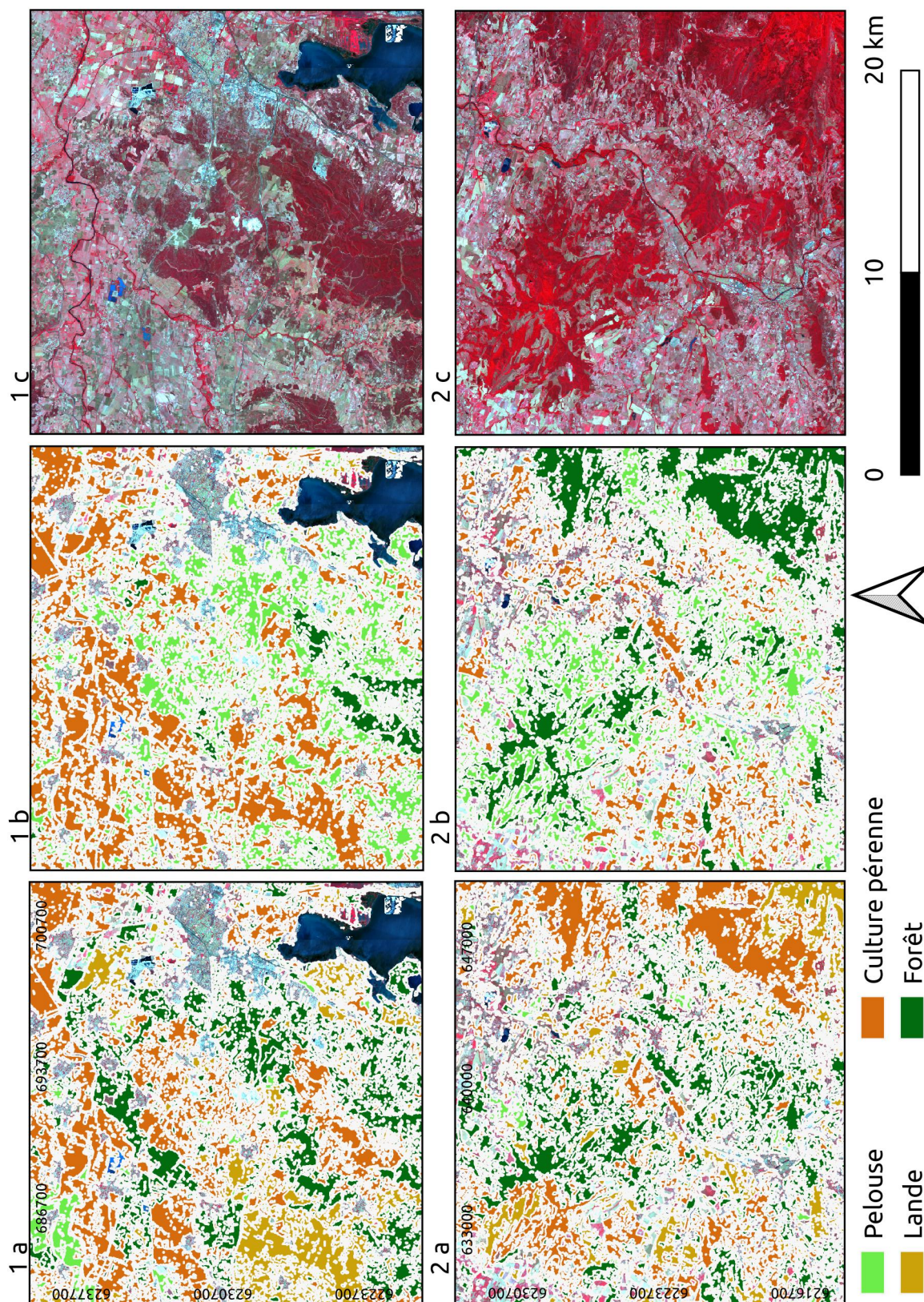


Figure 4.6 – Comparaison entre l'occupation des sols produite, OSO 2019 et la composition colorée infrarouge fausses couleurs.

4.4 Prétraitements des données de localisation des troupeaux

Les données de localisation des troupeaux sont importantes car elles traduisent le comportement des animaux au sein des exploitations pastorales. Dans le cadre du projet P@stor_All, elles sont collectées au niveau des exploitations pastorales privées et expérimentales de l'UMT Pasto⁹ (La Fage et Le Merle). Comme nous l'avons indiqué dans la section 4.2, la précision des données GPS dépend des caractéristiques propres à l'outil. C'est ainsi que le choix du type de capteur GPS est primordial pour assurer la meilleure précision des données collectées.

Dans le cadre du projet P@stor_All, deux marques de GPS ont été choisies pour la collecte des données : GeoVie G1000¹⁰ et E-Pasto commercialisé par la compagnie Aguila Technologies¹¹. Une balise GeoVie a été distribuée dans chaque exploitation pastorale privée et expérimentale. Environ 30 balises GPS E-Pasto ont quant à elles été uniquement déployées au sein des exploitations expérimentales pour enregistrer les données en continu sur les périodes de pâturage en vue d'analyser le comportement des animaux. La configuration de la fréquence temporelle de référence pour l'utilisation de ces balises GPS correspond à celle établie dans le cadre des activités du projet CLOChèTE qui est de 5 minutes. Au bout de 24 heures, dans les cas idéal (ce qui est rarement le cas), un nombre de 288 points GPS devrait être collecté.

Ces types de capteurs GPS ont été conçus pour résister aux contraintes caractéristiques des parcours méditerranéen et tropicaux et enregistrent des données avec une bonne précision pour l'élevage pastoral. Cependant, certaines conditions (milieux fermés, météorologie particulière) peuvent quelques fois perturber la précision des données enregistrées.

4.4.1 Caractéristiques des données GPS collectées

Les caractéristiques des données GPS collectées par les différentes exploitations pastorales sont présentées dans le Tableau 4.9.

-
9. https://idele.fr/umt_pasto/
 10. <https://www.geovie.eu/>
 11. <https://www.aguila.fr/societe/qui-sommes-nous>

Table 4.9 – Types d'utilisation des balises GPS au sein des exploitations pastorales privées et expérimentales.

Exploitation	Caractéristiques des données GPS
Privées	<p><i>Caractéristiques</i></p> <ul style="list-style-type: none"> - 7 exploitations pastorales - Balises GPS GéoVie déployées dès l'été 2021 - Une balise GPS embarquée sur un animal du troupeau - Fréquence d'acquisition configurable (peut aller de 30 secondes à 1 heure) - Format d'enregistrement : XML <p><i>Utilisation potentielle</i></p> <ul style="list-style-type: none"> - Localisation des troupeaux à travers l'interface utilisateur - Identification des milieux utilisés par les troupeaux - Suivi de l'utilisation des parcours : trajectoire, distance parcourue, ect. - Comportement des animaux
Ferme expérimentale La Fage	<p><i>Caractéristiques</i></p> <ul style="list-style-type: none"> - 23 balises E-Pasto déployées entre l'été et l'automne 2021 - Fréquence d'acquisition de 5 minutes - Pâturage en parcs : 2 lots d'agnelles utilisent des parcs différents à chaque saison de pâturage - 12 balises utilisées dans un lot de 20 agnelles naturellement allaitées - 11 balises utilisées dans un lot de 20 agnelles artificiellement allaitées - Format d'enregistrement : CSV <p><i>Utilisation potentielle</i></p> <ul style="list-style-type: none"> - Analyse de l'utilisation spatiale des parcours par les deux lots d'agnelles * Intensité d'utilisation des zones pastorales * Dispersion des animaux au cours du pâturage * Utilisation des massifs de ligneux
Ferme expérimentale Le Merle	<p><i>Caractéristiques</i></p> <ul style="list-style-type: none"> - 23 balises E-Pasto déployées du printemps à l'hiver 2022 - Fréquence d'acquisition de 5 minutes - Pâturage en milieux variés : début de printemps sur COUSSOUL, fin de printemps en COLLINES (Eygalières), été en montagne (ESTIVE dans le Mercantour) automne sur prairies (PRES DE CRAU) - Format d'enregistrement : CSV <p><i>Utilisation potentielle</i></p> <ul style="list-style-type: none"> - Analyse de l'influence de la variation des milieux sur le comportement des animaux au pâturage. Avec deux granularités de variation : i) changement de conditions pastorales entre saisons, et ii) variations de conditions du milieu (en particulier météo, mais aussi évolution de la disponibilité de la ressource) au cours de l'utilisation de chaque milieu. Analyse basée sur : * La distance journalière parcourue * L'intensité ou temps de pâturage * La dispersion des animaux au cours du pâturage * L'utilisation des massifs de ligneux

4.4.2 Préparation des données GPS

Les différences de ces données GPS peuvent concerner leur formats (GPX, XML, ou CSV). Ainsi, un reformatage aux formats binaires (SHP et gpkg) ou textuel (GeoJSON) est réalisé afin de préparer ces données pour des analyses et des requêtes spatiales et la recherche

et les agrégations à partir des moteurs dédiés tels qu'Elastic Search. Lors de l'analyse des données au sein du SI, les filtres spatial et temporel peuvent être appliqués aux données afin d'éliminer les informations qui ne concernent pas l'étendue spatiale ou la période visée.

Les données de localisation des animaux traduisent leurs activités spatio-temporelles. Il est dès lors utile de transformer ces données sous la forme de Tableau spatio-temporel qui constituera la base de toute analyse (cf. Tableau 4.10). La dimension temporelle est stockée au format timestamp qui correspond à une mesure de date qui représente le nombre de secondes écoulées depuis le 1^{er} janvier 1970. Ce format a plusieurs avantages :

- Simplicité d'enregistrement par rapport à d'autres format de temps ;
- Meilleure précision temporelle (à la seconde voire à la millisecondes) ;
- Interopérabilité avec différents langages de programmation et les systèmes de gestion de base de données ;
- Indépendance avec tous les fuseaux horaires.

Table 4.10 – Exemple de tableau spatio-temporel des données de localisation des animaux.

#	Timestamp	Longitude	Latitude
1	1624966877	43,91839	3,094610
2	1624967177	43,92163	3,092905
3	1624967477	43,91819	3,094735
4	1624967777	43,92165	3,092833
5	1624968077	43,92166	3,092793
...

4.5 Densité de végétation et topographie du milieu

Dans cette section, nous présentons les différentes procédures de production des données de densité de végétation et des variables topographiques.

4.5.1 Densité de végétation

Le comportement alimentaire et spatial des animaux est directement lié à la disponibilité (Chebli et al., 2020) et à la distribution spatiale (Wang et al., 2009b) des ressources pastorales. Ces deux facteurs varient en fonction des saisons et de l'utilisation à travers le pâturage par les animaux. Pour mieux les caractériser, les données fréquentes sur l'état de la végétation sont importantes.

Le choix des données a été réalisé sur base des critères de résolution spatiale et temporelle. Actuellement, les deux sources pouvant permettre de répondre à ce critère sont les données Sentinel 2 et Planet. La fréquence d'acquisition des deux satellites Sentinel 2 permet de

couvrir les zones de la terre à une fréquence de 5 jours. Ceci constitue une opportunité pour suivre l'état des ressources pastorales. Cependant, la seule contrainte pour les données Sentinel 2 est la présence de couvertures nuageuses pour la plupart des zones (montagneuses notamment) utilisées pour le pâturage des animaux. Lors de l'acquisition des données Sentinel 2 pour une zones explorée par les animaux au domaine expérimental de la Fage, nous avons observé la présence de nuages sur plus de la moitié des données pour une période d'environ 30 jours pour l'été 2021. Les données issues de la constellation Planet¹², en plus de leur résolution spatiale de 3m, possèdent une fréquence d'acquisition quotidienne. Ce qui permet à la fois d'obtenir des données de meilleure précision spatiale et de maximiser les possibilité d'obtenir les données sans couvertures nuageuses. Une utilisation de ce type de données pourrait également être une alternative pour palier la présence de nuages des données Sentinel 2.

La production de données de densité de végétation se base sur le calcul de deux indices spectraux : l'indice différentiel de végétation normalisé (Rouse et al., 1974) (NDVI, équation 4.4), et l'indice de végétation optimisé ajusté par rapport au sol (Rondeaux et al., 1996) (OSAVI, équation 4.5), construits à partir des bandes R et proche infrarouge (Rouse et al., 1974). Cet indice permet de caractériser la quantité de biomasse d'un milieu. Dans notre cas d'étude, l'utilisation des données temporelles permet de suivre l'évolution de la biomasse au fur et mesure de l'utilisation des parcours par les troupeaux d'animaux (Ma et al., 2019; Cheng et al., 2021).

$$NDVI = \frac{PIR - R}{PIR + R} \quad (4.4)$$

$$OSAVI = \frac{(PIR - R)}{PIR + R + 0,16} \quad (4.5)$$

Où PIR correspond à la bande proche infrarouge, R à la bande rouge. 0,16 correspond au facteur d'ajustement de l'indice.

Les données sont par la suite ré-échantillonnées à une résolution spatiale de 10m pour maintenir la cohérence spatiale.

4.5.2 Pente et orientation

Les caractéristiques topographiques des milieux influencent également l'utilisation des ressources pastorales par les animaux. Ces derniers ont tendance à éviter les zones de fortes pentes (Bailey et al., 2015) et valorisent plus des zones facilement accessibles. Ces zones inaccessibles par les animaux présentent un fort risque d'embroussaillage par le fait qu'elles sont évitées. L'analyse de l'influence de la pente et de l'orientation du terrain peut également

12. <https://www.planet.com/products/>

être faite afin de mieux comprendre leur influence sur les comportements des animaux.

Les données MNT RGE ALTI ont été utilisées à cette fin. Elles ont été collectées pour tous les départements des régions occitanie et Provence-Alpes-Côte d'Azur à partir de la plateforme de l'IGN. La production de la pente et de l'orientation du terrain a été réalisée grâce aux fonctions de calcul morphométrique implémentées dans SAGA GIS¹³. La chaîne de traitement a consisté en la production des pentes et orientations pour chaque département puis au rééchantillonnage des données résultantes à 10 m de résolution spatiale. Étant donné l'aspect statique de la topographie du milieu, les données de pente et d'orientation produites n'auront pas besoin d'une actualisation.

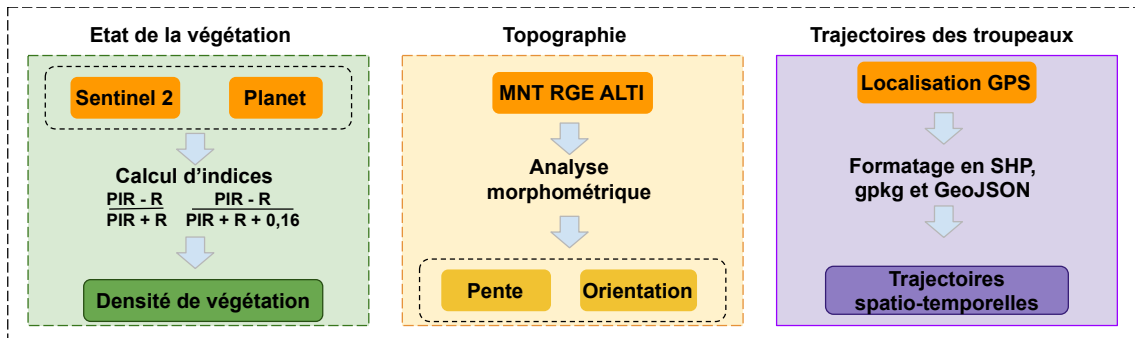


Figure 4.7 – Aperçu des prétraitements réalisés pour la production des données de densité de végétation, des variables topographiques et des trajectoires de troupeaux.

4.6 Conclusion

Dans ce chapitre, nous avons présenté les différents processus de traitement et de préparation des données spatiales en vue de rehausser leur qualité pour leur intégration dans le système d'information. La qualité des données spatiales a été évaluée par rapport à la précision, l'échelle de représentation et la sémantique, et la granularité temporelle. Nous nous sommes basés à la fois sur la contrainte imposée par les données et les possibilités de représenter les phénomènes au sein des exploitations pastorales. Les critères minimums pour représenter ces phénomènes doivent être considérés : résolution spatiale de 10m pour les données raster, précision de l'occupation des sols minimisant la confusion des classes pastorales, fréquence d'acquisition de 1 à 5 jours pour les données de végétation et de 5 minutes pour la localisation des troupeaux. Pour rejoindre ces critères, la préparation des données a concerné : l'amélioration des F-scores des classes pastorales, la production des densités de végétation, de la pente et de l'orientation, le formatage des données d'occupation des sols, et le prétraitement des données GPS.

L'amélioration de l'occupation des sols a concerné les landes, les pelouses, les cultures pérennes (vignes et vergers) et les forêts. Les résultats produits à partir du modèle Random

13. <https://saga-gis.sourceforge.io/en/index.html>

Forest optimal ont donné de bonnes précisions pour toutes les classes sauf pour les landes à cause de leur forte confusion avec les forêts et les pelouses. L'utilisation des données temporelles de haute résolution spatiale (telle que SPOT 6/7, avec 1,5m et 6m de résolution spatiale respectivement pour les bandes panchromatique et multispectrales) pourrait offrir la possibilité de mieux discriminer les landes, spécifiquement à partir de leur texture. La seule contrainte pourrait être la disponibilité de ce type de données pour couvrir l'ensemble de la région et le coût élevé de calcul. Pour pallier cette contrainte, les analyses peuvent être faites sur de petites surfaces (telles que les communes), suivies d'une agrégation des résultats pour représenter toute les régions concernées.

Pour décrire les ressources pastorales, l'information de densité de végétation produite à partir du NDVI et de l'OSAVI peut compléter l'occupation des sols. Ces deux indices de biomasse végétale peuvent apporter à la fois un niveau de détail sur l'état de la végétation et sur son évolution au fil du temps et de l'utilisation par les animaux. Ces indices peuvent être produits à partir des séries temporelles Sentinel 2 ou Planet. La fréquence temporelle d'analyse peut ainsi être adaptée en fonction des objectifs des utilisateurs.

Les données de localisation des animaux utilisées pour décrire leur comportement durant le pâturage sont collectées sous différentes conditions en fonctions des exploitations privées et expérimentales. Une harmonisation de ces données en formats interopérables binaires (SHP, gpkg), et textuel (GeoJSON) facilite leur exploration par les utilisateurs. Pour réaliser des analyses spatio-temporelles sur ces données, une structuration sous forme de tables spatio-temporelle est utile.

Enfin, cette approche préparatoire des données privilégie la conservation des informations détaillées que peuvent avoir les données brutes. Elle se base plus sur l'harmonisation des caractéristiques des données et sur la production des données dérivées permettant d'apporter une complémentarité pour la caractérisation des dynamiques au sein d'un écosystème agro-sylvo-pastoral.

Chapitre 5

Croisement des informations spatiales pour la production d'indicateurs d'aide à la décision en élevage pastoral

Dans le chapitre 3 nous avons abordé les questions liées à la définition et à la conception d'un système d'information destiné à servir comme référentiel pour intégrer les données hétérogènes du pastoralisme. Ce lac de données spatiales dispose des fonctionnalités offertes par les modules d'Hadoop HDFS ainsi que l'implémentation de gestion de métadonnées spatiales grâce à GéoNetWork. Ces fonctionnalités lui confèrent de bonnes capacités de gestion des données. Certains composants de requête et de recherche, respectivement basés sur PostgreSQL et la suite Elasticsearch, intègrent également des fonctionnalités d'exploration des données stockées.

Cependant, pour assurer la prise de décision en élevage pastoral, le lac de données doit proposer plus que le simple stockage et l'exploration des données. L'implémentation des modules analytiques est importante en vue de transformer les informations ingérées en de nouvelles informations avec plus de valeur, pour appuyer les éleveurs pastoraux dans la gestion à court et à long terme du pâturage des animaux.

Dans ce chapitre, nous abordons les questions liées à l'analyse des données spatiales dans le but de produire des bouquets d'informations utiles pour la documentation du fonctionnement au sein des exploitations pastorales en zones méditerranéennes françaises.

5.1 Introduction

La variabilité et l'hétérogénéité des parcours en zones méditerranéennes françaises, en termes de végétation, de topographie, créent une grande variabilité dans le mode de conduite du pâturage dans les exploitations pastorales. Cette variabilité est perceptible entre les exploitations et au sein des exploitations pastorales dans le temps et dans l'espace, obligeant ainsi les éleveurs pastoraux à ajuster continuellement la conduite du pâturage à court et à long terme. Des adaptations à court terme peuvent inclure des modifications du circuit de pâturage quotidien, du temps et de la répartition des animaux au sein des parcs, une réduction temporaire des performances des animaux ou une complémentation. Des adaptations à long terme peuvent quant à elles inclure la mobilité vers de nouvelles zones de pâturage, des changements dans l'utilisation saisonnière des ressources pastorales, l'aménagement de points d'eau et de clôtures ou le défrichement (de certaines zones - en général de couloirs). Afin d'avoir une maîtrise de la production et d'accroître l'efficacité dans la prise de décision, la connaissance des différents facteurs qui caractérisent l'écosystème agro-sylvo-pastoral et de leurs impacts sur le mode d'utilisation des parcours par les animaux est utile pour les éleveurs pastoraux.

La modélisation conceptuelle du comportement alimentaire et spatial des animaux présentée au chapitre 3, identifie les composants ressources pastorales, environnement, animaux et mode de conduite, comme les principaux facteurs qui impactent l'utilisation des parcours par les animaux. L'appréhension de ces facteurs contribuera à la compréhension du comportement des animaux, caractérisé par la manière dont ils se déplacent et occupent les milieux explorés. Cette modélisation a permis d'identifier les données hétérogènes permettant de caractériser ces facteurs, dont la qualité pour documenter les exploitations pastorales a été présentée dans le précédent chapitre 4.

Les informations contenues dans ces différentes sources de données hétérogènes, offrent un potentiel non négligeable pour documenter le fonctionnement d'un système agro-sylvo-pastoral. Cependant, ces informations prises individuellement ont souvent de faible précision pour mieux représenter les phénomènes spatiaux (Bahlo et al., 2019). Certains éleveurs peuvent être amenés à utiliser ces informations de manière partielle, avec des succès variables quant à la contribution à la prise de décision. La valeur ajoutée de ces données peut être trouvée dans la combinaison de différentes informations qu'elles possèdent (Capalbo et al., 2017), afin de permettre aux éleveurs d'adapter la gestion du pâturage en fonction des contraintes spécifiques à chaque milieu exploré. Le défi reste donc de trouver des moyens méthodologiques et analytiques pertinents permettant de combiner ces informations (Bahlo et al., 2019). En plus de ces considérations, les analyses doivent conduire à la production d'indicateurs simples et utilisables par les éleveurs pour être facilement adoptés dans le processus de prise de décision.

De nos jours, plusieurs expérimentations dans le cadre de l'élevage ont exploré les poten-

tialités d'usage du GPS pour des applications diverses : (i) localiser et suivre les trajectoires des animaux pâturant librement les milieux (Buerkert and Schlecht, 2009; Akasbi et al., 2012), (ii) évaluer les différences dans l'utilisation spatio-temporelles des milieux par les animaux (Feldt and Schlecht, 2016), (iii) évaluer les comportements des animaux et leur interaction avec l'environnement (Handcock et al., 2009), et (iv) évaluer la variation du temps de pâturage en fonction de types de milieux utilisés (Rutter et al., 1997). Cependant, ces travaux proposent des indicateurs limités aux contextes spécifiques à leurs expérimentations et n'approfondissent pas les analyses pour décrire l'influence des facteurs du milieu et de la végétation sur le comportement des animaux. Dans des situations complexes comme celles rencontrées dans les élevages pastoraux, il serait plus efficace d'identifier d'autres types d'analyse complémentaires pour parvenir à une bonne caractérisation du comportement des animaux et de l'influence des facteurs du milieu et de la végétation pastorale.

Dans ce chapitre, nous nous intéressons à deux principales questions : "Comment les différentes données identifiées, peuvent-elles être utilement combinées en vue de contribuer à la documentation du fonctionnement des exploitations pastorales, et de caractériser l'influence des différents facteurs sur le comportement des troupeaux au pâturage?", et "Comment produire des indicateurs utiles afin de répondre aux besoins des éleveurs pastoraux et de les accompagner dans le processus de prise de décision?". Pour répondre à ces questions, nous présentons dans la section 5.2, les différentes approches adaptées à l'analyse des données de spatio-temporelles de localisation des animaux dans le but de produire des indicateurs spatio-temporels. Dans la section 5.3 nous présentons les approches de croisement des informations spatiales afin de caractériser l'influence des facteurs du milieu sur le comportement des animaux. Pour l'évaluation de ces différentes approches, avons expérimenté leur application dans une situation d'élevage en conditions expérimentales que nous présentons à la section 5.4. Dans la section 5.4.6 nous discuterons des résultats obtenus et de leur pertinence pour la prise de décision en élevage pastoral. Enfin, la Section 5.4.7 conclura ce chapitre.

5.2 Approches d'analyse des données spatio-temporelles spatiales pour la production d'indicateurs du pastoralisme

Le principal objet de cette section concerne l'analyse des données de localisation des troupeaux pour la production d'indicateurs qui traduisent leur mode d'utilisation des parcours. Les données de localisation sont à la fois caractérisées par leur dimensions spatiale et temporelle qui représentent l'occupation de l'espace pastoral au fil du temps par les animaux.

La production d'indicateurs spatiaux et temporels qui traduisent l'utilisation de l'espace par les animaux dépend de la nature des données. Nous avons identifié différentes analyses permettant d'explorer les possibilité d'extraire des connaissances de ces données spatio-temporelles. Nous les regroupons en 2 : (i) l'extraction de séquences fréquentes et (ii) l'analyse de motifs de points. Ces deux concepts sont présentés dans les sous sections qui suivent.

5.2.1 Extraction de séquences fréquentes

L'extraction des séquences fréquentes est une technique d'exploration des données qui permet d'extraire des connaissances au sein des données issues des phénomènes qui se produisent au fil du temps. Elle permet de présenter les événements qui apparaissent de manière successive. L'intuition de ce concept se fonde sur le fait que les phénomènes qui se produisent dans la nature, sont liés par une relation de succession ou d'ordre d'apparition.

L'extraction des motifs séquentiel s'applique à différents domaines : bio-informatique (Wang et al., 2007), e-learning (Fournier-Viger et al., 2008), analyse du panier de la ménagère (Srikant and Agrawal, 1996), analyse de texte (Pokou et al., 2016). Elle étend également ses applications à l'analyse de séries de données temporelles (Lin et al., 2007).

Pour comprendre les séquences fréquentes, nous allons tout d'abord définir quelques concepts qui leur sont liés (dans l'encadré ci-dessous d'après les éléments tirés de Fournier-Viger and Lin (2017)) : les itemsets, les items, la séquence ainsi que la base de données de séquence. Nous nous basons sur l'article de Fournier-Viger and Lin (2017), où les auteurs présentent un état de l'art sur l'extraction des motifs fréquents. Nous représenterons également en quoi ils peuvent être associés à l'utilisation des milieux pastoraux.

Soit un ensemble de éléments (items) $I = \{i_1, i_2, \dots, i_m\}$. Un itemsets X est un ensemble d'éléments (items) tel que $X \subseteq I$. La notation $|X|$ désigne la cardinalité de l'ensemble ou, en d'autres termes, le nombre d'items dans un itemset X . Un itemset X est dit de longueur k ou un k -itemset s'il contient k items ($|X| = k$). Par exemple, considérons que l'ensemble de symboles $I = a, b, c, d, e, f, g$ représente les produits vendus dans un magasin de détail. L'ensemble a, b, c est un itemsets contenant trois articles, qui peuvent représenter les achats effectués par un client un jour donné.

Sans perte de généralité, supposons qu'il existe un ordre total sur les articles, tel que l'ordre lexicographique (par exemple, $\{a \prec b \prec c \prec d \prec e \prec f \prec g\}$). Une séquence est une liste ordonnée d'éléments $s = I_1, I_2, \dots, I_n$ telle que $I_k \subseteq I (1 \leq k \leq n)$. Par exemple, considérons la séquence $(\{a, b\}, \{c\}, \{f, g\}, \{g\}, \{e\})$ représentant cinq transactions effectuées par un client dans un magasin de détail. Chaque lettre représente un article. Les éléments entre crochets représentent un ensemble d'articles. Cette séquence indique qu'un client a acheté les articles a et b en même temps, puis a acheté l'article c , puis a acheté les articles f et g en même temps, puis a acheté g , et enfin a acheté e .

Une séquence $s_a = (A_1, A_2, \dots, A_n)$ est dite de longueur k ou une k -séquence si elle contient k éléments, ou en d'autres termes si $k = |A_1| + |A_2| + \dots + |A_n|$. Par exemple, la séquence $(\{a, b\}, \{c\}, \{f, g\}, \{g\}, \{e\})$ est une 7-séquence.

Une base de données de séquences SDB est une liste de séquences $SDB = (s_1, s_2, \dots, s_i)$ ayant des identités de séquence (SIDs) $1, 2, \dots, p$. Par exemple, une base de données de séquences est présentée dans le Tableau 5.1, qui contient quatre séquences ayant les SIDs 1, 2, 3 et 4. Ces séquences peuvent, par exemple, représenter des achats effectués par quatre clients.

Table 5.1 – Base de données de séquences

ID_séquence	Séquence
1	$(\{a, b\}, \{c\}, \{f, g\}, \{g\}, \{e\})$
2	$(\{a, d\}, \{c\}, \{b\}, \{a, b, e, f\})$
3	$(\{a\}, \{b\}, \{f, g\}, \{e\})$
4	$(\{b\}\{f, g\})$

Le but de l'extraction de motifs séquentiels est de découvrir des sous-séquences intéressantes dans une base de données de séquences, c'est-à-dire des relations séquentielles entre des éléments qui sont intéressants pour l'utilisateur. La fréquence d'apparition de ces sous séquences peut être représentée dans l'intervalle d'un nombre compris entre 0 et 1 appelé support.

Dans la conduite du pâturage sur parcours, les animaux se déplacent et opèrent des

choix alimentaires en fonction des facteurs identifiés. Ces choix résultent en la succession de l'utilisation des lieux identifiés par les animaux et les éleveurs pour assurer leurs besoins en fonction des périodes journalières. En pâturage libre, les animaux se construisent des circuits qu'ils parcourent quotidiennement. En pâturage en parc, ils réalisent des aller-retours entre différentes zones au cours de la journée. Et En pâturage gardé, l'éleveur conduit son troupeau dans une suite de secteurs identifiés au préalable pour assurer leurs besoins. Ainsi, l'application des concepts d'extraction de séquences fréquentes peut permettre d'extraire des connaissances sur les endroits fréquemment utilisés par les animaux.

Dans ce cas, l'itemset pour l'élevage pastoral correspondra à l'ensemble des zones du milieu utilisables par les animaux. La combinaison des différents endroits utilisés (items) chaque jour par les animaux forment les séquences d'utilisation journalières pour la composition de la base de données. Une transposition de la base de données devra donc être adaptée à cette définition.

5.2.2 Analyse de motifs de points

Un motif de points est un ensemble de localisation (I_1, I_2, \dots, I_n) qui représentent les événements enregistrés dans une région d'étude donnée (R) (Gatrell et al., 1996). La distribution des points qui constituent un motif est souvent influencée par deux propriétés (Wiegand and Moloney, 2014) : de premier et de second ordre. La propriété de premier ordre concerne les variables qui influencent la variation dans l'intensité de distribution des points qui représentent un événement. Ces variables correspondent souvent aux facteurs du milieu. La propriété de second ordre concerne les événements qui résultent de l'interaction entre les éléments symbolisé par les points (par exemple, les animaux utilisent souvent les zones proches de certaines localisation d'intérêt particulier telles qu'un point d'eau, une mangeoires, ou une zone de repos).

Dans le cadre de l'élevage pastoral en zones méditerranéennes françaises, les motifs de points peuvent représenter la distribution des animaux d'un troupeau soit à partir d'un animal leader ou explorateur, représentatif du troupeau (Grisot et al., 2022), porteur de GPS, soit à partir d'un nombre d'animaux porteurs de GPS. Le premier cas représente la situation souvent rencontrée dans les élevages privés ayant tout au plus un GPS pour suivre leurs troupeaux. En effet, les élevages équipés ne possèdent pas plusieurs GPS pour des raisons de coût d'acquisition et d'abonnement pour l'accès aux services de géolocalisation. Le second cas représente la situation des exploitations pastorales expérimentales, notamment les sites de l'UMT Pasto¹ tels que la Fage et le Merle qui possèdent quant à eux plus de moyens d'expérimenter le potentiel des données GPS en termes d'indicateur d'aide à la conduite du pâturage. L'équipement de plusieurs animaux dans les sites expérimentaux permet de palier à l'insuffisance des données en cas de problème au niveau de certains GPS et d'acquérir des données plus réalistes sur la distribution des animaux (en cas de séparation

1. https://idele.fr/umt_pasto/

en deux ou plusieurs groupes par exemple). Il permet ainsi d'étudier avec plus de précision le comportement des animaux au pâturage. Les méthodes testées et validées à partir des données des sites expérimentaux pourront être utilisées en situation opérationnelle dans les élevages privés.

Si on imagine les milieux pastoraux comme étant neutres, c'est-à-dire non polarisés, on peut considérer que les animaux se répartissent aléatoirement dans l'espace. Cependant, ceci n'est pas le cas dans la vie pratique. La distribution des phénomènes spatiaux est toujours gouvernée par l'action des facteurs au sein des milieux et aussi par leurs interactions. Pour le cas de l'élevage pastoral ces facteurs correspondent à : i) la disponibilité, l'accessibilité et la distribution spatiale des ressources pastorales, ii) la météorologie, la topographie et les points d'attraction, iii) la physiologie, la morphologie, la mémoire spatiale et les relations sociales des animaux, et iv) le mode de conduite. Ces propriétés définissent également la distribution de l'utilisation des parcours par les animaux. Elles sont à la base du comportement alimentaire et spatial des animaux que nous avons définis au chapitre 3.

Les différences dans l'utilisation des parcours peuvent s'observer à travers la variation de la dispersion des animaux à différents moments de la journée. Quelques situations peuvent être observées :

- Tout d'abord, lorsqu'ils sont au pâturage, les animaux se dispersent plus en explorant la ressource alimentaire et se resserrent au moments de repos ou de rumination (Lécrivain et al., 1993).
- Durant les moments de pâturage, l'étendue de la surface pâturée et la distribution spatiale de la ressource pastorale peuvent faire varier la dispersion des animaux (Wang et al., 2009b). Ils auront tendance à se disperser plus lorsqu'ils pâturent de grandes étendues alors qu'ils seront plus resserrés lorsqu'ils pâturent des zones plus petites. Dans un milieu où la végétation pastorale est aléatoirement distribuée, les animaux auront également tendance à se disperser plus pour pâturer qu'en cas de végétation pastorale continue.
- L'effet de la pente peut aussi être remarquée (Bailey et al., 2015). En évitant de se déplacer sur de longues distances dans des zones à fortes pentes, les animaux auront tendance à rester plus serrés durant le pâturage que lorsqu'ils se trouvent dans des zones à pente modérée.
- Les conditions météorologiques extrêmes peuvent également modifier les temps de pâturage des animaux (Manteca i Vilanova and J. Smith, 2014). En été par exemple, les conditions de températures élevées obligent les animaux à développer le pâturage de nuit plutôt qu'aux moments les plus chauds de la journée. On peut ainsi observer de grandes dispersions au coucher du soleil qu'en journée.
- Les facteurs propres à l'animal : la physiologie Manteca i Vilanova and J. Smith (2014), la morphologie Shipley (1999), la mémoire spatiale (Laca, 2009), et les relations sociales (Barroso et al., 2000) ; peuvent aussi être déterminants pour caractériser le

comportement spatial des animaux.

- Enfin, comme facteurs correspondant à l'effet de second ordre, les points d'attraction peuvent également modifier le mode d'utilisation de l'espace par les animaux (Poilecot et al., 2013). L'utilisation des zones pâturées à proximité des points d'intérêt sera beaucoup plus dense que dans une zone sans aucune attraction particulière.

Une variété de modalités d'utilisation des parcours par les animaux peut donc être observée en fonction des milieux, des espèces animales et la façon dont ils sont élevés. L'analyse des motifs de points permet de caractériser ces différences et de fournir une compréhension de l'impact des facteurs agissant. Plusieurs mesures ont été définies dans des études antérieures pour les décrire. Nous pouvons les regrouper en quatre catégories : les mesures de centrographie, d'étendue, de densité et la série d'analyses basées sur la mesure de distance. Une brève description des différentes catégories de ces mesures peut être trouvée dans (O'Sullivan D. and Unwin, 2010; Spielman, 2017).

Mesures de centrographie

Les mesures de centrographie concernent les analyses de centralité d'un motif de points. Elles déterminent la position d'un point représentatif de l'ensemble. Elles permettent ainsi de résumer la localisation des phénomènes spatiaux ponctuels, et la régularité de leur distribution.

Deux principaux indicateurs peuvent être calculés : la moyenne centrale et la médiane centrale. La moyenne centrale représente la moyennes en X et en Y des coordonnées géographiques. Et la médiane centrale représente la point représenté par les valeurs médianes en X et Y. En d'autres termes, elle représente le point où la moitié des données est au-dessus ou au-dessous du point et l'autre moitié est à sa gauche ou à sa droite.

La distance standard, qui représente la variance entre la distance moyenne des points et le centre moyen, permet de représenter la dispersion et l'orientation des points dans l'espace. Deux types d'ellipse peuvent être tracés pour représenter le motif de points : basée sur la distance moyenne en X et Y (produit une seule mesure considérée comme rayon) et basée sur la distance moyenne en X d'une part et en Y d'autre part (produit deux distances en X et Y utilisées comme paramètres).

L'utilisation de ces indicateurs dans le cadre de l'élevage pastoral permettra de représenter de manière générale, les endroits fréquentés par les troupeaux et une approximation de leur dispersion en fonction des périodes précises. Ces périodes peuvent être des moments particuliers de la journée, les jours ou encore plusieurs jours en fonction des besoins des éleveurs.

Les mesures de centrographie sont plus générales et ne présentent pas plus de détails concernant la distribution des points. Pour extraire plus de connaissances sur l'utilisation des milieux pastoraux par les animaux, des analyses plus poussées doivent être faites. Nous les

présentons dans les points qui suivent.

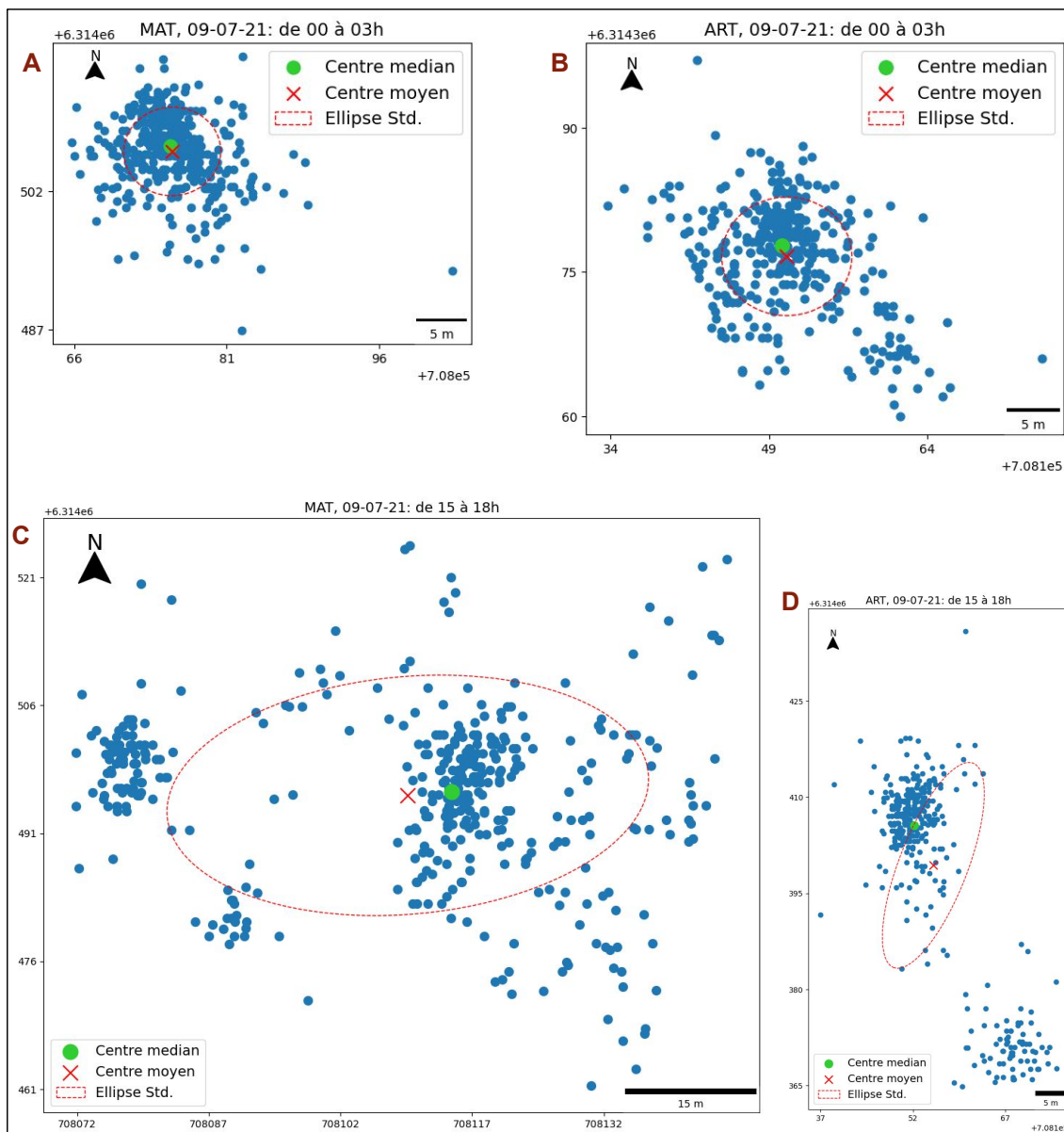


Figure 5.1 – Exemple d'indicateurs de centrographie : moyenne et médiane centrales du motif de points (en bleu) représentant l'utilisation des parcs par les deux groupes d'agnelles allaitées par les mères (MAT) et allaitées artificiellement (ART) dans le site expérimental de la Fage. L'ellipse standard représente la tendance du motif de points de localisation des animaux. Les graphes A et B représentent l'utilisation de deux différents parcs par les deux groupes d'agnelles en été, le 09-07-2021 entre 00h et 03h. La moyenne et médiane centrales y sont confondues, les deux motifs étant à peu près groupés. Les graphes C et D représentent l'utilisation des mêmes groupes d'agnelles, le même jour entre 15h et 18h. La moyenne et médiane centrales y sont bien séparées. Dans le graphe C, les agnelles MAT se dispersent plus dans la direction Ouest-Est et occupent une plus grande étendue que les agnelles ART qui se dispersent plus dans la direction Sud-Nord

Mesures d'étendue

Les mesures d'étendue permettent de représenter différentes formes de surfaces (rectangle minimum, cercle minimum, enveloppe convexe) occupées par un ensemble de points. Le but étant de déterminer la surface approximative qu'occupe un ensemble de points. Pour notre étude, l'intérêt serait de générer les étendues représentant l'occupation de l'espace par les animaux pendant les périodes définies. Cela permet de caractériser leur manière d'explorer l'espace.

Les différences d'utilisation peuvent être évaluées à travers les valeurs d'étendues explorées à des périodes définies du jour. Les zones les plus petites détermineront un regroupement qui peut être associé au repos ou à un regroupement des animaux autour d'un point d'attraction, tandis que les zones relativement grandes détermineront une dispersion et peuvent révéler les moments où les animaux paissent (Lécrivain et al., 1993). Dans le cadre de notre étude, nous avons appliqué ces mesures à un ensemble de points GPS caractérisant plusieurs animaux équipés, sur une période de temps donnée.

La Figure 5.2 présente les différentes formes d'indicateur d'étendue. Il apparaît que la forme qui se rapproche le plus de la superficie réelle explorée par les animaux est l'enveloppe convexe ou Alpha shape (Edelsbrunner et al., 1983; Akkiraju et al., 1995; Capinha and Pateiro-López, 2014). Elle peut permettre de déterminer la forme compacte qui représente occupée par les animaux. Le rectangle minimum et le cercle minimum peuvent être également utilisés pour représenter globalement l'espace occupé par le motif de points représentant la distribution des animaux.

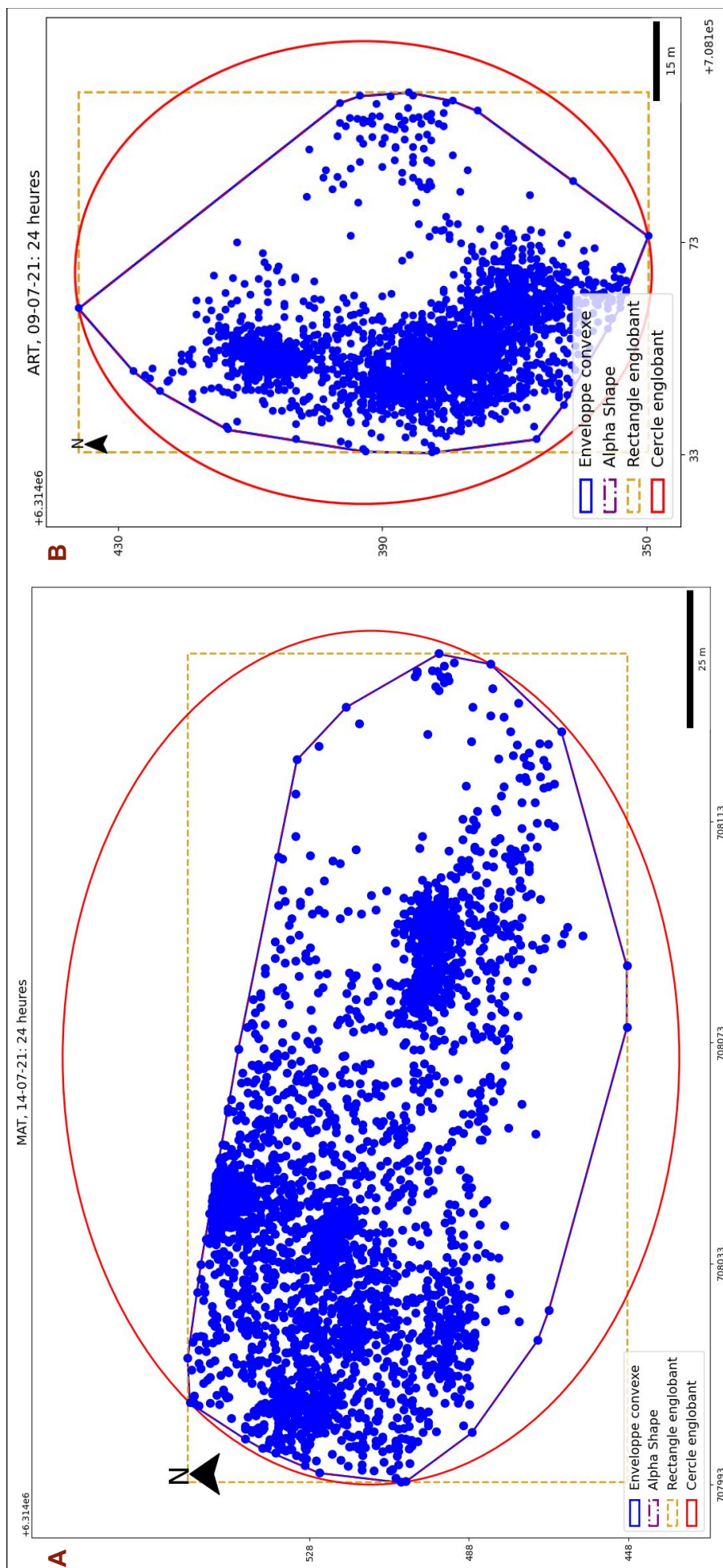


Figure 5.2 – Exemple d'indicateurs d'étendue. Ici, l'enveloppe convexe est confondue à l'Alpha shape qui est une de ses variantes.

Mesure de densité

Les mesures de densité permettent de caractériser l'intensité de la présence des animaux dans des zones explorées. Parmi les mesures de densité, on peut citer la densité globale, les statistiques par quadrats, et par noyau (López De La Cruz and Gutiérrez, 2008). Ces mesures devraient permettre de visualiser la distribution des points, pour tester ensuite la corrélation entre cette distribution et les facteurs du milieu (ressources pastorales, topographie, météo) qui représentent la propriété de premier ordre.

La densité globale consiste en une mesure globale de la densité par rapport à la zone explorée par les animaux. Il s'agit du rapport entre le nombre de points et la superficie explorée.

$$\hat{\lambda} = \frac{N}{A_t} \quad (5.1)$$

Avec $\hat{\lambda}$ la densité globale, N le nombre de points et A_t l'aire de l'étendue représenté par le motif de points à un instant t . La densité globale calculée ici se diffère du chargement qui correspond au nombre d'animaux par la superficie du parc. Elle ne tient pas compte du nombre total d'animaux mais plutôt du temps passé par les animaux équipés par rapport à une superficie qu'ils occupent à un instant t donné. Dans le cas où la superficie correspond à celle du parc, la densité globale serait corrélée au chargement.

La densité par quadrats consiste en une mesure locale de la densité, qui prend tout son intérêt lorsque la répartition des individus représentée par les points GPS, n'est pas constante dans l'espace. Pour l'élevage par exemple, les animaux utiliseront plus des endroits où la végétation fourragère est disponible en quantité et de meilleure qualité que ceux où elle est moins appréciée (Bailey and Provenza, 2008). La technique du Quadrat divise une zone d'étude en des cellules de taille déterminée appelées quadrats, qui représentent des sous régions de la zone. Un comptage de points est ainsi fait au sein de chaque quadrat pour mesurer la densité locale de l'ensemble de la zone explorée. La valeur de la taille de quadrats peut influencer la mesure de la densité locale. Les plus petites valeurs conduisent souvent à l'obtention de beaucoup de quadrats dépourvus de points. A l'inverse, de grandes valeurs peuvent empêcher la découverte de variations dans la distribution spatiale des densité.

Le choix de la taille des cellule doit être adapté au contexte d'élevage propre à chaque exploitation pour espérer d'obtenir plus d'information sur la distribution des animaux dans le milieu. Trois facteurs sont à considérer : le niveau d'hétérogénéité du milieu et la taille du troupeau (en considérant également qu'on n'équipe jamais tous les animaux). Plus le milieu est finement hétérogène, plus il sera pertinent de choisir des mailles de petite taille ; mais plus le troupeau est grand, plus cette précision sera illusoire car de fait des animaux non équipés, en général proches des animaux équipés, pourraient avoir occupé les quadrats voisins. En plus, il convient de prendre en compte la précision des mesures GPS qui, en conditions pastorales

et avec le matériel utilisé pour l'étude, est aux environs de 10m. A travers ces considérations, on choisira comme plus petite maille de 10 x 10m s'agissant de petits parcs ou des conditions expérimentales avec de petits groupes d'animaux, où de nombreux individus sont équipés ; et des mailles de plus grande taille (par ex. 25 x 25m ou plus) dans le cas de troupeaux de plus grande taille, et/ou lorsque peu d'animaux sont équipés.

En fin, **la densité par noyau** génère également une représentation de la densité localisée pour des portions de la zones explorée. Contrairement à la densité par quadrats, une grille de taille déterminée parcourt les cellules de la zone d'intérêt et calcule la densité par la méthode du noyau (KDE = Kernel Density Estimation). Cette méthode permet de représenter la continuité de la densité dans un milieu exploré par les animaux.

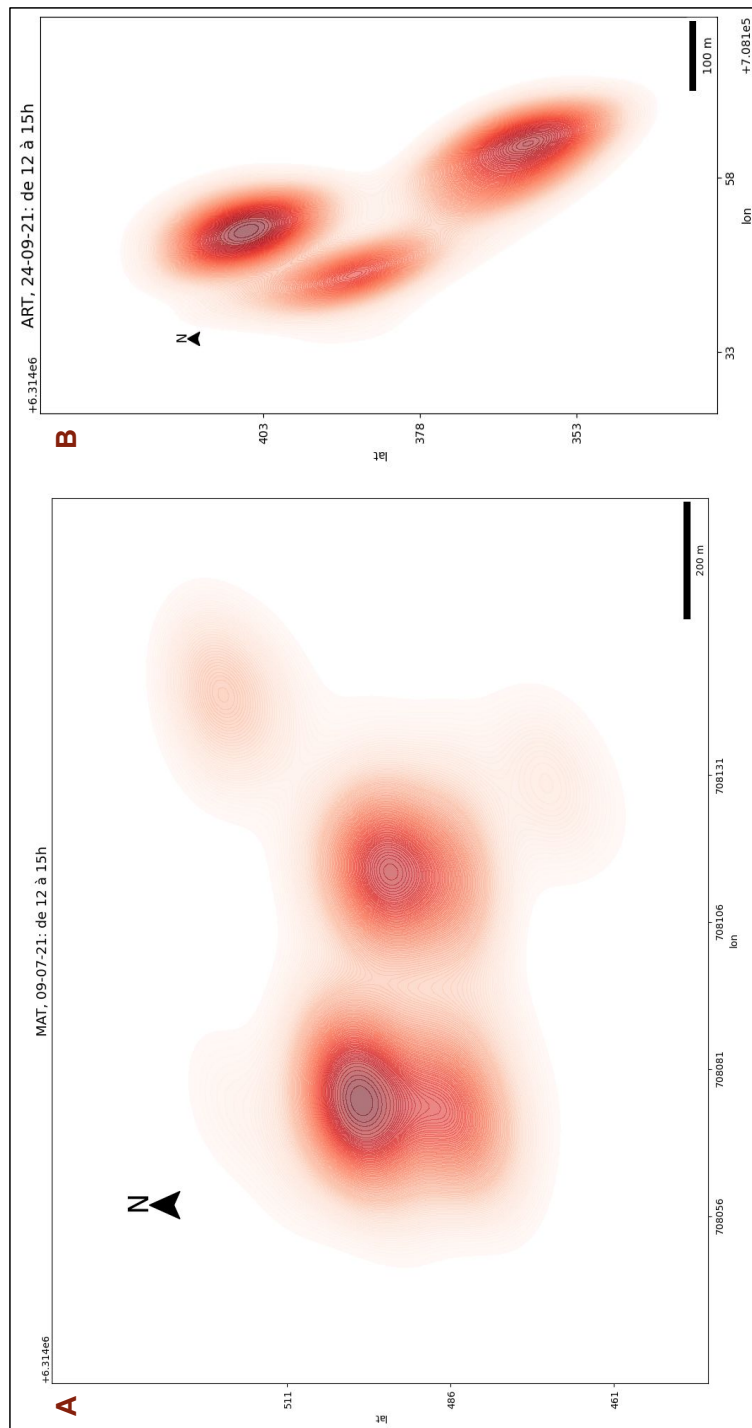


Figure 5.3 – Exemple d'indicateurs de densité. A gauche : densités locales calculées au sein de chaque cellule de 10m X 10m. A droite : Densité locale par noyau.

Parmi ces différents indicateurs de densité, la densité globale et par quadrat permettent d'extraire des valeurs numériques pouvant être utilisées pour des analyses à posteriori. La densité par noyau est plus utilisée pour le but de représentation spatial.

Indicateurs basés sur le calcul de distances

Ces indicateurs permettent de mesurer le niveau de dispersion ou de regroupement des motifs de points. Ils permettent de mesurer la façon dont les points sont distribués les uns par rapport aux autres et concernent plus la propriété de second degré, contrairement aux indicateurs de densité. La dispersion peut être calculée à partir des fonctions de Ripley (Ripley, 1988), et des mesures de distance aux plus proches voisins (Dixon, 2001).

Les résultats produisent des indices de dispersion qui peuvent être comparés en fonction des périodes de pâturage pour caractériser le comportement des animaux. La dispersion des animaux d'un troupeau peut varier à la fois au cours de la journée et au fil de l'utilisation des milieux pastoraux. Elle peut être mesurée au cours de la journée et au cours des jours et des phases de pâturage afin caractériser l'évolution de l'utilisation des parcours par les troupeaux. Au cours des journées, les animaux utilisent différemment l'espace. Une succession des comportements pâturage-repos-rumination-pâturage représente brièvement l'activité des animaux sur parcours (Lécrivain et al., 1993). Dans des conditions pastorales contrastées, un troupeau peut être amené à se disperser plus ou moins, notamment en fonction de la topographie (Bailey et al., 2015) et surtout de la densité de ressources pastorales (Bailey and Provenza, 2008; Chebli et al., 2020). Au fil de l'utilisation du parc, en terme de jours et de phase de pâturage, les troupeaux peuvent varier l'occupation de l'espace pour rechercher des zones avec de la ressource disponible pour se nourrir (Chebli et al., 2020). Par ailleurs, des troupeaux différents peuvent avoir des habitudes de pâturage différentes; notamment, les animaux gardés "serrés" (comme c'est le cas dans les Alpes) aiment bien pâturer près les uns des autres, alors que ceux qui évoluent tout le temps dans de grands parcs sont souvent plus autonomes les uns des autres et aptes à se disperser. En fonction de chaque contexte de pâturage, le niveau de dispersion des motifs de points peuvent permettre de caractériser les moments de pâturage et ceux de faibles activités (repos, rumination, sommeil) ou le regroupement autour d'un point d'intérêt.

Les fonctions de Ripley sont utilisées pour représenter graphiquement la dispersion ou le regroupement des motifs de points. Pour des analyses basées sur des données enregistrées sur plusieurs jours, nous avons besoin d'un indicateur numérique comparable afin de caractériser la dispersion des animaux au fil de l'utilisation des ressources du parcours. Les mesures de distance aux plus proches voisins conviennent le mieux à cet objectif. Dans cette étude nous nous pencherons plus sur ce type de mesure.

L'analyse de la moyenne des plus proches voisins (AND) mesure la distance moyenne entre chaque point de la zone d'étude et le point le plus proche. La mesure du regroupement ou de la dispersion des points est alors effectuée en comparant la moyenne des distances au plus proches voisins d'un motif de points observée à celle d'un motif de points issu d'une distribution aléatoire (CSR) (Clark and Evans, 1954) (cf. Équations 5.3, 5.2 et 5.4).

$$AND_{obs} = \frac{\sum_{i=1}^n d_{i_{obs}}}{N} \quad (5.2)$$

$$AND_{CSR} = \frac{\sum_{i=1}^n d_{i_{CSR}}}{N} \quad (5.3)$$

$$Ratio = \frac{AND_{obs}}{AND_{CSR}} \quad (5.4)$$

Avec : AND_{obs} la moyenne des distances aux plus proches voisins issus de la distribution observée, AND_{CSR} celle issue d'une distribution issu d'un processus aléatoire, d_i la distance au point le plus proche, N le nombre de points, et A l'air du rectangle minimum englobant du motif de points.

Si $Ratio = 1$, la distribution des points est considérée comme étant aléatoire, Si $Ratio < 1$, la distribution des points est considérée comme étant groupée, and Si $Ratio > 1$, la distribution des points est considérée comme étant dispersée.

Pour des motifs des points représentant la localisation des animaux, il est possible de trouver uniquement des motifs groupés ou dispersés. Dans ce cas, la caractérisation de l'activité des animaux est faite à partir de l'analyse des variations des valeurs produites.

Table 5.2 – Récapitulatif des indicateurs permettant de représenter et de caractériser l'utilisation des parcours par les animaux. Les indicateurs en gras sont ceux qui ont été retenus dans la suite du travail. Ils permettent de répondre aux deux premières questions faisant l'objet de ce chapitre.

Indicateur	Description	Applications
Indicateurs de tendance centrale		
Moyenne et médiane centrale	Moyenne ou médiane en X et Y des points	Représentation globale de la localisation des animaux sur parcours
Ellipse standard	Ellipse représentant l'orientation de la dispersion des points	Représentation du degré de dispersion des animaux par rapport à la moyenne ou la médiane lors de l'utilisation des parcours
Indicateurs d'étendue		
Rectangle minimum englobant	Rectangle minimum incluant l'ensemble des points	Représentation de l'étendue occupée par les animaux durant une période de pâturage
Cercle minimum englobant	Cercle minimum représentant l'ensemble des points	Représentation de l'étendue occupée par les animaux durant une période de pâturage
Enveloppe convexe ou Alpha shape	Etendue compacte représentant l'étendue occupée par les points	Représentation de l'étendue approximative explorée par les animaux durant une période de pâturage
Indicateur de densité		
Quadrat	Mesure du nombre de points au sein de petites zones (quadrats) pour représenter la densité	Représentation de la variation de la densité au niveau de petites zones de l'étendue explorée.
Densité par noyau	Mesure de la densité par une fonction permettant de lisser les valeurs	Représentation de la variation de densité de manière continue dans l'espace pâturé par les animaux
Indicateurs de distance		
Moyenne des plus proche voisins (AND)	Mesure le niveau de dispersion ou de regroupement des points.	Caractérisation du niveau de dispersion des animaux durant les périodes de pâturage

5.3 Croisement des informations spatiales pour modéliser et caractériser l'utilisation des parcours par les animaux

Les différents indicateurs d'utilisation de l'espace par les animaux, sont le reflet des effets de premier (facteurs environnementaux) et de second (interaction avec d'autres objets spatiaux) ordre. Ces effets peuvent être évalués en intégrant les facteurs sélectionnés. L'objectif étant de caractériser et de hiérarchiser les facteurs qui déterminent le mode d'utilisation de l'espace par les animaux. L'observation de différences dans l'utilisation des parcours à l'échelle de la journée, des jours et des phases de pâturage nécessite un découpage des données en périodes bien définies pouvant permettre de caractériser le comportement des animaux. Ainsi, plusieurs découpages des données peuvent être réalisés : en créneaux journaliers, au jour et à la phase de pâturage (plusieurs jours successives d'une saison).

Parmi les données identifiées, certaines ne présentent que l'information spatiale comme c'est le cas pour l'occupation des sols et la pente du terrain. D'autres présentent des informations temporelles telles que les données météorologiques horaires ou journaliers, et l'état de la végétation (basé sur le calcul d'indice) qui possèdent une temporalité qui dépend des caractéristiques de données satellites utilisées pour les produire (1 à 10 jours). Les analyses permettront d'explorer ces informations spatiales et temporelles pour viser l'extraction des connaissances au fil des journées et des phases de pâturage. Le Tableau suivant résume les différentes possibilités de découpage des données pour constituer les bases de données d'analyse.

Table 5.3 – Fréquences d'analyse des information pour caractériser l'utilisation des parcours par les animaux

Fréquence d'analyse	Facteurs	Application
Créneaux journaliers	Indicateurs d'étendue Indicateurs de dispersion Météorologie horaire Occupation des sols Pente Points d'intérêt	Caractérisation du comportement des animaux en cours des journées en fonction de l'influence des facteurs
Jours de pâturage	Indicateurs de densité Distance parcourue Météorologie journalière Occupation des sols Pente Points d'intérêt	Caractérisation de l'intensité d'utilisation des zones pastorales et de la distance parcourue par les animaux sur l'action des facteurs indiqués
Phase de X jours de pâturage	Indicateurs de densité Distance parcourue Météorologie journalière Occupation des sols Pente Points d'intérêt	Caractérisation de l'intensité d'utilisation des zones pastorales et de la distance parcourue par les animaux sur l'action des facteurs indiqués

La caractérisation de l'influence des facteurs du milieu sur les indicateurs de l'utilisation des parcours peut en suite être faite à partir d'analyse de corrélation et la régression linéaire multiple. L'interaction entre variables pourra ensuite être évaluée en utilisant le coefficient de détermination (R^2). Un test de significativité peut également permettre d'estimer la probabilité de liaison des facteurs en utilisant la p valeur de référence de 0,1.

Des analyses spatiales (agrégation, intersection) spatiales pourront aussi être réalisées en vue d'associer les comportements des animaux à des caractéristiques des variables liées au parcours. Il s'agira par exemple de distinguer les distances parcourues en fonction des types d'occupation de sol ou de catégorie de pente. Ou encore, de représenter l'intensité d'utilisation des zones en fonction du type l'occupation des sols ou de pente.

Dans la section qui suit nous présentons les résultats de l'application des méthodes décrites pour caractériser l'utilisation de l'espace par deux types d'animaux : les agnelles élevées sous la mère (naturellement allaitées) et les agnelles élevées sans la mère (artificiellement

allaitées).

5.4 Application en conditions expérimentales

5.4.1 Contexte

Cette étude concerne une expérience menée dans l'unité expérimentale INRAE de La Fage, située sur le plateau calcaire du sud du Massif Central français, à une altitude d'environ 800 m (cf. Figure 5.4). Elle est caractérisée par un climat méditerranéen avec un été tempéré selon la classification de Köppen, avec une température maximale moyenne de 16°C sur l'année (de 6°C en janvier à 27°C en août), et une pluviométrie annuelle de 966mm (minimum de 45mm en août et maximum de 108mm en avril). Dans cet environnement, les sols sont très peu profonds et la croissance de l'herbe est très saisonnière, principalement entre mai et juin. Les landes à buis et à genévrier constituent les principales formations végétales des parcours utilisées pour l'élevage extensif des ovins.

L'essai a été mené sur de petits parcs d'environ 1ha à la topographie relativement plane, avec une végétation composée d'un tapis herbacé diversifié et de massifs ligneux parsemés, une mangeoire où les animaux recevaient de petites quantités d'aliments concentrés, et un réservoir d'eau. Deux groupes d'agnelles de 3,5 mois issus d'un même troupeau ont été utilisés pour caractériser et analyser les variations dans leur utilisation des parcours. Ce troupeau dont les mères ont reproduit ces agnelles a été conduit en plein air intégral, en majorité sur parcours, sans complémentation entre mai et décembre. Un groupe d'une vingtaine d'agnelles ayant reçu un allaitement maternel (MAT) et un autre d'une vingtaine d'agnelles un allaitement artificiel (ART). Les deux groupes ont été placés dans des sous parcs différents d'environ 1ha durant les périodes estivale et automnale pendant respectivement 35 et 26 jours. Dans chaque sous parc, une mangeoire a été placée afin de compléter les agnelles chaque matin avec environ 5 kg de concentré.

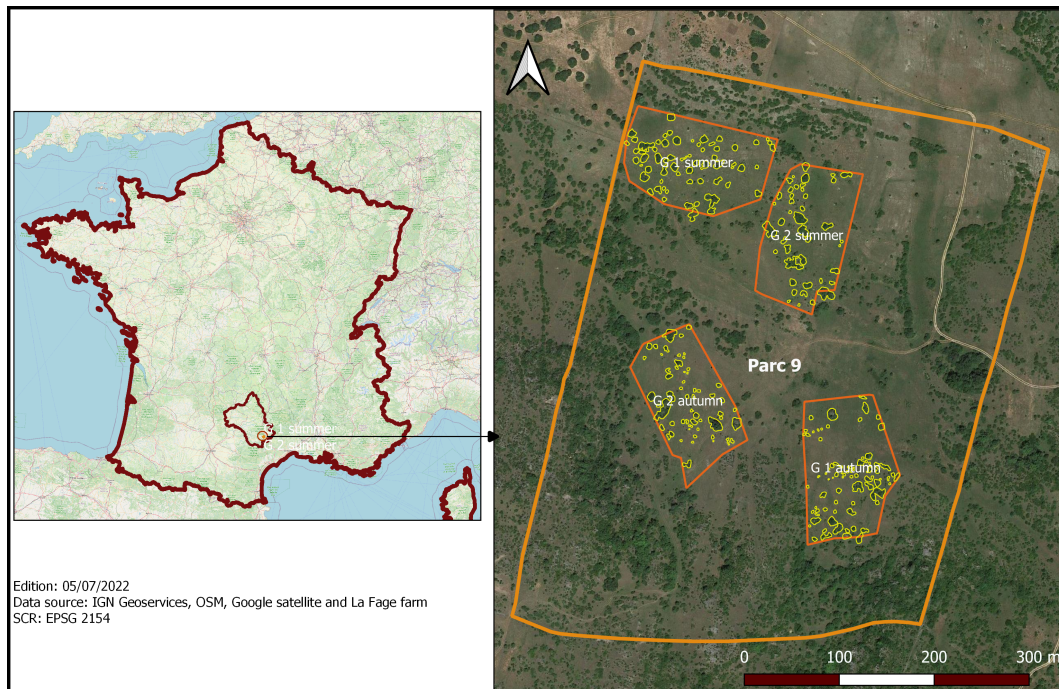


Figure 5.4 – Localisation de l'unité expérimentale de la Fage dans le département de l'Aveyron. Le grand parc 9 comprend les différents sous parcs pâturés séparément par les animaux en été et en automne. Les contours jaunes représentent les périmètres des massifs de plantes ligneuses.

Des colliers GPS E-Pasto fabriqués par la société Aguila technologies ont été placés sur 12 des 20 agnelles du troupeau MAT, et 11 des 20 agnelles du troupeau ART afin d'enregistrer leur localisation en continu pendant les périodes estivale et automnale avec une fréquence de 5 minutes. Une grande partie des agnelles (respectivement 60% et 55%) des deux groupes ont donc été équipées de GPS en vue d'analyser le comportement du troupeau dans son ensemble.

5.4.2 Données utilisées

Les données utilisées dans cette étude sont les points GPS enregistrés pour les agnelles des deux groupes, les NDVIs calculés pour la période de pâturage estivale, la pente des parcs, les paramètres météorologiques horaires et journaliers, et les données spatiales des limites des parcs et des contours des massifs de ligneux (cf. Tableau 5.5).

Les données GPS ont été collectées du 08/07/21 au 12/08/21 (35 jours enregistrés) pour la période estivale et du 23/09/21 au 19/10/21 (26 jours enregistrés) pour la période automnale. La fréquence du GPS Aguila a été réglée pour enregistrer les données toutes les 5 minutes. Dans les meilleures conditions, environ 288 points devraient être collectés sur 24 heures. Cependant, cela n'a pas été le cas pour toutes les périodes étudiées en raison des conditions particulières liées aux matériels GPS et à l'environnement. Nous avons défini un

seuil de 220 points GPS par jour comme critère de données à retenir. Pour ne conserver que les jours ayant complété le cycle de 24 heures, le premier et le dernier jour de présence dans le parc ont été exclus des analyses.

Le prétraitement des données GPS s'est fait en deux étapes. La première étape a concerné l'extraction spatiale des points intersectant chacun des parcs pâturés par les agnelles en été et en automne. Puis une seconde étape a consisté à éliminer les données des jours dont la moyenne est inférieure à 220 points. Le résultat de ce processus est présenté dans le Tableau 5.4).

Table 5.4 – Aperçu des données GPS utilisées dans les analyses. Le total des points représente la somme des données GPS avec plus de 220 points de localisation par jour.

Saison	Période	Groupe	Jours <220 points	Jours retenus	Total points
Eté	08/07 - 12/08/2021	Agnelles MAT	16/07/2021	31/35	99732
		Agnelles ART	03/08/2021	31/35	89154
Automne	23/09 - 19/10/2021	Agnelles MAT	13 et 15/10/2021	22/26	64313
		Agnelles ART	-	22/26	44902

Afin d'analyser la variation du comportement quotidien en termes d'utilisation de l'espace par les animaux, nous avons segmenté les données en huit périodes de 3 heures (0-3h, 3-6h, 6-9h, 9-12h, 12-15h, 15-18h, 18-21h et 21-24h). Le choix de cette segmentation tient compte du compromis entre le nombre de points pour chaque créneau (suffisamment élevé) et les moments de la journée permettant l'observation du comportement des animaux (repos et pâturage). En fin, pour analyser l'utilisation des parcs au fil du temps, les données GPS ont été découpées en phases de 10 jours. Cependant, en raison de l'élimination des jours avec moins de 220 points, nous avons dû aller au-delà de 10 jours (11 jours) pour certaines phases afin de construire les données pour 10 dates. Pour l'été, les données ont été segmentées en 3 phases : du 09 au 19/07/2021 (11 jours, sans le 16/07/2021), du 20 au 29/07/2021 (10 jours), et du 30/07 au 09/08/2021 (11 jours, sans le 03/08/2021). Et pour l'automne, les données ont été segmentées en 2 phases : du 24/09 au 03/10/2021 (10 jours) et du 04 au 14/10/2021 (11 jours, sans le 13/10/2021).

Les NDVI utilisés sont dérivés des données Planet rééchantillonnées à une résolution spatiale de 10m qui correspond à la fine résolution spatiale utilisable par rapport à l'utilisation de petits parcs par des lots avec plusieurs animaux équipés en GPS. Le choix des données Planet a été fait pour compenser l'indisponibilité des données Sentinel 2 sur les périodes d'intérêt pour notre étude. Le quota de données Planet n'était intéressant que pour la période estivale pour laquelle nous avons obtenu des images pour 18 dates. Pour la période d'automne, seules 5 dates ont été collectées, ne permettant pas de réaliser les analyses prévues pour la majeure partie de la période de pâturage. La présence de nuages en automne n'a pas favorisé la prise de données Planet de meilleure qualité. Le NDVI a été généré automatiquement

avec le module python eo-crops² utilisant la bibliothèque eo-learn³ et OrfeoToolBox⁴ pour télécharger et analyser les données à l'échelle parcellaire.

Le MNT téléchargé pour le département de l'Aveyron a été utilisé pour produire une carte de pente de 10m de résolution spatiale à partir du logiciel SAGA GIS⁵. Les données de pente pour chaque parc ont été obtenues par extraction de la couche vectorielle correspondante. Les intervalles de pente représentant les parcs pâturés sont les suivantes : $[0^\circ - 11.75^\circ]$; $[0.57^\circ - 11.75^\circ]$; $[1.15^\circ - 12.78^\circ]$ et $[0^\circ - 17.70^\circ]$ respectivement pour les parcs utilisés par les groupes 1 et 2 en été et en automne.

Les données météorologiques ont été enregistrées par la station INRAE CLIMATIK implantée dans l'unité expérimentale de la Fage. Les enregistrements contiennent des informations horaires et journalières sur la température, la vitesse du vent, l'humidité de l'air et le rayonnement global. Les données météorologiques ont été extraites pour les dates correspondant aux périodes de pâturage d'été et d'automne. Toutes les variables ont été retenues pour être incluses dans l'analyse afin de comprendre l'effet de la météo sur le comportement des animaux au pâturage.

Les massifs ligneux présents dans les pâturages influencent le comportement des animaux. Ils créent les conditions pour le maintien d'une végétation herbacée de meilleure qualité et offrent également un fourrage complémentaire provenant d'espèces ligneuses. De plus, ils servent d'abri aux animaux en cas d'intempéries, de vents violents ou de températures élevées (Roudaut et al., 2007). Dans le cas du site expérimental de La Fage, ils sont composés de buis et de genévriers, qui sont peu consommés par les animaux. Pour évaluer leur influence sur l'utilisation de l'espace par les deux groupes d'animaux, nous avons digitalisé des polygones de chaque massif ligneux en les numérisant sur une image Google Earth de 2019. La comparaison avec des images capturées bien avant cette date montre que l'emplacement de ces espèces ligneuses reste semblable. En plus la croissance des buis (*Buxus sempervirens*) est très lente surtout en milieu pauvre et aucune opération de débroussaillage n'a été réalisée dans le parc concerné par l'étude. Nous supposons que cette distribution d'arbustes n'a pas changé pour la période qui concerne nos analyses.

2. <https://github.com/j-desloires/eo-crops>

3. <https://eo-learn.readthedocs.io/en/latest/>

4. <https://www.orfeo-toolbox.org/CookBook/index.html>

5. <https://saga-gis.sourceforge.io/en/index.html>

Table 5.5 – Données utilisées et applications relatives aux facteurs influençant l'utilisation de l'espace par les animaux. Ces données peuvent également être combinées pour analyser leur interaction sur l'utilisation de l'espace par les animaux.

Données	Granularité spatiale	Granularité temporelle	Application
Localisation GPS	-	5 minutes	Caractérisation du comportement des animaux et de leur mode d'utilisation des parcours
NDVI	10m	1 à 5 jours	Caractérisation de l'influence de la végétation sur le comportement des animaux.
Massifs de ligneux	-	-	Caractérisation de l'influence des ligneux sur le comportement des animaux.
Pente	10m	-	Caractérisation de l'effet de la pente sur le comportement des animaux.
Météorologie	-	Horaire et Journalier	Caractérisation de l'influence des paramètres météorologiques sur le comportement des animaux

5.4.3 Production d'indicateurs du comportement des agnelles

Pour les besoins de cette étude, nous avons choisi la moyenne centrale comme mesure de tendance centrale pour représenter la position du groupe d'agnelles. Comme mesure d'étendue, nous avons choisi l'Alpha shape en vue de représenter la zone explorée par les agnelles. Comme indicateur de densité nous avons opté pour les statistiques par quadrats afin de mieux représenter la présence des agnelles dans les sous parcs et de produire des mesures pouvant être intégrées dans des analyses ultérieures. La maille de 10m a été jugée idéale en considérant la taille des lots (20) et des nombres d'agnelles porteuses de GPS (12 et 11). Elle permet également de garder la correspondance avec les autres données spatiales. Les densités par noyau n'ont pas été considérées étant donné qu'elles sont plus utilisées à des fins de visualisation. Enfin, nous avons utilisé la moyenne des distances aux plus proches voisins (AND) comme mesure de dispersion car elle représente une valeur globale qui peut être utilisée pour caractériser la dispersion des agnelles, contrairement aux fonctions de Ripley qui sont plus destinées à un usage graphique. Nous avons utilisé la distribution de Poisson pour générer le motif des points correspondant à la distribution aléatoire pour comparer le niveau de dispersion ou de regroupement des agnelles.

Pour chaque période journalière de 3 heures, une mesure de dispersion et d'étendue ont ainsi été générées pour effectuer les analyses statistiques. La moyenne centrale a été utilisée uniquement pour générer les distances de chaque groupe d'agnelles à la mangeoire et à l'abreuvoir mais pas comme indicateur de l'utilisation des sous parcs. Nous avons utilisé la bibliothèque libpysal (Rey and Anselin, 2007) pour générer la surface explorée et pointpat (Rey et al., 2019) pour générer les valeurs de dispersion et de moyenne centrale sous python 3.8.

5.4.4 Analyse des informations

Le but des analyses a été de caractériser l'utilisation des parcs par les agnelles, quotidiennement et au fil des jours, à partir des indicateurs sélectionnés; et d'expliquer l'influence des variables au sein des parcs sur le comportement des agnelles.

Caractérisation de l'utilisation des parcs par les agnelles

Nous avons appliqué l'analyse de la variance (ANOVA) à un facteur pour effectuer 3 niveaux de comparaison de l'utilisation de l'espace par les deux groupes d'agnelles. Premièrement, pour les différentes phases de pâturage représentées (début, milieu et fin), nous avons effectué une comparaison deux à deux des mêmes créneaux horaires (avec H_0 : l'équivalence de la distribution des valeurs d'indicateurs pour les mêmes créneaux horaires entre deux phases) afin d'observer d'éventuelles variations au sein d'un même groupe. Ensuite, pour chaque phase d'été et d'automne, une comparaison deux à deux a été faite entre les créneaux horaires (avec H_0 : l'équivalence de la distribution des valeurs pour chaque paire de créneaux horaires de la journée), afin de différencier l'utilisation journalière de l'espace par les agnelles. Enfin, une comparaison deux à deux entre les mêmes créneaux horaires des deux groupes d'agnelles pour chaque phase de pâturage (avec H_0 : l'équivalence de la distribution des valeurs pour les mêmes créneaux horaires des deux groupes à chaque phase de pâturage), afin d'observer les différences dans le mode d'utilisation des parcs entre les deux groupes d'agnelles. Nous avons utilisé le niveau de significativité de 95% ($p < 0.05$).

Caractérisation de l'intensité d'utilisation des parcs par les agnelles

L'activité des animaux a également été représentée sous forme de carte de densité à partir des quadrats. L'emplacement des points pour chaque phase de 10 jours a été ainsi produit à partir d'une grille représentant le nombre de points dans chaque cellule de grille de 10m x 10m. Cette grille vectorielle a ensuite été transformée en image matricielle pour représenter la distribution de la densité de points dans chaque pixel de 10m.

Les valeurs de densité étant très variables pour les phases observées, une reclassification a été effectuée pour simplifier l'interprétation et rendre les résultats comparables. Les valeurs de densité maximale et minimale se situaient entre 1 et 4877 points par maille de 10m X 10m pour toutes les phases de 10 jours considérées. Les mailles dont les valeurs de densité étaient égales à 1 ont été considérées comme ayant une densité nulle car étant très faible pour 10 jours de pâturage et pouvant également provenir d'une imprécision de mesure de GPS. Pour effectuer le classement, nous avons généré des déciles sur les valeurs de densité observées. Le Tableau 5.6 résume les plages de valeurs correspondant aux classes générées. On remarque que l'amplitude des valeurs de densité augmente en fonction des classes. Cela montre que les plus grandes densités de présence par maille de 10m x 10m (cas de la classe

10), sont moins fréquentes.

Les massifs de ligneux devraient plus être utilisés par les animaux en cas de fortes températures, de fort vent et de pluies (Manteca i Vilanova and J. Smith, 2014). Les données météorologiques enregistrées à La Fage montrent que ces paramètres sont plus élevés durant l'été que l'automne, avec les valeurs minimales, moyennes et maximales respectives de : 8,8°, 17,6°, 30,4° et 1,8°, 11,63°, 22,3° de température; 00mm, 0,2mm, 15mm et 0,0mm, 0,0mm et 2,5mm de précipitation; et 1,0m/s, 4,5m/s, 12,0m/s et 0,0m/s, 4,4m/s et 12,0m/s de vitesse de vent. Par conséquent, le temps passé sous les ligneux devrait être plus important pendant la période estivale que pendant l'automne. Afin d'observer cela, nous avons utilisé l'intersection entre les contours des ligneux et les points représentant la localisation des agnelles pour les périodes de pâturage d'été et d'automne. Cela nous a permis d'extraire pour chaque polygone représentant un massif de ligneux, le nombre de total de points pour chacune des périodes de pâturage étudiées. Nous avons utilisé des histogrammes afin d'observer la fréquence des différentes classes de densité décrits dans le Tableau 5.6.

Table 5.6 – Intervalles de classes de valeurs de densité de présence au sein des zones des sous parcs par les agnelles. Découpage en classe réalisé à partir des déciles sur les valeurs de densités de présence des agnelles pour l'ensemble des phases d'été et d'automne.

Intervalle par phase	Classe	Interprétation
[2 - 12]	1	Présence rare
[12 - 31]	2	Présence rare
[31 - 52]	3	Présence occasionelle
[52 - 77]	4	Présence occasionelle
[77 - 103]	5	Faible utilisation
[103 - 133]	6	Faible utilisation
[133 - 179]	7	Utilisation modérée
[179 - 245]	8	Utilisation modérée
[245 - 393]	9	Utilisation fréquente
[393 - 4877]	10	Forte utilisation

Analyse de l'influence des facteurs du milieu

Les variables d'étendue explorée et de dispersion sont des indicateurs spatiaux étant donné qu'elles sont localisées dans les parcs et temporels étant donné qu'elles varient en fonction des périodes (créneaux) journalières. La variable de densité quant à elle est un indicateur spatial uniquement car elle représente pour une période donnée, la densité de présence des animaux dans les différentes zones des parcs.

Pour effectuer l'analyse de la corrélation entre les indicateurs produits et les variables environnementaux et météorologiques, nous avons constitué 2 types de jeux de données : spatiales et temporelles. Les jeux de données spatiales correspondent au regroupement au sein de mêmes tables, des indicateurs spatiaux (étendue, dispersion et densité) et des valeurs des variables environnementaux (pente, NDVIs, distance à la mangeoire et distance au massif de ligneux le plus proche). Les jeux de données temporelles correspondent quant à elle au

regroupement des indicateurs temporels (étendue et dispersion) et des valeurs des paramètres météorologiques. Pour ces analyses, les indicateurs d'étendue explorée et de dispersion ont été utilisés par créneaux journaliers de 3 heures en vue d'étudier le comportement journalier des agnelles. L'indicateur de densité a été utilisé à une fréquence journalière en vue d'analyser l'utilisation spatiale des agnelles.

La pente, la mangeoire et les contours de ligneux étant spatialement fixes, leurs données représentant chacune des parcs d'été et d'automne ont été intégrées de manière unique pour l'extraction des valeurs. Pour chaque indicateur spatial, les valeurs du NDVI de la date la plus proche ont été utilisées étant donné le manque de données pour certains jours (18 dates sur 32 pour l'été). L'intégration du NDVI n'a été faite que pour les données du pâturage d'été vu que pour l'automne très peu d'images de qualité ont été trouvées.

Pour analyser l'influence des facteurs du milieu sur le comportement des agnelles, nous avons réalisé dans un premier temps une régression linéaire entre chaque indicateur et chacune des variables environnementales ou météorologiques. Nous avons appliqué la méthode de moindres carrés de la fonction `linregress` implémentée dans la librairie `scipy`⁶ pour réaliser ces opérations. Nous présentons les différents graphes produits en Annexes (voir Figures A.3 à A.22).

Nous avons ensuite effectué une régression linéaire multiple sur chacune des données constituées pour le pâturage d'été et d'automne pour les deux groupes d'agnelles. Nous avons utilisé la méthode de moindres carrés ordinaire pour réaliser cette tâche. L'interaction de toutes les variables a été évaluée en utilisant le coefficient de détermination (R^2), et l'influence des facteurs individuellement a été évaluée en utilisant le seuil de significativité de 95%. Ainsi, le but a été d'éliminer itérativement les facteurs non liés aux indicateurs produits en vue de ne garder que ceux qui leur sont corrélés. Ces analyses ont été réalisées à l'aide du module `statmodels`⁷ sous Python 3.8.

Table 5.7 – Constitution de deux types de données pour l'analyse de l'impact des facteurs du milieu sur l'utilisation de l'espace par les agnelles : i) association entre indicateurs et variables spatiales et ii) association entre indicateurs et variables temporelles. Les paramètres météorologiques horaires ont été moyennés par période de 3 heures pour la correspondance avec les indicateurs d'étendue et de dispersion.

Croisement entre indicateurs et variables spatiaux		Croisement entre indicateurs et variables temporelles	
Indicateurs spatiaux	Variables spatiales du milieu	Indicateurs temporels	Variables temporelles
Étendue explorée - Par créneau de 3h	Pente NDVI Distance à la mangeoire Distance au massif de ligneux le plus proche	Étendue explorée - Par créneau de 3h	Humidité instantanée (U) Rayonnement global (RG) Température (°C) Vitesse du vent (VX) Précipitations (PP)
Dispersion - Par créneau de 3h		Dispersion - Par créneau de 3h	
Densité de présence - Journalière (24h)			

6. <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.linregress.html>

7. <https://www.statsmodels.org/stable/index.html>

5.4.5 Résultats

Évolution des indicateurs au fil de l'utilisation du parc

Étendue explorée et dispersion des agnelles au fil du temps. La distribution de la surface explorée et la dispersion par les deux groupes d'agnelles sont présentées dans les figures 5.5, 5.6, 5.7 et 5.8.

Durant l'été (cf. Figure 5.5), pour les agnelles MAT, les faibles valeurs d'étendues apparaissent aux créneaux de 00 à 06h durant la phase 1. Durant la phases 2 on observe en plus les créneaux de 12 à 18h et de 21-00h, et de manière similaire pour la phase 3 sauf au créneau de 15-18h qui présente des valeurs élevées. A ces périodes de la journée des 3 phases respectives, les agnelles MAT pourraient être au repos ou en rumination étant donné qu'elles se regroupent plus qu'à d'autres moments. Les grandes valeurs d'étendue apparaissent pour la phase 1 aux créneaux de 06 à 12h et de 15 à 21h, avec plus de variabilités. Pour les phases 2 et 3 en revanche, on observe une démarcation aux périodes 06-09h et 12-21h qui montrent de grandes valeurs d'étendue explorée avec moins de variabilité qu'à la phase 1. La distribution de ces valeurs d'étendue pourrait traduire l'activité de pâturage au sein du parc. On observe également d'autres distributions intermédiaires pouvant traduire des moments de pâturage modéré ou de faible activité des agnelles.

Les mêmes tendances d'utilisation journalière du parc sont observées pour les agnelles ART. On peut cependant observer quelques différences dans les phases 2 et 3 où la distribution des valeurs d'étendue explorée est mieux démarquée aux créneaux de 06h-09h et 18h-21h pour les agnelles MAT, alors qu'elle est plus ou moins confondue avec celle des créneaux intermédiaires pour les agnelles ART (cf. Figure 5.5). Ces différences peuvent permettre de caractériser les mode de pâturage journalier entre les deux groupes d'agnelles.

Quant à l'indicateur de dispersion, on observe à peu près les mêmes motifs de distribution des valeurs que ceux de l'indicateur d'étendue explorée pour les agnelles MAT. La variabilité des valeurs est faible aux périodes auxquelles les agnelles seraient en train de pâturer contrairement à ce qu'on observe avec l'indicateur d'étendue explorée. Le cas le plus visibles concerne les créneaux de 06-09h et de 18-21h des phases 2 et 3, avec beaucoup moins de variabilité pour l'indicateur de dispersion pour les agnelles MAT. En ce qui concerne les agnelles ART, la distribution des valeurs de dispersion ne suit pas le même motif que celles d'étendues explorées. Elle ne permet pas de bien distinguer les différents modes d'utilisation de l'espace par les agnelles, à l'inverse de celle des agnelles MAT.

Pour l'indicateur d'étendue explorée et de dispersion, les comparaisons statistiques réalisées à partir de l'ANOVA (cf. Tableau B.1) ne relèvent pas de différence significatives entre les 3 phases de pâturage pour chaque groupe d'agnelle, sauf pour l'étendue explorée des agnelles ART aux phases 2 et 3 entre 12h et 15h. A cette période, les agnelles ART ont occupé de petites étendues à la phase 2 alors qu'elles ont occupé des étendues relativement

élevées à la phase 3.

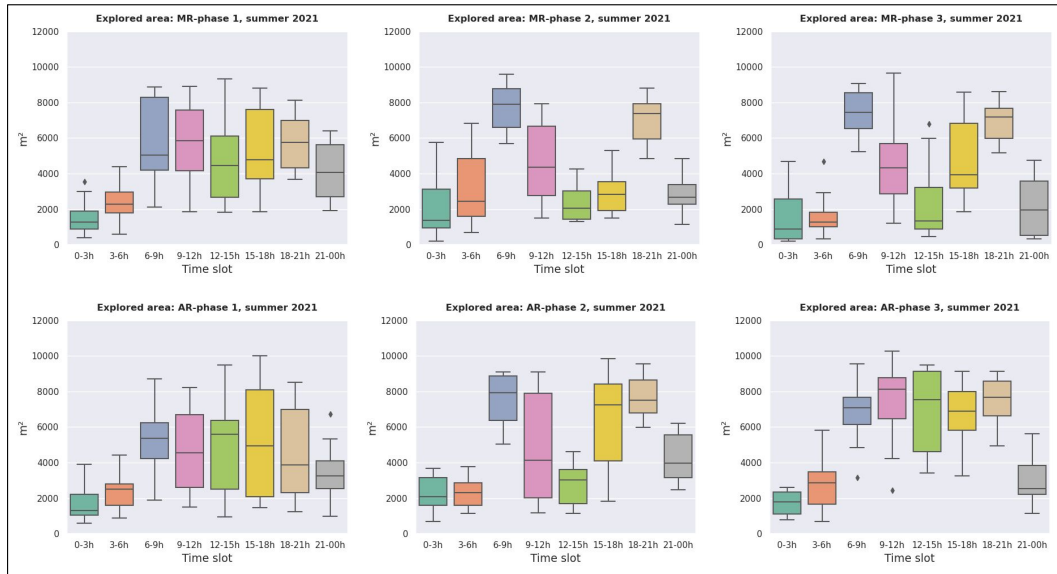


Figure 5.5 – Distribution des valeurs d'étendues explorées des agnelles MAT (ayant reçu un allaitement maternel) et ART (artificiellement allaitée) pour les créneaux horaires des phases de pâturage d'été.

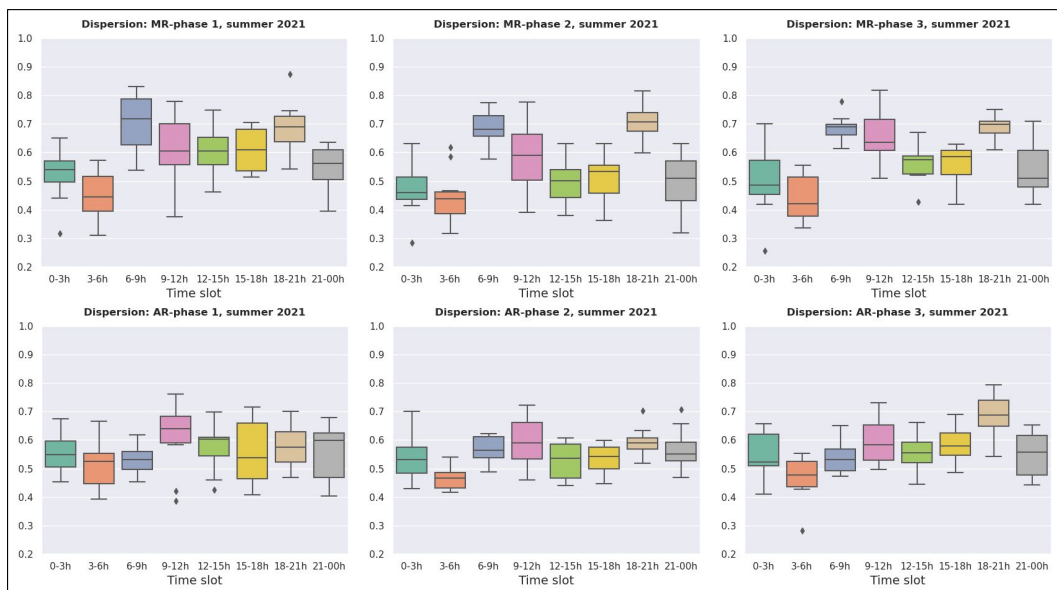


Figure 5.6 – Distribution des valeurs de dispersion des agnelles MAT (ayant reçu un allaitement maternel) et ART (artificiellement allaitée) pour les créneaux horaires des phases de pâturage d'été.

Durant l'automne, on observe moins de variabilité pour les créneaux horaires qu'en été. Trois modes d'utilisation de l'espace semblent apparaître pour les agnelles MAT en ce qui concerne l'étendue explorée (cf. Figure 5.7). Entre 21h00 et 06h00, les agnelles semblent occuper de petites proportions de l'espace. Cela pourrait correspondre au temps de repos ou de sommeil. Ensuite, elles se répandent à partir de 06h00 pour couvrir de grandes surfaces

du parc aux créneaux horaires de 06h00 à 12h00 et de 15h00 à 18h00. Une occupation des surfaces intermédiaires est observée au créneau horaire de 12-15h. Entre les deux phases de pâturage, dans l'ensemble, on n'observe pas de différence remarquables dans la distribution de l'étendue explorée pour les périodes similaires de la journée. La seule différence est observée au créneau de 06-09h où la distribution des valeurs est relativement faible en phase 2 par rapport à la phase 1. Un motif similaire à celui des agnelles MAT est également observé pour les agnelles ART en phase 1. En phase 2, uniquement 2 modes d'utilisation spatiale peuvent être observés : petites étendues explorées aux mêmes créneaux qu'en phase 1 et valeur d'étendue explorée élevée aux périodes de 06 à 21h. L'observation des deux groupes d'agnelles montrent que celles MAT utiliseraient plus les espaces relativement grandes en pâturant que celle ART, comme c'est le cas pour la phase 1. Quant à la phase 2, les valeurs d'étendues explorées sont plus variables pour les agnelles ART que les agnelles MAT.

En ce qui concerne l'indicateur de dispersion (cf. Figure 5.8), le motif de distribution des valeurs est différent de celui de l'étendue explorée pour les deux groupes d'agnelles. Il apparaît moins évident de distinguer les différents modes d'utilisation de l'espace étant donné les similitudes de distribution des valeurs pour les différentes périodes journalières.

L'ANOVA entre les périodes des deux phases de pâturage (cf. Tableau B.1) n'a montré aucune différence significative sur l'étendue explorée et la dispersion pour les deux groupes d'agnelles.

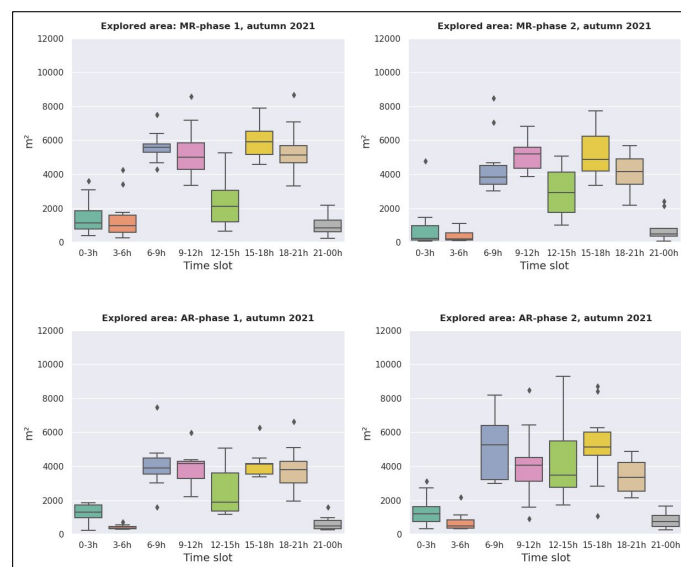


Figure 5.7 – Distribution des valeurs d'étendues explorées par les agnelles MAT (ayant reçu un allaitement maternel) et ART (artificiellement allaitée) pour les créneaux horaires des phases de pâturage d'automne.

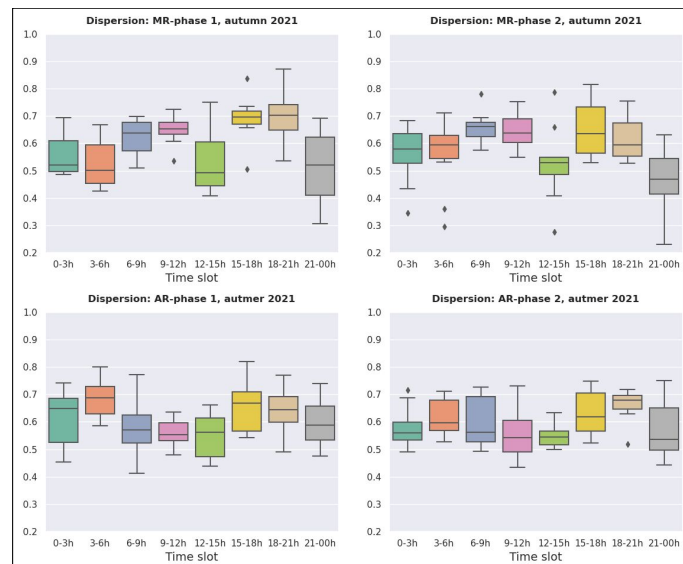


Figure 5.8 – Distribution des valeurs de dispersion par les agnelles MAT (ayant reçu un allaitement maternel) et ART (artificiellement allaitée) pour les créneaux horaires des phases de pâturage d'automne.

Densité de présence au fil de l'utilisation des parcs. Pour le pâturage d'été, la différence de densité d'utilisation des parcs est marginale pour les deux groupes d'agnelles (voir les figures 5.9 et 5.10). Elles utilisent de manière à peu près similaire les zones du parc pour s'alimenter et se reposer au cours des trois phases de pâturage.

En automne, la différence d'utilisation de l'espace est plus marquée au début et à la fin du pâturage (voir, figures 5.11 et 5.12). Au début, les agnelles MAT utilisent plus densément la partie nord du parc, probablement à cause de la disponibilité à cet endroit, d'une végétation pastorale appréciée par les agnelles, alors que la partie sud du parc est moins fréquentée. Dans la deuxième phase, on observe la situation inverse. Après avoir probablement épuisé les ressources fourragères au-dessus du parc lors de la première phase, les animaux se sont déplacés vers le centre et le sud du parc où une végétation pastorale a dû probablement pousser. Pour les agnelles ART, l'utilisation est plus perceptible dans la partie sud pendant la première phase de pâturage. Elles y passent probablement plus de temps à pâturer et à se reposer. A la deuxième phase de pâturage, les agnelles ont changé de zone de pâturage. Elles utilisent beaucoup plus la partie nord-est où la végétation auraient poussé.

Au cours des deux phases de pâturage pour les deux groupes d'agnelles, de fortes densités de présence sont observées autour de l'auge, car les agnelles s'y rassemblent tous les jours pendant une partie de la matinée pour manger des aliments concentrés et pour s'abreuver. La qualité non homogène des NDVIs (cf. Figures A.24, A.25, A.26 et A.27 en Annexe) des trois phases de pâturage d'été ne permet pas de décrire l'utilisation des zones du parc en fonction de leur densité de végétation. De même, les parcs étant presque plats avec de faibles amplitudes de pentes (cf. Figure A.23 en Annexe), l'observation visuelle du lien entre la densité de présence des agnelles et les types de pente n'est pas facile à établir.

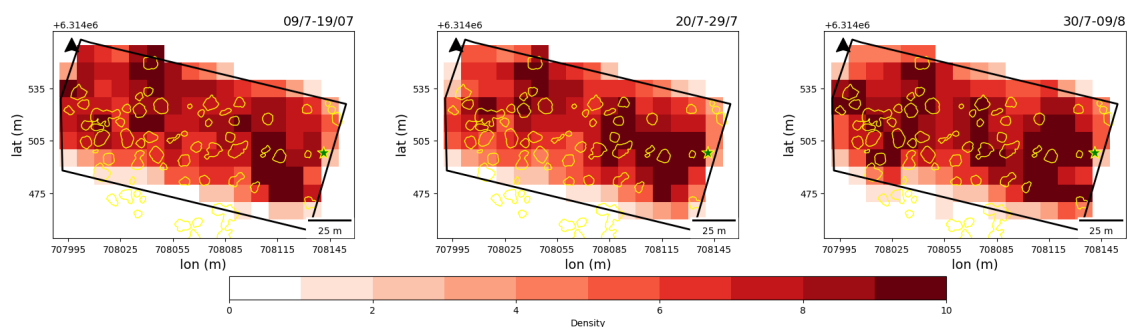


Figure 5.9 – Distribution spatiale de la densité d'utilisation de l'espace dans les zones de 10 m x 10 m pour le pâturage d'été par les agnelles MAT. Les contours (en jaune) représentent les limites des massifs de ligneux. L'étoile verte bordée de jaune représente la localisation de l'auge.

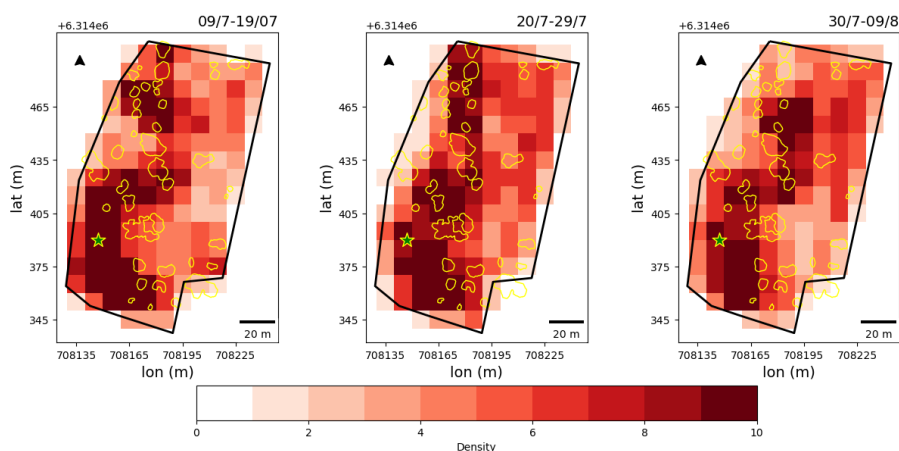


Figure 5.10 – Distribution spatiale de la densité d'utilisation de l'espace dans les zones de 10 m x 10 m pour le pâturage d'été par les agnelles ART. Les contours (en jaune) représentent les limites des massifs de ligneux. L'étoile verte bordée de jaune représente la localisation de l'auge.

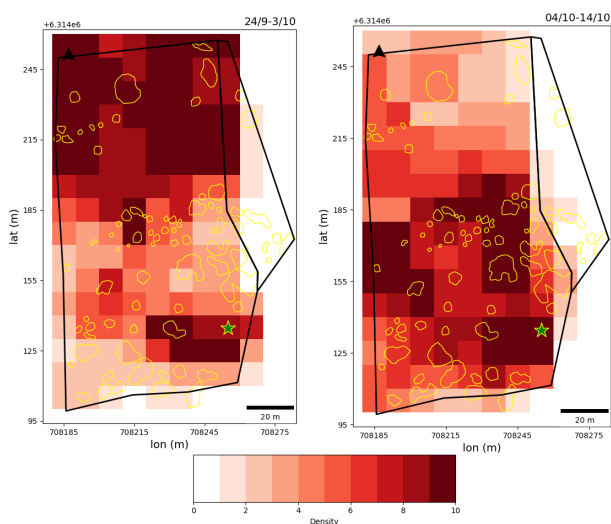


Figure 5.11 – Distribution spatiale de la densité d'utilisation de l'espace dans les zones de 10 m x 10 m pour le pâturage d'automne par les agnelles MAT. Les contours (en jaune) représentent les limites des massifs de ligneux. L'étoile verte bordée de jaune représente la localisation de l'auge.

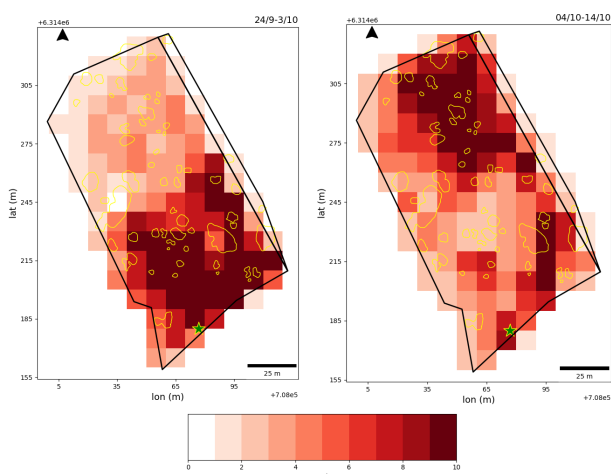


Figure 5.12 – Distribution spatiale de la densité d'utilisation de l'espace dans les zones de 10 m x 10 m pour le pâturage d'automne par les agnelles ART. Les contours (en jaune) représentent les limites des massifs de ligneux. L'étoile verte bordée de jaune représente la localisation de l'auge.

Densité de présence des agnelles au sein des massifs de ligneux. Les Figures 5.13 et 5.14 montrent la distribution de différentes densités d'utilisation des ligneux pour le pâturage d'été et d'automne par les agnelles MAT et ART. Elles pourraient permettre d'estimer l'utilisation des ligneux par les agnelles à différentes phases et saisons de pâturage.

Pour le pâturage d'été, on observe à travers la Figure 5.13 que les agnelles MAT utilisent plus les massifs de ligneux que les agnelles ART étant donné qu'elles possèdent plus de pixels de fortes densité. La densité de présence sous les ligneux apparaît similaire à toutes les phases

de pâturage pour chaque groupe d'agnelles.

En automne (cf. Figure 5.14) les agnelles MAT visitent plus souvent les massifs de ligneux que les agnelles ART à la première phase de pâturage. A la deuxième phase, elles visitent les ligneux à peu près de la même manière quant au nombre de pixels de forte densité. La densité de présence sous les ligneux aux deux phases de pâturage apparaît différente pour chacune des groupes d'agnelles contrairement au pâturage d'été.

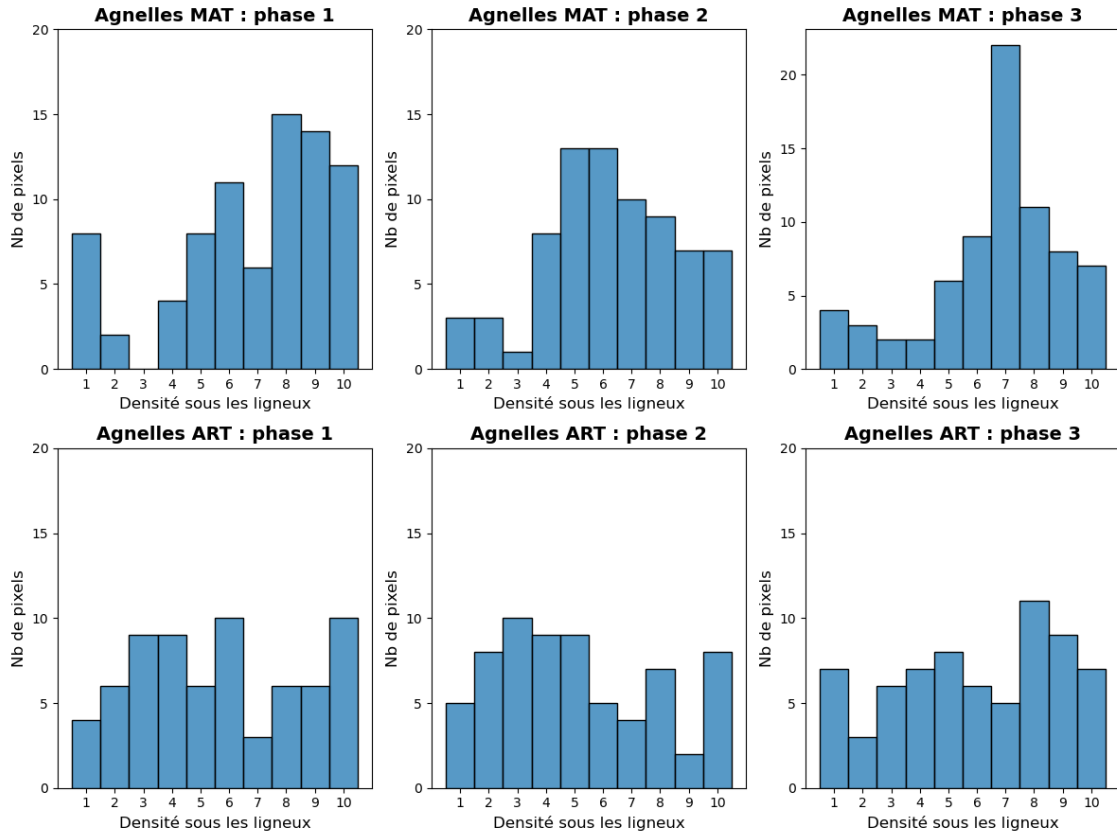


Figure 5.13 – Nombre de pixels par catégorie de densité de présence des agnelles sous les massifs de ligneux pour le pâturage d'été.

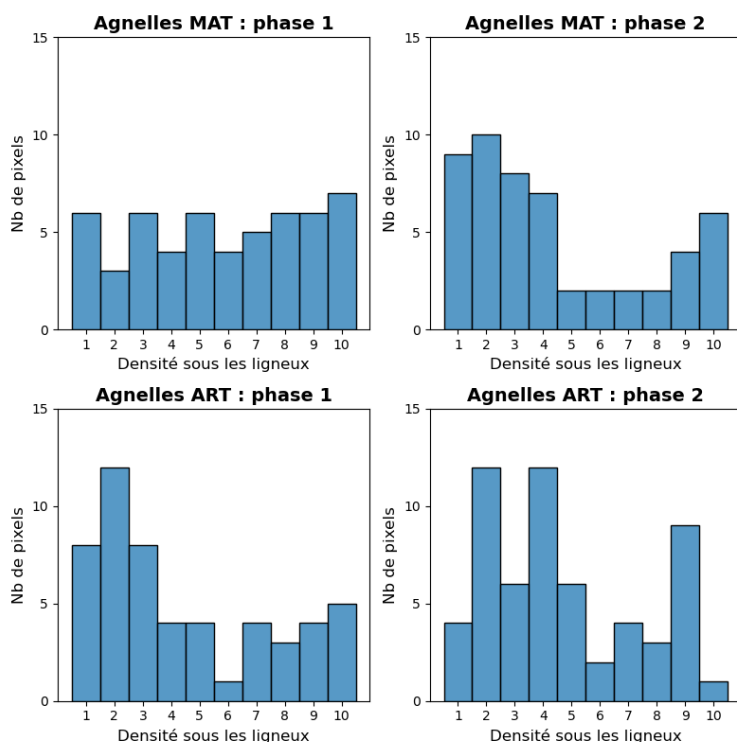


Figure 5.14 – Nombre de pixels par catégorie de densité de présence des agnelles sous les massifs de ligneux pour le pâturage d'automne.

Différences d'utilisation des parcs entre les deux groupes d'agnelles

Pour les deux périodes de pâturage et pour toutes les phases (début, milieu et fin de l'utilisation du parc), les résultats de l'ANOVA ($p < 0,1$) ont montré plus de similitudes que de différences dans la distribution des valeurs des indicateurs (étendue explorée et dispersion des animaux) pour chaque paire de créneaux horaires entre les deux groupes d'agnelles (cf. Tableau 5.8). En été, des différences significatives ont été observées pour l'étendue explorée pour les créneaux horaires de 15h-18h de la phase 2 et 12h-15h de la phase 3. Pour ces créneaux horaires respectives, les agnelles MAT pourraient être au repos car elles occupent de petites surfaces par rapport aux agnelles ART qui semblent occuper des surfaces relativement grandes, probablement pour le pâturage. Pour l'indicateur de dispersion, des différences sont observées en été au créneau horaire de 06h-09h aux phases 1 et 3. A ces moments, les agnelles MAT se dispersent plus que les agnelles ART probablement pour le pâturage. En automne, la seule différence significative a été observée pour l'indicateur de dispersion à la première phase de pâturage au créneau horaire de 03h-06h. Les agnelles ART semblent être encore actives à ce moment, alors que les agnelles MAT semblent être au repos. Globalement, par la lecture visuelle de la distribution des valeurs, on peut observer que les agnelles MAT se dispersent davantage au moment du pâturage que les agnelles ART.

Table 5.8 – P valeurs obtenus à partir de l'ANOVA entre les indicateurs de l'utilisation journalière de l'espace pour les mêmes créneaux horaires entre les deux groupes d'agnelles pour chaque phase de pâturage. P = Phase, NS = non significatif.

Créneau horaire	Pâturage d'été						Pâturage d'automne			
	Etendue explorée			Dispersion des agnelles			Etendue explorée		Dispersion des agnelles	
	P 1	P 2	P 3	P 1	P 2	P 3	P 1	P 2	P 1	P 2
0-3h	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
3-6h	NS	NS	NS	NS	NS	NS	NS	NS	0,01	NS
6-9h	NS	NS	NS	0,01	NS	0,00	NS	NS	NS	NS
9-12h	NS	NS	0,10	NS	NS	NS	NS	NS	NS	NS
12-15h	NS	NS	0,00	NS	NS	NS	NS	NS	NS	NS
15-18h	NS	0,00	NS	NS	NS	NS	0,06	NS	NS	NS
18-21h	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
21-0h	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS

Interaction des facteurs du milieu

Les résultats de la régression linéaire multiple entre les indicateurs spatiaux (étendue explorée, dispersion et densité d'utilisation de l'espace), temporelles (étendue explorée et dispersion) et les facteurs du milieu n'a pas montré de liaison significative. Les coefficients de détermination présentés dans les Tableaux 5.9 et 5.10 sont très bas. L'élimination progressive des variables non significatives n'a pas permis d'améliorer le modèle de régression pour expliquer les indicateurs sélectionnés.

Table 5.9 – Coefficients de détermination de la régression linéaire multiple entre les indicateurs sélectionnés et les variables spatiales du milieu : *pente, NDVI, distance à l'auge, distance au ligneux le plus proche.*

Variables utilisées	R ² par groupe d'agnelles			
	MAT été	ART été	MAT automne	ART automne
Étendue explorée	0,17	0,18	0,02	0,1
Dispersion	0,08	0,01	0,02	0,06
Densité de présence	0,03	0,08	0,01	0,01

Table 5.10 – Coefficients de détermination de la régression linéaire multiple entre les indicateurs sélectionnés et les variables temporelles du milieu (paramètres météorologiques) : *humidité instantanée, rayonnement global, température horaire, vitesse du vent et précipitations².*

Variables utilisées	R ² par groupe d'agnelles			
	MAT été	ART été	MAT automne	ART automne
Étendue explorée	0,08	0,04	0,16	0,17
Dispersion	0,03	0,03	0,01	0,03

5.4.6 Discussion

Qualité des indicateurs testés pour la caractérisation du comportement des animaux

Dans cette étude, nous avons testé des indicateurs permettant de caractériser l'utilisation spatiale et temporelle de l'espace par les animaux. Nous avons par la suite analysé l'influence des facteurs du milieu sur ces indicateurs. Compte tenu de la nature des données utilisées, principalement des données GPS, nous avons eu l'opportunité d'aborder la question par l'analyse de motifs de points (Gatrell et al., 1996; Ben-Said, 2021) pour observer comment les agnelles, utilisent les parcs à différents moments de la journée et à différentes phases de pâturage. Parmi les informations essentielles à la gestion du bétail, l'intensité du pâturage est souvent utilisée (Kawamura et al., 2005; Akasbi et al., 2012; Cheng et al., 2021). Elle permet aux éleveurs de moduler l'utilisation du milieu pour assurer la conduite du pâturage. En plus de l'intensité du pâturage, la valorisation des ressources des parcours par le pâturage est cruciale. Elle dépend des possibilités des éleveurs et de la capacité des animaux à profiter de la variabilité des ressources offertes par les milieux explorés (Riedel et al., 2013). Ces aspects de l'utilisation du territoire peuvent être appréhendés à travers la manière dont les animaux s'étalent ou se dispersent lors du pâturage. Dans cette étude, nous avons choisi d'utiliser les indicateurs de surface explorée, de dispersion dans l'espace et de densité d'utilisation de l'espace (Spielman, 2017) pour caractériser ces différents aspects du comportement animal.

Les indicateurs de surface explorée et de dispersion ont permis de caractériser l'utilisation quotidienne et saisonnière des parcs par les agnelles. A travers ces indicateurs, il est donc possible d'observer à quel moment de la journée les animaux sont probablement au repos ou au pâturage. La complémentarité de ces deux indicateurs permet, d'une part, de confirmer les résultats obtenus par chacun d'eux et, d'autre part, de compenser les difficultés d'interprétation dues à la variabilité des valeurs pour certains moments de la journée. En revanche, la représentation spatiale de l'utilisation des parcs par les agnelles a été faite par l'indicateur de densité. Nous avons pu observer les différences de niveau de présence dans les zones des parcs lors des différentes phases de pâturage d'une même saison.

Des différences dans l'utilisation quotidienne des parcs par les agnelles MAT et ART ont été observées à certains créneaux horaires pendant certaines phases de pâturage. Globalement, les résultats ont montré qu'il n'y avait pas de différence significative entre les deux groupes. Des différences d'utilisation de l'espace entre l'été et l'automne pour les mêmes créneaux horaires ont été notées. Les agnelles étaient moins dispersées dans les parcs pendant certains moments de la journée en l'été, alors qu'en automne, elles étaient relativement dispersées. Les différences plus notables aux créneaux de 09-12h et 15-18h.

Contribution et applications des indicateurs à la diversité des conditions pastorales

Dans les conditions variées et hétérogènes des écosystèmes agro-sylvo-pastoraux, ce travail propose une nouvelle manière d'appréhender l'utilisation des parcours par les animaux à partir des données issues des capteurs GPS. En plus de la connaissance du parcours, les éleveurs ont besoin de savoir comment les animaux explorent les ressources végétales dans le temps et dans l'espace pour ajuster le pâturage (Guérin et al., 2001). Ensemble, les trois indicateurs utilisés permettent de mieux comprendre l'utilisation spatiale et temporelle des parcours par les animaux et peuvent contribuer à la prise de décision sur la gestion du pâturage à plusieurs échelles d'utilisation des parcours.

Au-delà des caractéristiques de notre étude où les animaux étaient placés en conditions expérimentales, ces indicateurs peuvent également être utilisés dans le cadre d'un élevage en conditions opérationnelles. Dans ces conditions, les animaux exploitent des milieux plus variables et hétérogènes en termes de ressources pastorales (Guérin et al., 2001). L'indicateur de dispersion basé sur la distance moyenne la plus proche (Clark and Evans, 1954; Dixon, 2001) permet de définir le niveau de dispersion d'un ensemble de points par rapport au milieu utilisé. Quel que soit le milieu exploré, il permet de distinguer comment les animaux se dispersent pendant la journée ou la phase de pâturage. L'indicateur de surface explorée (Akkiraju et al., 1995) variera en fonction du milieu utilisé (les grandes surfaces explorées seront observées dans les milieux plus étendus). Comme l'indicateur de dispersion, il peut être utilisé pour caractériser et différencier l'utilisation temporelle de l'espace (sur la journée et sur une phase de pâturage) par les animaux, comme le montrent nos résultats. L'indicateur de densité est également généralisable à différents environnements. Cependant, dans les parcours, la connaissance du milieu est nécessaire pour interpréter l'intensité de pâturage des animaux. En effet, l'intensité d'utilisation des espaces par les animaux est définie en tenant compte de la diversité des milieux pastoraux (Guérin et al., 2001). Une même densité d'utilisation de l'espace ne détermine pas la même intensité d'utilisation pour deux milieux différents. Cet indicateur sera utile pour aider l'éleveur à identifier les zones fréquemment utilisées et celles qui le sont moins afin d'ajuster l'utilisation de l'espace pastoral sur de ses objectifs.

Les indicateurs choisis reflètent l'utilisation de l'espace par un groupe d'animaux. Dans notre étude de cas, plus de la moitié des agnelles des deux groupes étaient équipées de GPS. Dans certaines exploitations privées, il est difficile de trouver plus d'un GPS par troupeau pour des raisons pratiques, économiques et environnementales. L'applicabilité de ces indicateurs dans ces conditions d'exploitation est discutable. En effet, l'indicateur de dispersion et la surface explorée conduiront à des résultats similaires du point de vue de l'interprétation, tant que le groupe d'animaux reste ensemble. Dans ce cas, il est nécessaire de placer le GPS sur un animal leader qui puisse représenter l'activité du groupe. Les travaux du projet CLOCHÈTE (Grisot et al., 2022) sur l'intégration des outils numériques dans la gestion de petits ruminants ont défini les critères de sélection d'un animal leader pour représenter le groupe.

Avec l'indicateur de densité, les résultats permettront également des interprétations cohérentes. Toutefois, les éleveurs peuvent interpoler la densité par rapport au nombre d'animaux pour évaluer approximativement l'activité des animaux dans l'environnement.

Enfin, d'autres considérations pratiques, telles que la précision des données GPS et l'influence des caractéristiques des différents milieux sur la qualité de ces données et son impact sur les résultats, doivent être prises en compte. Plusieurs facteurs, tels que le relief, les arbustes, la météo, etc., peuvent avoir un impact sur la précision des GPS et donc sur la qualité des résultats des indicateurs. Dans cette étude, nous n'avons pas évalué l'influence de ces facteurs. Nous avons défini un critère d'un nombre minimum de 220 points GPS quotidiens pour l'intégration dans nos analyses. C'était pour nous un moyen simple de définir les données à prendre en compte.

Effet de la taille des parcs et des facteurs du milieu sur le comportement des agnelles

L'analyse de la corrélation entre les indicateurs d'utilisation des parcs par les agnelles et les variables du milieu a produit des coefficients de détermination faibles (cf. Tableaux 5.9 et 5.10). Le cadre expérimental caractérisant notre étude ne nous a pas permis de mettre en évidence l'influence des facteurs du milieu sur le comportement des animaux. Nous considérons que la taille des parcs (environ 1ha) est déterminante étant donné qu'elle limite les animaux à varier leur utilisation en fonction des facteurs du milieu. Les animaux auront tendance à utiliser le parc de la même manière pendant une phase ou une saison de pâturage. L'effet du parc sur le comportement des animaux a été mentionné par les éleveurs lors de nos entretiens.

En raison de la taille du parc, il n'est pas facile d'observer l'impact de la distance à la mangeoire, car les animaux n'ont pas la possibilité de se déplacer sur de longues distances. En plus de l'effet de la taille des parcs, leurs faibles pentes et n'ont eu aucune influence sur le déplacement des animaux. La distribution des paramètres météorologiques montre que la température a atteint 30 °C le jour le plus chaud. Dans le travail de (?), les auteurs indiquent que la température critique qui peut impliquer un changement dans l'utilisation de l'espace par les herbivores domestiques est de 25 °C. À ces températures, les animaux ont tendance à rester relativement groupés et à développer le pâturage nocturne. Les indicateurs de surface explorée et de dispersion produits ont montré que les deux groupes d'agnelles avaient tendance à rester groupés pendant les moments chauds de la période estivale (c'est-à-dire de 12 à 18 h). La différence était plus prononcée pour les agnelles MR. Ceci est probablement dû en partie au fait que lorsqu'il fait chaud, les animaux se regroupent sous les massifs de ligneux. Cependant, dans le cas des températures inférieures à la température critique (plus de 75% dans notre cas), l'impact de ce facteur sur le comportement des animaux n'est pas évident.

Nous avons considéré que des analyses en conditions opérationnelles dans des milieux ouverts et variés permettront de mettre en évidence l'influence des facteurs du milieu sur le

comportement spatial des animaux.

5.4.7 Conclusion

Dans cette section du travail, nous avons expérimenté l'application des méthodes permettant de caractériser l'utilisation des parcours par deux groupes d'agnelles, l'un avec allaitement maternelle (MAT) et l'autre avec allaitement artificiel (ART) à partir des données GPS. Nous avons par la suite visé l'évaluation de l'influence des facteurs du milieu sur le comportement spatial des animaux. La caractérisation spatio-temporelle du comportement alimentaire des animaux a été réalisée à partir des indicateurs d'étendue explorée, de dispersion et de densité. Et les informations sur la pente, l'état de la végétation, les paramètres météorologiques et les points d'intérêts, ont été utilisées comme facteurs explicatifs du comportement spatial des animaux.

Les résultats obtenus ont montré que les indicateurs choisis peuvent permettre de caractériser l'utilisation journalière des parcs (repos et pâturage) par les animaux. La comparaison du comportement des agnelles au fil de l'utilisation du parc (début, milieu et fin de saison) montre quelques variations au sein de chaque groupe d'agnelles. Cependant, les différences observées n'ont pas été statistiquement significatives. La comparaison de la distribution des indicateurs d'étendue explorée et de dispersion entre les deux groupes d'agnelles montre que les agnelles MAT restent plus dispersées pendant certains moments de pâturage que les agnelles ART. Les indicateurs utilisés dans cette étude peuvent également être généralisés en conditions opérationnelles étant donné leurs caractéristiques permettant d'observer les différences dans le mode d'utilisation spatiale et temporelle des parcours.

L'analyse de corrélation entre les indicateurs du comportement des animaux et les facteurs du milieu n'a pas permis d'observer de liaison. Les conditions expérimentales, notamment la taille du parc n'a pas favorisé les possibilités des agnelles à explorer l'espace dans les conditions variables des facteurs considérés. Nous estimons que l'application de nos analyses dans des conditions d'élevage utilisant des parcours plus larges et variés permettra de mettre en évidence l'influence des facteurs du milieu sur le comportement des animaux.

Chapitre 6

Discussion générale et conclusion

6.1 Vue générale des travaux

Nous avons mené les études pour proposer des approches techniques et méthodologiques permettant d'intégrer les données hétérogènes du pastoralisme et de produire des indicateurs d'aide à la décision en élevage pastoral. Nos travaux ont permis de répondre aux principaux objectifs : i) d'identification des besoins des éleveurs et des données hétérogènes du pastoralisme, et de définition et conception d'un système d'informations (SI) capable d'intégrer ces données, ii) de collecte, prétraitement et préparation des données hétérogènes pour faciliter la caractérisation de l'utilisation des milieux par les troupeaux et servir de référence pour la conduite du pâturage, et iii) de définition des techniques et méthodes d'analyse adaptées à la production d'indicateurs d'aide à la décision en élevage pastoral.

Pour atteindre ces objectifs, nous avons privilégié l'approche de co-conception du SI ensemble avec les parties prenantes constituées d'éleveurs, chercheurs, enseignants et de l'institut d'élevage (IDELE). Cette approche nous a permis de définir les considérations techniques et fonctionnelles du SI en adéquation avec les besoins des éleveurs et les caractéristiques des données hétérogènes identifiées. L'analyse de l'utilisation des parcours par les troupeaux s'est basée sur la caractérisation des milieux pâturés d'une part et de l'utilisation spatio-temporelle des parcours par les troupeaux d'autre part. Nous avons privilégié les méthodes d'extraction de séquences fréquentes et d'analyse de motifs de points pour caractériser l'utilisation spatio-temporelle des parcours par les troupeaux. Ces méthodes sont adaptées aux données spatio-temporelles collectées dans le cadre du pastoralisme.

Les fonctionnalités du SI ont été définies sur base de l'analyse des besoins des éleveurs pastoraux en rapport avec la conduite du pâturage sur parcours. Cela a permis par la suite d'identifier les données à intégrer dans le SI pour répondre aux besoins des éleveurs en termes d'indicateurs et de documents techniques pour la conduite du pâturage sur parcours. Le lac de données spatiales a été défini comme architecture du SI en adéquation avec les contraintes techniques et fonctionnelles imposées par les données hétérogènes identifiées. Le choix de l'architecture du lac de données a été fait sur la technologie Hadoop d'Apache qui propose des applications libres, améliorées par de grandes communautés et pouvant former un écosystème répondant aux besoins d'intégration des données hétérogènes du pastoralisme. Pour parvenir à une bonne caractérisation des milieux pastoraux et de l'utilisation des parcours par les troupeaux, nous avons défini les critères de préparation et de prétraitement des données spatiales. Ces critères sont basés sur les récents travaux autour de la cartographie des milieux et du comportement spatial des troupeaux en régions Occitanie et PACA. En fin, à travers une analyse de données en situation expérimentale, nous avons commencé à évaluer la pertinence des méthodes d'analyse choisies pour caractériser le comportement spatial des troupeaux et hiérarchiser l'influence des facteurs du milieu.

Les travaux réalisés nous ont permis d'apporter deux principales contributions : opérationnelle et scientifique. La contribution opérationnelle concerne la définition et la conception

d'un SI en adéquation avec les besoins des éleveurs pastoraux et les enjeux liés aux données hétérogènes, et pouvant servir de référence pour documenter la conduite du pâturage sur parcours. La contribution scientifique concerne la définition des méthodes d'analyse permettant de produire des indicateurs pouvant traduire l'utilisation spatiale des milieux pastoraux variés par les troupeaux et de hiérarchiser l'influence des facteurs du milieu sur le comportement spatial des troupeaux.

Dans ce chapitre dédié à la discussion de nos résultats, nous évaluons l'intérêt du SI dans le cadre de l'élevage pastoral en zones méditerranéennes et montagneuses françaises, et sa généralisation dans d'autres systèmes d'élevage pastoraux. Nous présentons également les perspectives de recherche visant à améliorer les fonctionnalités du SI en vue d'enrichir la documentation du pastoralisme en régions Occitanie et PACA.

6.2 Évaluation du lac de données spatiales dans le cadre de l'élevage pastoral

6.2.1 Intérêt des données pour la gestion de l'élevage pastoral en zones méditerranéennes et montagneuses françaises

Le lac de données spatiales élaboré dans le cadre de notre étude propose des solutions pouvant permettre de répondre aux enjeux de production et environnementaux du pastoralisme en zones méditerranéennes et montagneuses françaises. Il s'appuie sur la caractérisation des facteurs du milieu et de l'utilisation spatiale des parcours par les troupeaux pour documenter leurs comportements alimentaires et spatiaux. Cette documentation contribuera à la fourniture d'informations variées sur l'utilisation des parcours par les troupeaux afin de servir de référence pour la conduite du pâturage en zones méditerranéennes et montagneuses françaises. Une meilleure connaissance de ces deux aspects de gestion du pâturage sur parcours devra permettre en même temps de valoriser les milieux pâturés (Agreil et al., 2004; Gautier et al., 2011) et faciliter aux troupeaux l'accès à une diversité des ressources pastorales pour leurs besoins alimentaires (Meuret, 2007).

La caractérisation des facteurs du milieu pastoral dépend de la qualité des données spatiales utilisées. En effet, la variété des données spatiales conduit à une caractérisation variable des facteurs du milieu en termes de niveau de détails d'informations qu'elles procurent. Compte tenu de la disponibilité des données, nous avons considéré comme facteurs du milieu : la topographie, la végétation et la météorologie. Les données altimétriques de l'IGN (MNT RGE ALTI) possèdent une résolution spatiale suffisante (5m puis rééchantillonnées à 10m dans cette étude) pour représenter la topographie des milieux utilisés par les exploitations pastorales. Ces données peuvent être utilisées pour représenter avec une meilleure précision la topographie des milieux pâturés à différents niveaux de l'exploitation au territoire. Seules, les données d'occupation des sols utilisées pour décrire la végétation des parcours se limitent

à la description de classes de végétation. En plus de leur limites en termes de précisions des classes pastorales, elles ne permettent pas de décrire la végétation consommées ou consommables par les troupeaux. De ce fait, nous avons privilégié la complémentarité des données de densité de végétation issues du NDVI pouvant être utilisées comme indicateur d'état de la végétation au fil de l'utilisation des parcours par les troupeaux mais aussi des variations des conditions météorologiques. Le NDVI (Normalized Difference Vegetation Index) est un indicateur intéressant pouvant être utilisé dans la modélisation de la biomasse des pâturage (Cheng et al., 2021; Kowal et al., 2021), et la représentation des massifs de ligneux qui présentent un enjeu important pour le pastoralisme. L'OSAVI (Optimized Soil Adjusted Vegetation Index) peut également être utilisé pour limiter l'influence de la réflectance dû au sol environnant pendant les périodes sèches. Enfin, l'intégration des paramètres météorologiques pour décrire les milieux explorés sera plus facile pour les sites expérimentaux qui sont pour la plupart mieux équipés en capteurs que les élevages privés.

La caractérisation du comportement spatial des troupeaux dépend de la fréquence d'acquisition des données GPS. Dans les sites expérimentaux, la fréquence de 5 minutes a été adoptée au regard des résultats obtenus dans des travaux précédents (Putfarken et al., 2008; Akasbi et al., 2012; Evangelou et al., 2014; Feldt and Schlecht, 2016) et des conclusions tirées des projets CLOCHÈTE (Grisot et al., 2022) et P@stor_All (Candela, 2022). Elle représente un compromis entre la précision et l'autonomie des capteurs GPS. Les données issues du site expérimental du domaine de la Fage auxquelles nous avons appliqué les méthodes d'analyse explorées, ont permis d'obtenir des indicateurs pouvant traduire le comportement spatial des troupeaux. Ces indicateurs spatiaux et temporels produits contribuent à la compréhension de l'utilisation de l'espace pâturé par les troupeaux. Ils décrivent la variation de l'utilisation de l'espace au fil de la journée et au cours des jours, et peuvent par la suite être associés aux facteurs du milieu pour hiérarchiser leurs influences.

6.2.2 Complémentarité par rapport à d'autres outils développés pour la gestion de l'élevage pastoral

Les applications du lac de données spatiales apportent une complémentarité aux outils déjà développés dans le cadre des projets de l'UMT Pasto pour accompagner les acteurs pastoraux dans la gestion de l'élevage pastoral. Parmi ces outils nous pouvons citer le référentiel Pastothèque, la méthode Mil'ouv qui visent principalement l'identification et le diagnostic des ressources pastorales, et le Rami Pastoral et Stratpasto destinés à évaluer et améliorer l'utilisation des parcours par les troupeaux (Etienne et al., 2021). Le référentiel Pastothèque propose des fiches décrivant les ressources végétales et les fonctions des milieux pastoraux sous plusieurs angles complémentaires : principales espèces et types biologiques, fonctionnalités pastorales, ressource pastorale disponible, enjeux environnementaux, critères pour une gestion pastorale durable du milieu. La méthode Mil'ouv (Péglion et al., 2017) permet quant à elle de réaliser un diagnostic et une évaluation écologique et pastorale des exploitations. Le

Rami Pastoral (Farrié et al., 2015) est un jeu sérieux qui représente le fonctionnement d'un système d'alimentation pastoral ou agro-pastoral. Il se réalise sur un plateau de jeu avec des cartes et des baguettes connecté à un simulateur informatique qui s'appuie sur une base de données qui rassemble les informations concernant la qualité des ressources pâturées et les variations probables d'états corporels des troupeaux en réponse à l'alimentation. Il peut ainsi servir de référence sur la conduite du pâturage en se basant sur les résultats obtenus dans différents cas observés. Enfin, Stratpasto (Malzac et al., 2018), est une base de données en ligne qui propose une analyse fine du système d'alimentation à l'échelle de l'atelier ou de l'exploitation.

Deux principales différences complémentaires peuvent être relevées entre ces différents outils et le lac de données spatiales proposé dans le cadre du projet P@stor_All : l'échelle d'analyse des ressources pastorales, et la caractérisation de l'utilisation des parcours par les troupeaux. En ce qui concerne l'analyse des ressources pastorales, ces outils apportent une profondeur de précision que les données identifiées pour le lac de données spatiales ne possèdent pas, à travers l'identification et les diagnostics réalisés sur terrain par les opérateurs concernés. De ce point de vue, ils permettent de collecter des informations réalistes sur les ressources pastorales. Les méthodes d'analyse devant être intégrées dans le lac de données spatiales contribuent plus par rapport à ces outils dans la documentation de l'utilisation spatio-temporelle des parcours par les troupeaux. Le Tableau 6.1 présente plus d'éléments de comparaison entre ces différents outils présentés. Nous ajoutons à la liste l'application de localisation des troupeaux en temps réel de GéoVie et Pâtur'Ajuste, le réseau de partage et d'échange d'expérience sur le pastoralisme. La complémentarité entre ces outils peut renforcer davantage la documentation de l'élevage pastoral en zones méditerranéennes et montagneuses françaises.

Table 6.1 – Comparaison des fonctionnalités proposées par les différents outils d'aide à la décision pour la conduite du pâturage sur parcours. **X** = forte intégration de la fonction, (x) = faible intégration de la fonction.

Fonction	SI P@stor-all	Pasto-thèque	Rami Pastoral	Mil'ouv	Strat Pasto	Pâtur'ajuste	dispositifs GPS d'élevage
Savoirs							
Partage d'expérience	X		X			X	
Connaissances scientifiques	X				X		
Références							
Mode d'exploitation des parcours	X	X	X			X	
Ressources pastorales selon les milieux	(x)	X	(x)	X	X	X	
Autres références techniques	X						
Données locales							
Description des milieux	(x)			X			
Localisation du troupeau en temps réel							X
Description du comportement spatial du troupeau	X						

Dans le SI P@stor-all, les modes d'exploitation et les ressources selon les milieux seront disponibles via des liens vers les références externes.

6.3 Contributions techniques et méthodologiques du lac de données spatiales pour la gestion et l'analyse des données du pastoralisme

6.3.1 Considérations pour l'hétérogénéité et la variété des conditions pastorales

Les applications du lac de données ont été expérimentées dans différentes études telles que : les villes intelligentes (Kafando et al., 2020), l'intelligence de marché (Llave, 2018), le traitement automatique des langues (Erturk and Shi, 2016), le traitement de données d'aviation (García et al., 2017) et l'analyse des réseaux sociaux (Farrugia et al., 2016). Ces différentes études ont montré les performances du lac de données pour gérer les données hétérogènes. La revue bibliographique réalisée ne nous a pas permis d'identifier les applications du lac de données dans le domaine agricole. Pour notre étude, le choix du lac de données s'est basé sur sa capacité à assurer l'intégration des données hétérogènes du pastoralisme non préalablement transformées, leur accès par différents acteurs et l'augmentation en volume des données. Les systèmes décisionnels basés sur les entrepôts de données utilisent les bases de données relationnelles dont ils extraient et transforment les données selon un formalisme bien défini afin de proposer les indicateurs d'aide à la décision (Colhoun, 2018). Pour des applications pastorales, les données hétérogènes collectées peuvent être manipulées à travers diverses méthodes pour produire des indicateurs adaptés à chaque contexte d'élevage. Les

critères d'analyse de données peuvent être différents d'une exploitation pastorales à une autre car les troupeaux utilisent différemment l'espace en fonction de différents facteurs présentés au chapitre 3 (les troupeaux élevés en parc utiliseront différemment l'espace que les troupeaux élevés en plein air). Ceci oblige de garder les données à leur état brutes au sein du SI afin d'offrir la possibilité de tester différentes approches pour décrire l'utilisation des parcours par les troupeaux.

Le lac de données spatiales que nous proposons met en avant les outils présentés dans le cadre de précédentes études, dont celle de (Kafando et al., 2020), pour répondre aux besoins de l'élevage pastoral. Il privilégie la conservation de la nature des données ingérées grâce au composant HDFS (Hadoop Distributed File System, (Honnutagi, 2014)) qui assure le stockage distribué des données. Grâce à HDFS le SI assure également les performances de stockage et de traitement de gros volumes de données et peut donc bien gérer la montée en charge. En plus, le lac de données Hadoop d'Apache permet l'intégration d'outils de gestion de métadonnées des données spatiales (Kafando et al., 2020), d'analyses et de visualisation formant ainsi un écosystème complet pour la prise de décision.

6.3.2 Pertinence des méthodes d'analyse d'indicateurs pour décrire le comportement spatial des troupeaux

Le but des analyses réalisées a été de proposer des indicateurs pouvant permettre de caractériser l'utilisation de l'espace par les troupeaux et de comprendre les mécanismes qui déterminent leur comportement spatial. Parmi plusieurs méthodes testées au cours de nos travaux, nous avons maintenu l'extraction de séquences fréquentes (Srikant and Agrawal, 1996; Fournier-Viger et al., 2008) et l'analyse de motifs de points (Ben-Said, 2021). Ces deux méthodes d'analyse ont été utilisées dans divers domaines d'application. Pour l'extraction de séquences fréquentes nous pouvons citer : la bio-informatique (Wang et al., 2007), le e-learning (Fournier-Viger et al., 2008) et l'analyse de texte (Pokou et al., 2016). L'analyse de motif de points trouve plus des applications en écologie (Velázquez et al., 2016; Ben-Said, 2021). Des applications en épidémiologie peuvent également être identifiées (Gatrell et al., 1996). Ces deux méthodes sont également adaptées à l'analyse des données spatiales décrivant les positions géographiques des troupeaux au cours de l'utilisation du parcours.

L'extraction de séquences fréquentes permet d'isoler les événements qui se succèdent dans le temps. Dans le cadre de l'utilisation de l'espace par les troupeaux, il peut permettre d'identifier et fournir une base de connaissance sur la succession des lieux visités par les troupeaux. Les caractéristiques de ces lieux peuvent par la suite être extraites pour expliquer les facteurs qui déterminent les choix opérés par les troupeaux. Le cadre expérimental sur lequel nous nous sommes basés n'a pas permis de tester cette méthode notamment à cause de la petite taille des parcs utilisés présentant à la fois une contrainte spatiale (les troupeaux étant limités dans un petit espace) et une variabilité moins importante des facteurs du milieu. Cependant, nous considérons que cette méthode d'analyse permettra d'extraire des connais-

sances sur le comportement spatial des troupeaux dans toutes les formes de conduites du pâturage : en pâturage libre parce que les troupeaux se construisent des circuits de pâturage, en pâturage en grand parc où les troupeaux réalisent des aller-retour entre zones au cours de la journée, et en pâturage gardé où l'éleveur conduit son troupeau dans une suite de secteurs.

La caractérisation de l'utilisation de l'espace par les troupeaux réalisées dans le cadre de l'expérimentation du domaine de la Fage a été faite à partir de l'analyse de motifs de points. Ces méthodes d'analyse nous ont permis de produire 3 indicateurs complémentaires pouvant traduire l'activité des troupeaux dans différentes conditions de pâturage. Dans le cadre de cette expérimentation, l'indicateur d'étendue explorée et de dispersion ont permis de distinguer le mode d'utilisation de l'espace par les troupeaux au cours de la journée. Associés à l'expertise du gestionnaire du troupeau, ces indicateurs peuvent permettre d'identifier les instants de pâturage et de repos des troupeaux à différents moments de la journée. L'indicateur d'intensité de présence aux endroits du parc offre quant à lui la possibilité d'observer la variation de l'utilisation de l'espace au fil des jours ou des phases de pâturage. Pour l'été et l'automne, le découpage en début, milieu et fin de pâturage a montré comment les troupeaux varient ou non l'utilisation des parcs utilisés.

Ensemble, ces trois indicateurs peuvent être utilisés dans différentes conditions pastorales pour mettre en évidence l'activité des troupeaux à partir de leurs localisations. La compréhension des comportements des troupeaux au cours des journées et entre les jours pourra être facilitée grâce à la connaissance experte des gestionnaires de troupeaux. Cette association de connaissance des indicateurs et de l'expertise des gestionnaires contribuera de manière efficace à la caractérisation du comportement spatial des troupeaux en zones méditerranéennes et montagneuses françaises. Elle permettra également d'analyser le lien entre les facteurs environnementaux et le comportement spatial des animaux en vue de produire des connaissances approfondies, qui pourront être mises à disposition des éleveurs.

Au travers d'analyses de corrélation et de régression, la hiérarchisation de l'influence des facteurs du milieu sur le comportement des troupeaux représenté par les indicateurs peut être faite. Il faudra cependant que les troupeaux soient placés dans des conditions pastorales où la variabilité des différents facteurs est susceptible d'impacter leur comportement spatial. Les trois sites expérimentaux associés à l'UMT Pasto (La Fage, Le Merle et Carmejane) ayant pour enjeux communs la génétique, l'alimentation (autonomie, place des parcours), et les technologies en élevage, peuvent contribuer à cette réalisation. Ils mobilisent différents types de parcours pour l'alimentation des troupeaux. A la Fage les parcours sont mobilisés (sur environ 9 mois sur 12) pour l'utilisation en plein air intégral. Au Merle la transhumance et les parcours de basse altitude sont mobilisés (environ 6 mois sur 12 sur parcours). Et à Carmejane les troupeaux exploitent les parcours boisés de moyenne altitude (plus de 4 mois sur 12 de parcours). En plus, ces sites expérimentaux peuvent faciliter la collecte d'informations permettant de renseigner les facteurs environnementaux et le comportement des animaux, et possèdent chacun une station météorologique. L'équipement des troupeaux de différentes races dans ces conditions variées des parcours peut permettre de rassembler des informations

pertinentes pour documenter le comportement spatial des troupeaux et hiérarchiser l'influence des facteurs environnementaux.

6.4 Valeur et portée des méthodes proposées

6.4.1 Qualité des données

La validité des méthodes abordées dans nos travaux repose sur la disponibilité et la qualité des données pour décrire le fonctionnement de l'écosystème agro-sylvo-pastoral. En l'absence des données, il est difficile voire impossible de caractériser certains composants de cet écosystème. Il en est de même pour la qualité de données qui doit être assez suffisante pour représenter l'information analysée. Les critères de qualité définis au chapitre 4 (précision, échelle de représentation et sémantique, et granularité temporelle) permettent d'évaluer la capacité des données à fournir des informations pertinentes.

Pour caractériser les milieux pastoraux, les différentes données identifiées présentent des qualités variables. La faible précision des classes pastorales (pelouses et landes) dans l'occupation des sols OSO (Inglada et al., 2017) peut être problématique pour l'interprétation des résultats finaux sur le comportement spatial des troupeaux. Les approches employées pour rehausser la discrimination de ces classes pastorales relèvent les difficultés d'obtenir de meilleurs résultats à cause de la forte hétérogénéité de ces classes. L'intégration des indices de végétation tels que le NDVI et l'OSAVI, peut dans ce cas permettre de caractériser l'état de la végétation et améliorer l'information sur la végétation. En effet, en l'absence des données de terrain, le NDVI est une données intéressante pouvant décrire l'état de la végétation (Hill et al., 2004; Júnior et al., 2019; Serrano et al., 2021). L'utilisation du NDVI et de l'OSAVI en conditions pastorales permettra également de suivre l'évolution de la végétation au fil du pâturage des troupeaux. Il suffira pour ces indicateurs de définir la résolution temporelle suffisante qui correspond aux périodes pour lesquelles on vise d'étudier le changement dans le cadre de l'utilisation des parcours par les troupeaux. Pour notre expérimentation, nous avons définis une fréquence temporelle de 10 jours pour analyser le changement dans l'utilisation des parcs par phase de 10 jours. La modélisation de la pente a été basée sur une donnée de meilleure qualité (modèle numérique de terrain de 5 m de maille) pour des applications pastorales pouvant offrir une opportunité pour analyser l'influence de la topographie sur le comportement des troupeaux.

Le premier critère de qualité pour les données de localisation GPS est la précision qui peut être influencée par les facteurs météorologiques ou les conditions du milieu (Buerkert and Schlecht, 2009). Dans le contexte des parcours, la prise en compte de ces facteurs pouvant impacter la qualité des données doit être assurée. Des prétraitements doivent être faits pour identifier les données aberrantes afin de mitiger leur influence sur les résultats produits. La fréquence d'acquisition des données de localisation représente un autre aspect

important dont dépend la qualité des indicateurs produits. Dans les fermes expérimentales de l'UMT Pasto, les GPS déployés ont été calibrés pour enregistrer la localisation toutes les 5 minutes sur base des recommandations des travaux récents (Grisot et al., 2022; Candela, 2022). En élevage privé par contre, la situation est plus hétérogène au sein d'une même exploitation et entre les exploitations. En plus, les éleveurs sont souvent limités à l'utilisation d'un seul GPS pour équiper leurs troupeaux produisant ainsi une faible quantité de données pouvant rendre difficile la production d'indicateurs. Pour économiser la batterie, les éleveurs peuvent réduire la fréquence d'acquisition, et cela limite la possibilité de calculer tous les indicateurs. Dans des situations où les éleveurs équipés de GPS veulent connaître la position instantanée du troupeau il y a un risque de perdre des données et d'avoir ainsi des "trous" dans la séquence de points et un nombre de points insuffisant pour le calcul des indicateurs. La fréquence d'acquisition impacte également les résultats de certains indicateurs tels que la distance parcourue par les troupeaux et la vitesse de déplacement (cf. Figure 6.1).

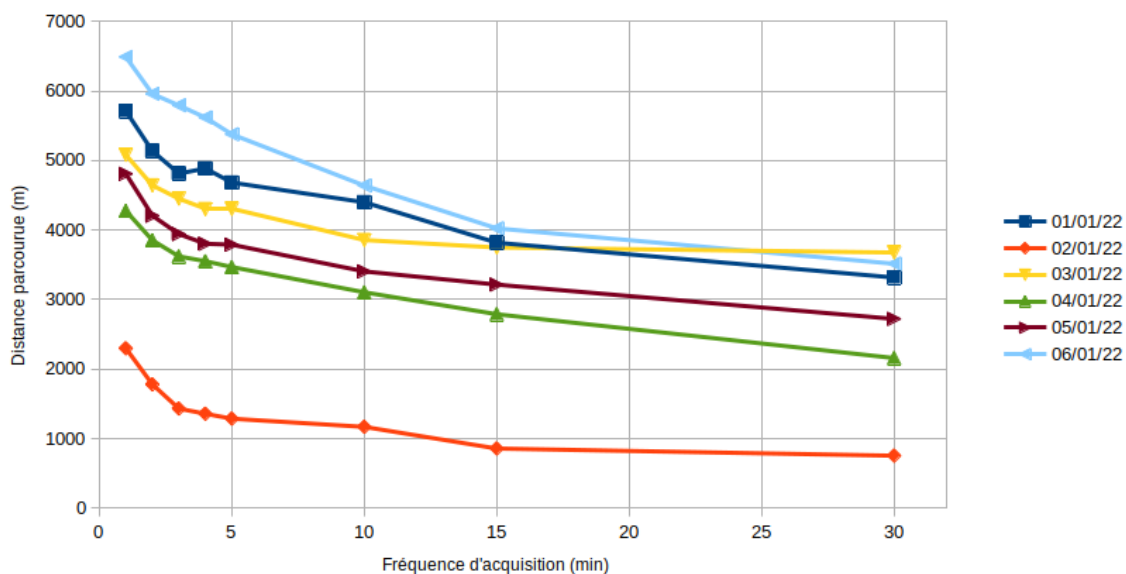


Figure 6.1 – Variation des distances parcourues par un troupeau en fonction de la fréquence d'acquisition. Données issues d'un élevage privé. Source : Candela (2022)

6.4.2 Validité des analyses et des indicateurs d'utilisation de l'espace par les troupeaux

Les résultats produits peuvent être impactés par différentes causes notamment les points aberrants, les points manquants et l'échelle temporelle d'analyse.

Les points aberrants représentent l'imprécision inhérente à la mesure des coordonnées GPS, compte tenu des conditions particulières du matériel ou du milieu. Ils s'éloignent des motifs de points et peuvent constituer un biais dans le calcul d'indicateurs tels que l'étendue explorée qui sera plus grande que normalement. Plus il y a de points aberrants plus les valeurs d'indicateurs produits seront impactées. Les points manquants représentent les trous dans

l'acquisition des données GPS réduisant la quantité de données pour une période donnée. Avec 5 minutes de fréquence d'acquisition, environ 288 points devraient être enregistrés dans 24h. En raison des conditions météorologiques particulières ou de passage des troupeaux dans une zones moins ouverte les GPS perdent intermittemment la liaison satellitaire et capturent les données de manière irrégulière. Pour notre expérimentation à la Fage nous avons eu de cas avec moins de 180 points sur 288. Les points aberrants peuvent être détectés et éliminés à partir des méthodes non paramétriques telle que les diagrammes en boîte. D'autres méthodes paramétriques telles que le Z-score (Aggarwal et al., 2019) ou DBSCAN (Thang and Kim, 2011; Celik et al., 2011) peuvent aussi être employées. Face aux points manquants la seule option consiste à définir un seuil de proportion des données en dessous de laquelle les résultats d'analyse risquent d'être impactés. Ce critère sera à définir en fonction de chaque condition d'élevage. On retiendra ainsi les données supérieur ou égal à ce seuil pour effectuer les analyses. Pour le cas de nos analyses à la Fage, pour chaque GPS, nous avons retenu un minimum de 220 points par jours sur le maximum de 288 points (soit environ $\geq 76\%$).

Le choix de l'échelle temporelle devrait être basé sur le but visé par l'analyse en fonction des possibilités offertes par les données. L'utilisation des parcours par les troupeaux peut être analysée à l'échelle horaire ou de créneau horaire, journalière ou de phase de pâturage intégrant plusieurs jours. L'échelle horaire ou de créneaux horaire peut permettre de suivre la variation de l'utilisation des parcours au cours de la journée. Elle peut donc permettre d'observer les différents modes d'utilisation journalière des parcours par les troupeaux pouvant être associés aux comportements des troupeaux (pâturage, repos, rumination). L'échelle journalière permettra quant à elle d'observer la variation au cours des jours de l'utilisation spatiale des troupeaux sans fournir les détails concernant leurs comportements journaliers. A l'échelle de phase de pâturage (plusieurs jours) il s'agira d'avoir une vue globale de l'utilisation des parcours par les troupeaux (par exemple la répartition spatiale des trajectoires parcourues par les troupeaux ou de l'utilisation de l'espace pâturés).

Les proportions de points aberrants et manquants peuvent considérablement influencer le choix de la fréquence temporelle d'analyse. De grandes proportions de points manquants et aberrants peuvent rendre difficile la caractérisation de l'activité journalière des troupeaux. Également, les grandes périodes (échelle de plusieurs heures, 5 à 10 heures par exemple) pourraient noyer certains comportements des troupeaux. Les analyses réalisées dans le cadre des données de la Fage ont permis de maintenir une échelle temporelle de 3 heures pour un minimum de 220 points par jour (avec 11 et 12 agnelles équipées sur 20 respectivement pour les deux groupes) après élimination des points aberrants. Cette échelle temporelle nous a permis de définir 8 créneaux journaliers de 3 heures et d'observer les variations dans le mode d'utilisation des parc par les deux groupes d'agnelles. La situation est différente en élevage privé où le plus souvent un seul animal est équipé de GPS. En effet, dans le cas idéal 288 points sont collectés par GPS par jour pour représenter les localisations d'un troupeau, soit 12 points par heures ou 36 points pour 3 heures. Cependant, à cause de points aberrants et manquants, il sera souvent difficile d'obtenir le maximum de points. En plus, la densité de

points par échelle temporelle sera irrégulièrement répartie et peut impacter significativement les résultats obtenus et par conséquent l'interprétation des indicateurs. Ces aspects montrent la nécessité dans les élevages privés mais aussi expérimentaux, de vérifier non seulement la proportion de points aberrants et manquants, mais aussi la régularité de distribution de la densité des points pour les différentes périodes définies.

Des indicateurs de distance parcourue, de superficie explorée, de vitesse instantanée et de densité ont été calculés à l'échelle journalière pour les élevages privés dans le cadre des travaux du projet P@stora_All réalisés par Candela (2022) (cf. Annexes A.28 à A.35 pour les indicateurs de distance et de surface). Ces indicateurs montrent la variation sur une période de pâturage de l'utilisation de l'espace en comparant les jours. Pour l'indicateur de densité, différentes échelles temporelles (au jour, à la semaine et au mois) ont permis de visualiser la densité de présence des troupeaux dans les zones explorées. Par rapport à nos travaux, l'échelle temporelle utilisée dans le cadre des élevages privés ne permettent pas d'identifier les variations de comportements des troupeaux au cours du jours. L'indicateur de vitesse instantanée pourrait permettre d'y parvenir mais trouvent une limite causée par l'imprécision des GPS (notamment les point aberrants) pouvant biaiser les résultats.

Les deux groupes d'indicateurs issus de notre expérimentation et des travaux réalisés par Candela (2022) sur les élevages pastoraux privés peuvent se compléter pour apporter plus d'informations sur l'utilisation des parcours par les troupeaux. Les indicateurs produits à partir des données expérimentales peuvent permettre de produire des connaissances sur le comportement spatial des troupeaux sur parcours et ses variations durant une saison de pâturage donnée. Les indicateurs produits à partir des données privées renseignent sur les variations globales de l'utilisation des parcours entre les jours. Les deux groupes d'indicateurs combinés présentent un intérêt opérationnel pour les élevages pastoraux afin de suivre à différentes échelles temporelle (au cours des jours et entre les jours) l'activité des troupeaux.

6.4.3 Application des approches développées à d'autres contextes d'élevage pastoral

L'application de nos approches développées à d'autres contextes d'élevage pastoral sera conditionnées par la disponibilité des données et l'engagement des parties prenantes, principalement les éleveurs pastoraux dans le processus de conception du SI. Malgré les contextes pédo-climatiques et écologiques différents, les systèmes pastoraux des zones méditerranéennes d'Europe ont en commun l'hétérogénéité de de la topographie, une présence de ligneux et des milieux embroussaillés, et une diversité d'animaux avec une grande part de petits ruminants (Bernués et al., 2011; Aubron et al., 2014; Jouven, 2018). Avec la mobilité des troupeaux qui y est traditionnellement pratiquée et la diversité d'espèces locales associée à chaque région spécifique (de Rancourt et al., 2006), les enjeux de développement d'outil de prise de décision sont rapprochés de ceux de notre cas d'étude. Les approches présentées dans cette thèse peuvent également être reprises dans d'autres zones méditerranéennes au

sud de l'Europe. Les enjeux au sud de la Méditerranée se présentent différemment, étant donné les sécheresses, l'intensification de la production et la mise en culture de terres de parcours au détriment de l'élevage (Bourbouze, 2018). Cela nécessite pour la mise en place d'un outil de prise de décision, l'application de nouvelles approches adaptées au contexte de ces régions.

Dans les systèmes d'élevage pastoraux des pays sahéliens ou le pastoralisme revêt d'une importance capitale pour l'économie des pays et la sécurité alimentaire des populations (Cedea and Csao/Ocde, 2008; De Haan et al., 1999), les défis à relever sont encore plus complexes qu'en régions méditerranéennes. Dans ces pays la disponibilité en quantité, en qualité et en distribution spatiale des fourrages est fortement impactée par les effets du climat (notamment la sécheresse) et la nature des sols (Hiernaux et al., 2014). Les ressources pastorales évoluent de manière complexes (Dardel et al., 2014) et affectent variablement les exploitations pastorales (Ibra et al., 2012). La gestion du pastoralisme dans ces pays se conçoit plutôt à l'échelle nationale ou régionale. Un SI adapté aux enjeux dans ces régions devrait se baser sur la caractérisation des évolutions à long terme et la prévention des crises que subissent les systèmes pastoraux sahéliens (Ibra et al., 2012). Ainsi, les besoins en outil de prise de décision dans ces régions se situent à différentes échelle spatiale et temporelle que celles abordées dans notre SI. Le SI SIPSA (De Cao et al., 2008; Ibra et al., 2012) développé dans ces régions permet d'intégrer les enjeux spécifiques à ces régions. Ce SI apporte des informations sur l'état et l'évolution des zones pastorales pour permettre l'interprétation, faciliter la décision et améliorer le ciblage des interventions.

6.5 Perspectives de recherche

Les études menées dans le cadre de notre thèse ouvrent différentes perspectives pouvant être pris en compte dans le cadre de prochains travaux, notamment pour l'enrichissement des fonctionnalités du SI. Premièrement, les approches d'analyses développées à partir des données expérimentales, et celles utilisées pour les élevages privés, pourraient être testées en conditions opérationnelles à partir des données enregistrées sur différents types de troupeaux pâturant différents milieux (tels que ceux du domaine expérimental du Merle). Nous estimons que des analyses dans de telles conditions permettront de produire d'autres connaissances complémentaires et de mettre en avant les effets dus aux facteurs du milieu. L'extraction de séquences fréquentes non utilisée dans nos analyses pourront produire des résultats valorisables dans ces conditions. Deuxièmement, la définition des moyens permettant d'intégrer au lac de données spatiales, les outils développés dans le cadre des projets précédents (Mil'ouv, Pastothèque, Rami Pastoral et StratPast, Pâtur'Ajuste) permettra de palier aux difficultés de caractérisation des milieux pastoraux et d'améliorer l'analyse de l'interaction entre les facteurs du milieu et le comportement spatial des troupeaux. La pente souvent associée à l'embroussaillage constituant un enjeu important pour le pastoralisme, une étude focalisée sur la compréhension de la variation du comportement des troupeaux dans des milieux

aux pentes variées pourrait contribuer à la documentation de l'élevage pastoral.

6.6 Conclusion générale

Pour atteindre les trois objectifs poursuivis dans cette thèse, trois principaux résultats ont été produits : i) la définition d'un lac de données spatiales en adéquation avec les besoins des éleveurs et l'hétérogénéité des systèmes agro-sylvo-pastoraux, ii) la définition des critères de qualité des données spatiales pour mieux décrire les milieux pastoraux et le comportement spatial des troupeaux, et iii) la définition des approches d'analyse, et en particulier l'analyse de séquences fréquentes et l'extraction de motifs de points pour décrire le comportement spatial des troupeaux. Le SI a été défini comme un lac de données spatiales basé sur Hadoop qui propose grâce à son système de fichier basé sur HDFS, un stockage distribué et la réalisation d'analyse parallélisé. Ce SI offre une flexibilité dans un contexte de diversité des systèmes agro-sylvo-pastoraux où une diversité d'analyses peut être réalisée en fonction des conditions spécifiques à chaque exploitation. Les critères de qualité des données spatiales retenus sont : i) 10m de résolution spatiale pour bien décrire l'hétérogénéité du milieu, ii) une précision suffisamment élevée de l'occupation des sols en vue de mieux caractériser les classes d'intérêt pastoral, et iii) une fréquence d'acquisition de 5 minutes pour les données de localisation des troupeaux afin de faciliter l'extraction de connaissances à différentes échelles temporelles (horaire, journalière, phase de pâturage). Cependant, les pré-traitements réalisés pour améliorer les classes pastorales (pelouses, landes et forêts pâturées), n'ont pas permis d'obtenir de meilleurs résultats à cause de leur forte hétérogénéité. Les données complémentaires telles que le NDVI et l'OSAVI peuvent être mobilisées pour décrire l'état et la dynamiques des végétations. L'analyse de motifs de points appliquée aux données GPS a permis de produire des indicateurs d'étendue explorée, de dispersion et de densité de présence au sein du parc. Malheureusement, les conditions expérimentales et notamment la taille des parcs (environ 1ha), n'ont pas favorisé une variabilité des facteurs du milieu et ainsi n'ont pas permis de caractériser leur influence sur le comportement des troupeaux.

Les travaux réalisés contribuent à deux niveaux. Premièrement à la proposition d'un SI adapté aux contextes et aux données hétérogènes de l'élevage pastoral en zones méditerranéennes et montagneuses françaises. Deuxièmement à la définition des méthodes d'analyse adaptées aux données hétérogènes du pastoralisme afin de proposer des indicateurs d'aide à la décision pour la conduite du pâturage sur parcours. Ces travaux présentent un intérêt pour la documentation de l'élevage pastoral en zones méditerranéennes et montagneuses françaises en vue de répondre aux enjeux de production et de préservation des écosystèmes dans le cadre du pastoralisme. A court terme, le SI proposé est un outil d'aide à la décision qui peut permettre de faire un suivi au quotidien de l'utilisation de l'espace par les troupeaux et d'ajuster la conduite du pâturage en fonction des variations des conditions du milieu. L'alimentation du SI en données et en résultats d'analyses permettra d'accroître les cas des exemples sur le comportement spatial des troupeaux et donc d'enrichir à long termes, la

documentation de l'utilisation des parcours par les troupeaux.

Le SI proposé présente une complémentarité avec d'autres outils d'aide à la gestion du pâturage développés dans le cadre des travaux de l'UMT Pasto en apportant une contribution sur la caractérisation de l'utilisation spatio-temporelles des parcours par les troupeaux. En revanche le SI peut tirer profit de la caractérisation des ressources pastorales qui constitue le points forts de ces outils. L'application des approches développées durant nos travaux peut être valable dans les régions méditerranéennes européennes étant caractérisées par des enjeux et des opportunités à peu près similaires.

Annexes

Annexe A

Figures

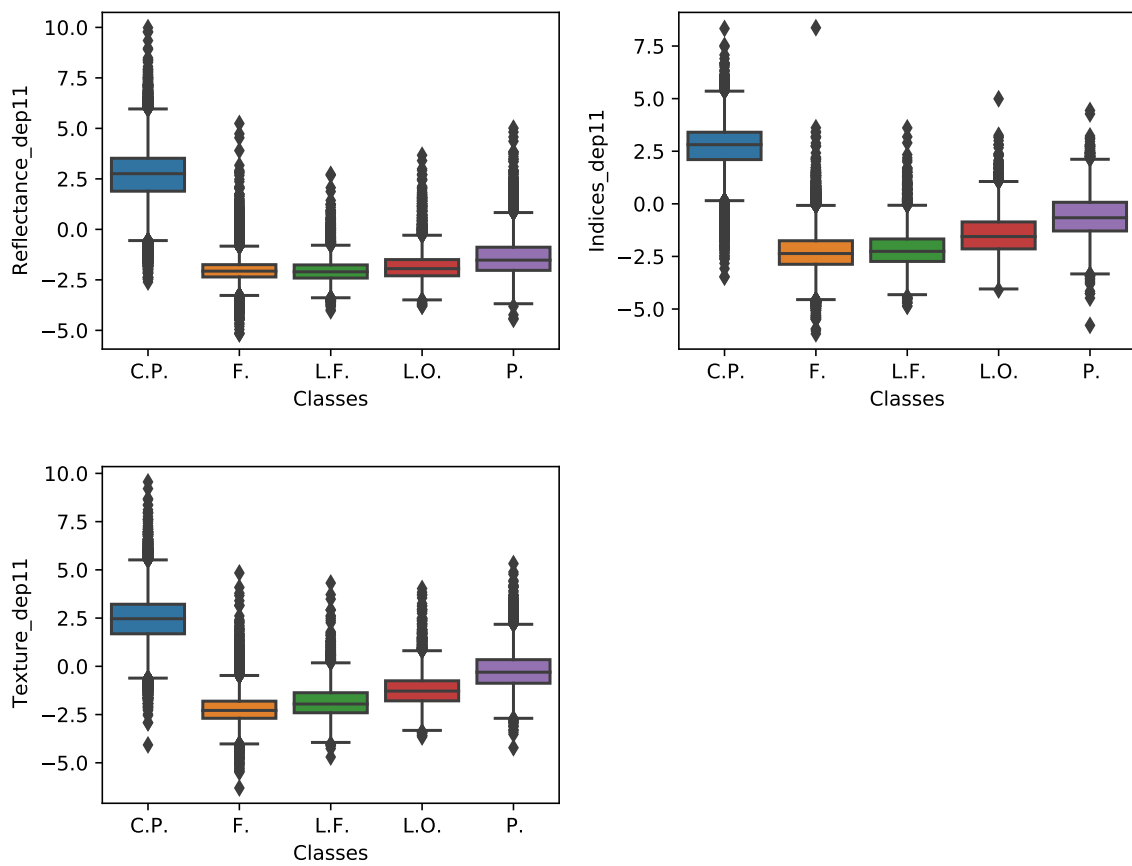


Figure A.1 – Distribution des variables issues des premiers composants de l'ADL appliquée pour analyser la séparabilité des 5 classes à partir des variables de reflectance, spectrales et texturales, pour le département de l'Aude (11).

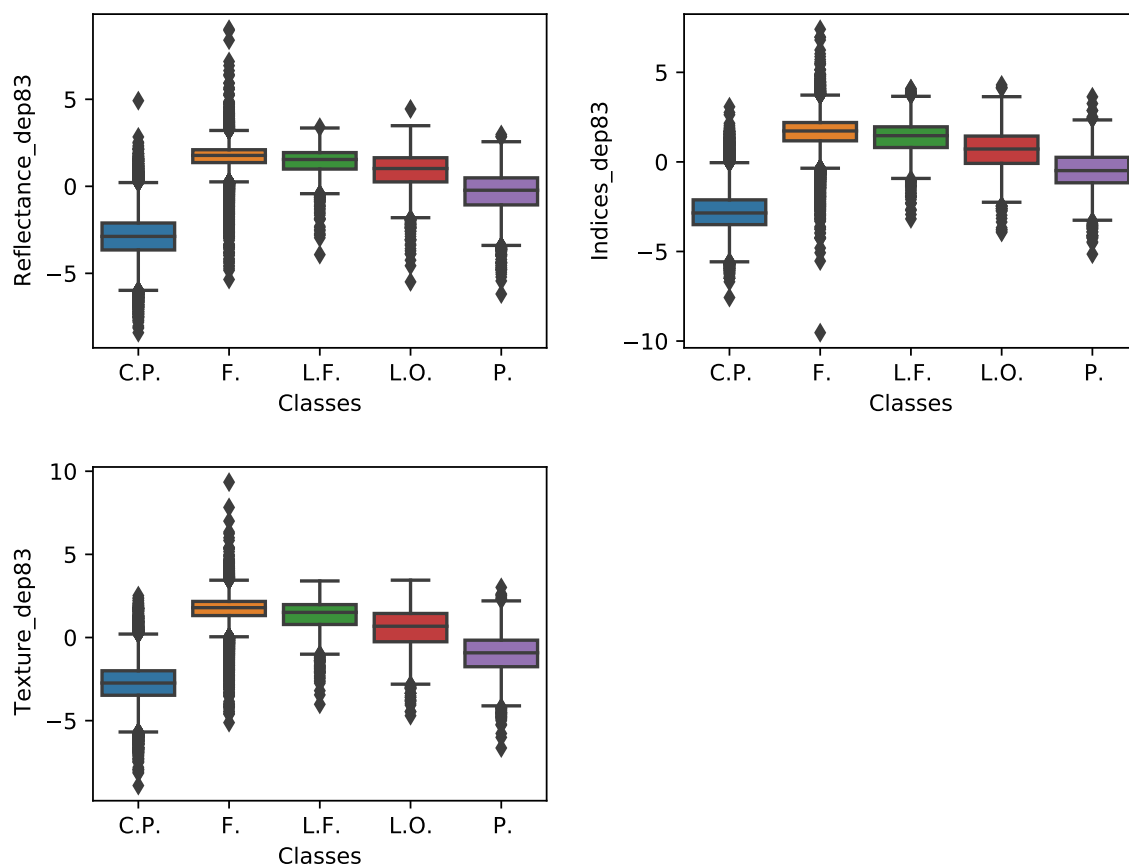


Figure A.2 – Distribution des variables issues des premiers composants de l'ADL appliquée pour analyser la séparabilité des 5 classes à partir des variables de reflectance, spectrales et texturales, pour le département du Var (83).

Résultats de la régression linéaire entre indicateurs spatiaux et variables du milieu

Indicateurs d'étendue explorée

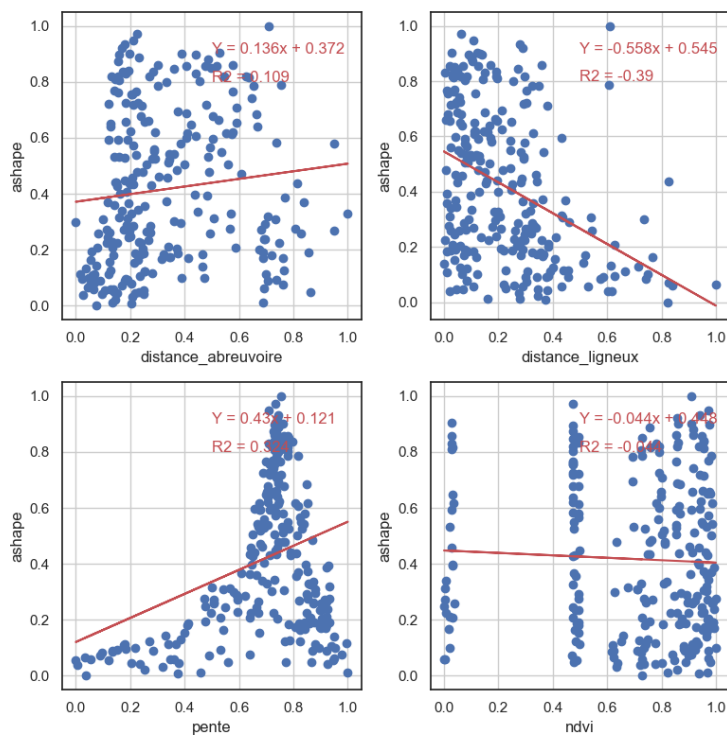


Figure A.3 – Résultats de la régression linéaire entre l'indicateur d'étendue explorée (Alpha shape) et les variables spatiales du milieu pour les agnelles MAT pâturant l'été.

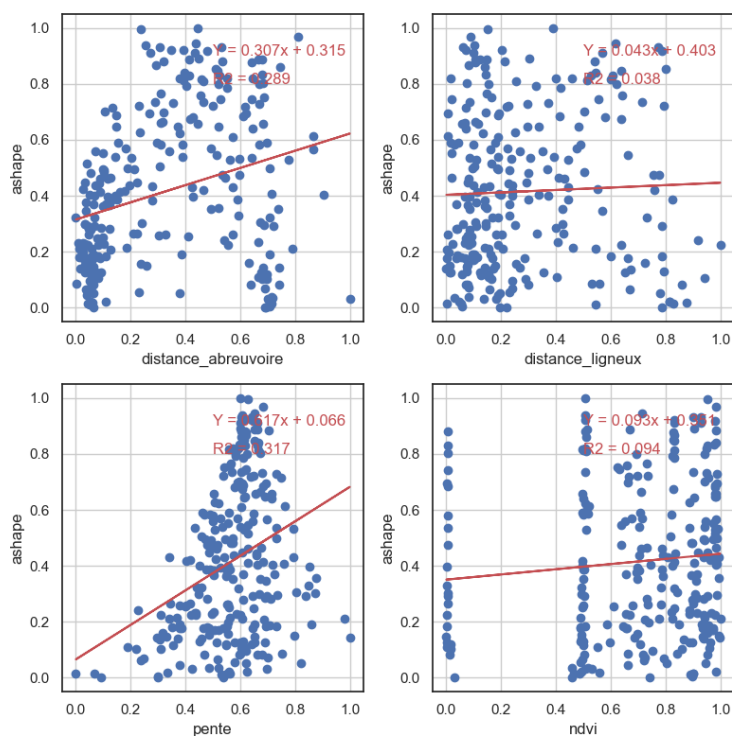


Figure A.4 – Résultats de la régression linéaire entre l'indicateur d'étendue explorée(Alpha shape) et les variables spatiales du milieu pour les agnelles ART pâturant l'été.

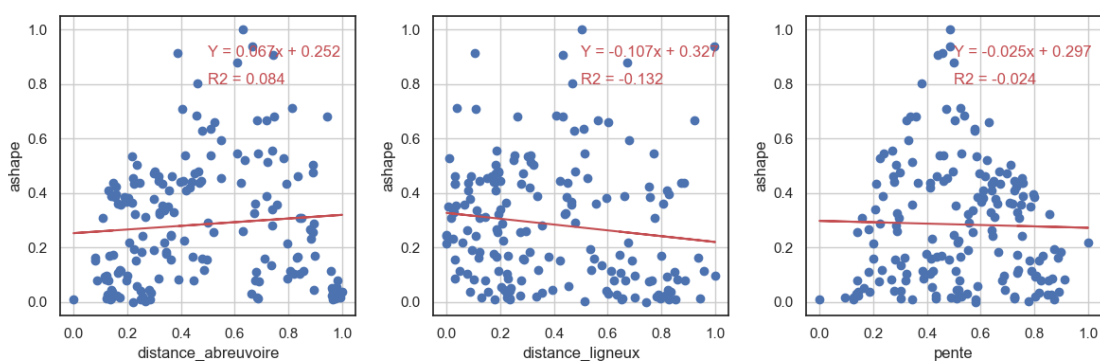


Figure A.5 – Résultats de la régression linéaire entre l'indicateur d'étendue explorée(Alpha shape) et les variables spatiales du milieu pour les agnelles MAT pâturant l'automne.

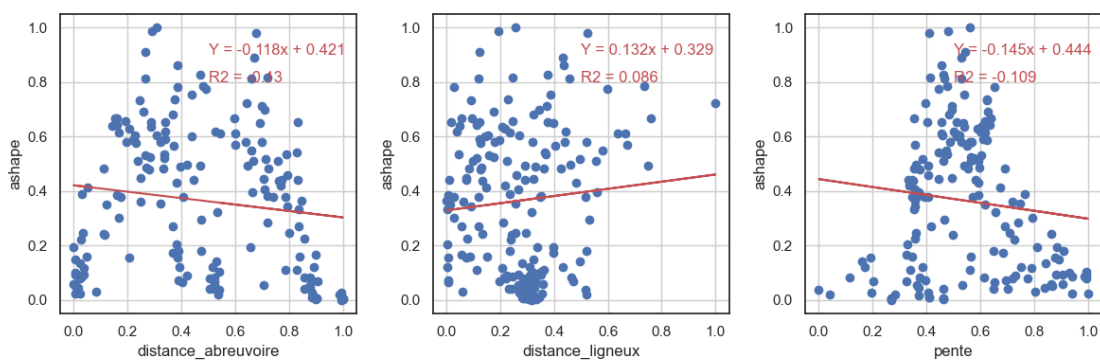


Figure A.6 – Résultats de la régression linéaire entre l'indicateur d'étendue explorée (Alpha shape) et les variables spatiales du milieu pour les agnelles ART pâturant l'automne.

Indicateurs de dispersion

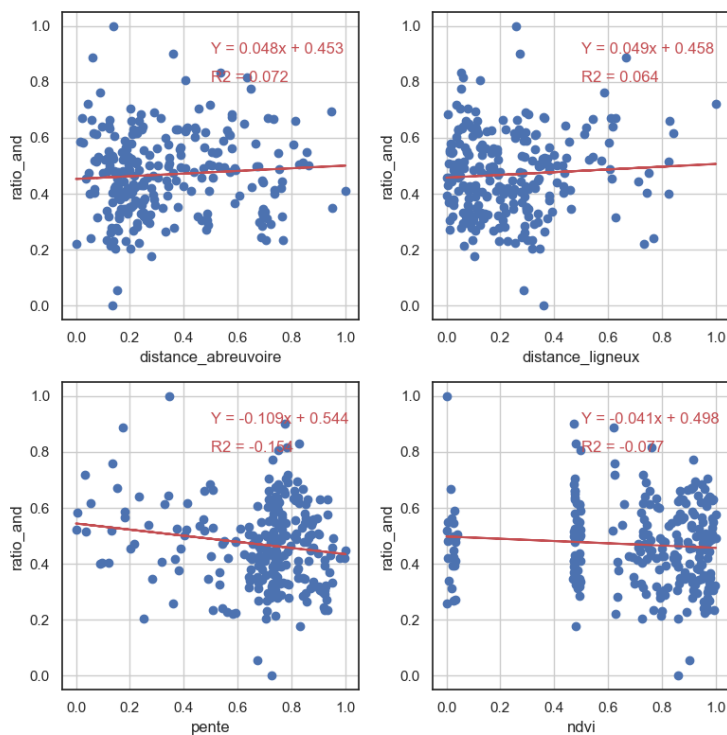


Figure A.7 – Résultats de la régression linéaire entre l'indicateur de dispersion (AND) et les variables spatiales du milieu pour les agnelles MAT pâturant l'été.

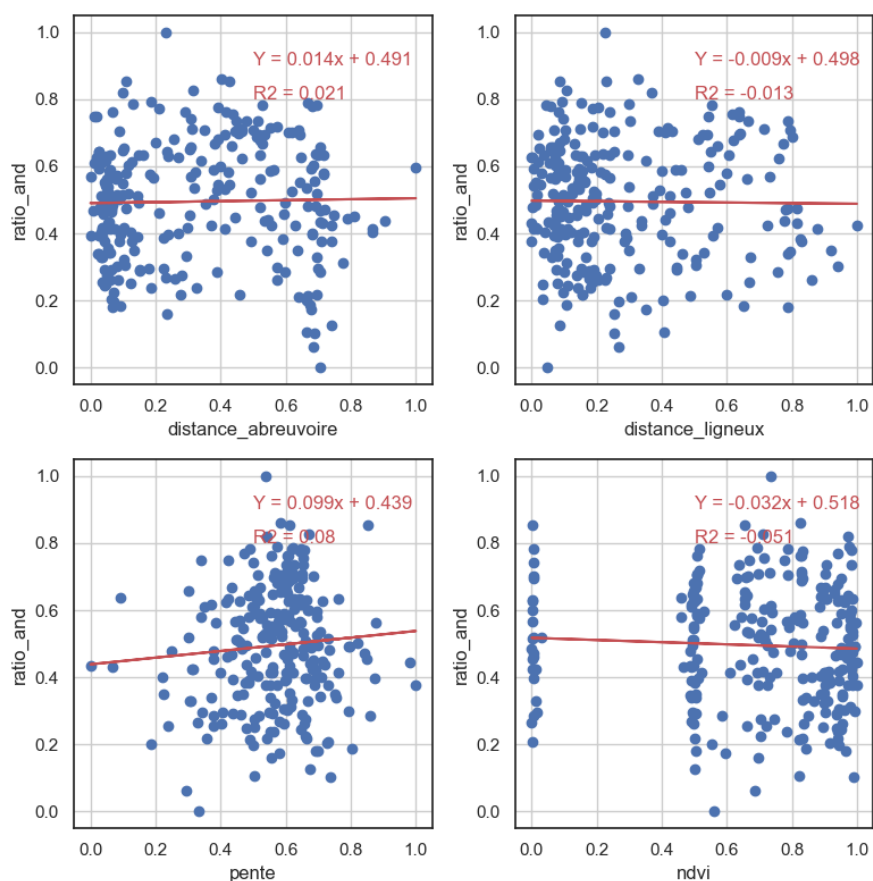


Figure A.8 – Résultats de la régression linéaire entre l'indicateur de dispersion (AND) et les variables spatiales du milieu pour les agnelles ART pâturant l'été.

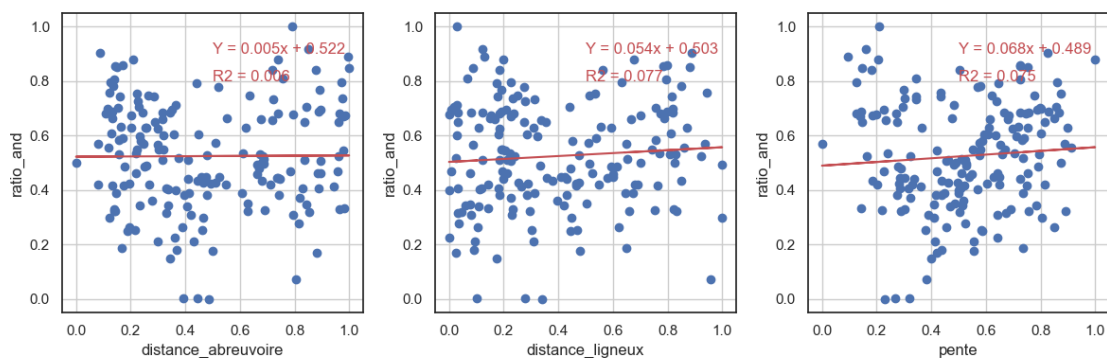


Figure A.9 – Résultats de la régression linéaire entre l'indicateur de dispersion (AND) et les variables spatiales du milieu pour les agnelles MAT pâturant l'automne.

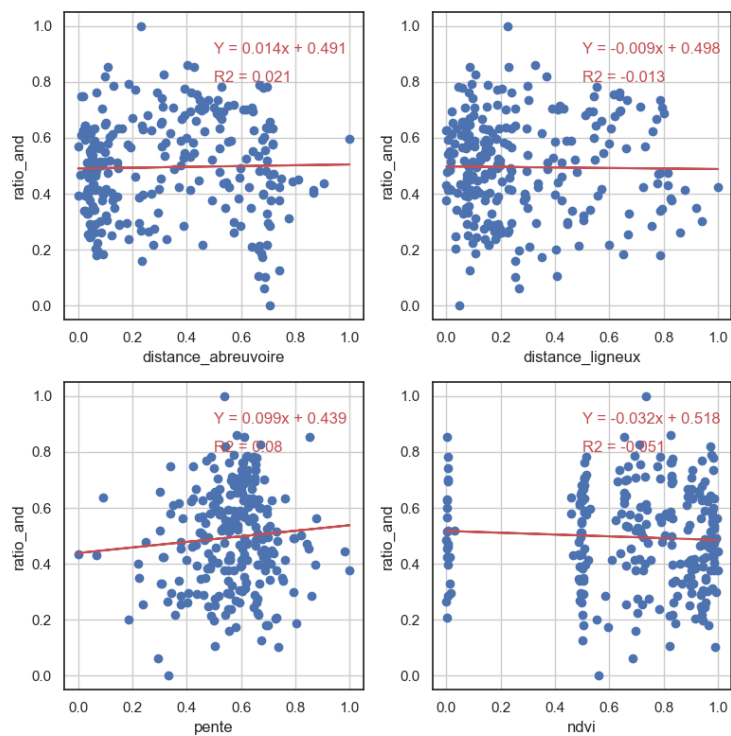


Figure A.10 – Résultats de la régression linéaire entre l'indicateur de dispersion (AND) et les variables spatiales du milieu pour les agnelles ART pâturant l'automne.

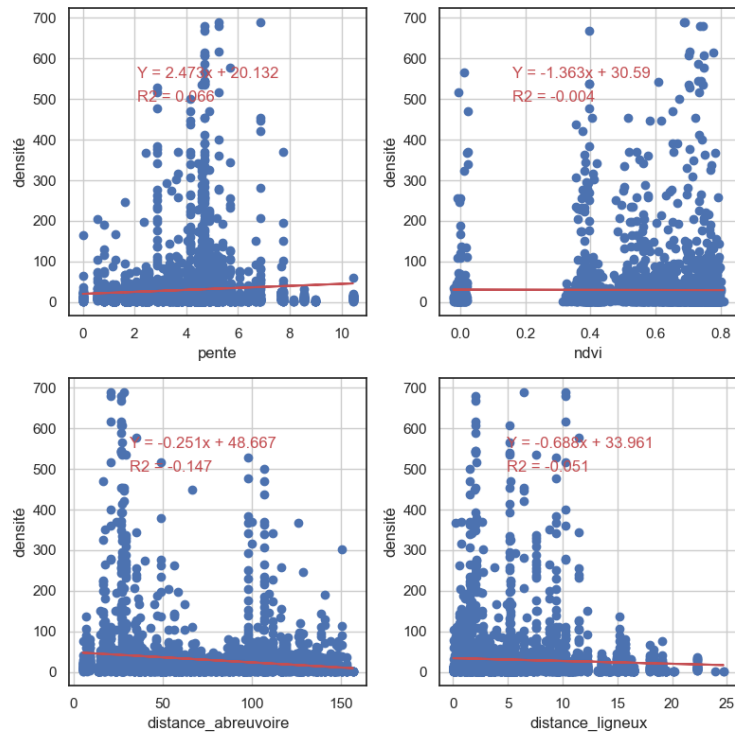
Indicateur de densité

Figure A.11 – Résultats de la régression linéaire entre l'indicateur de densité et les variables spatiales du milieu pour les agnelles MAT pâturant l'été.

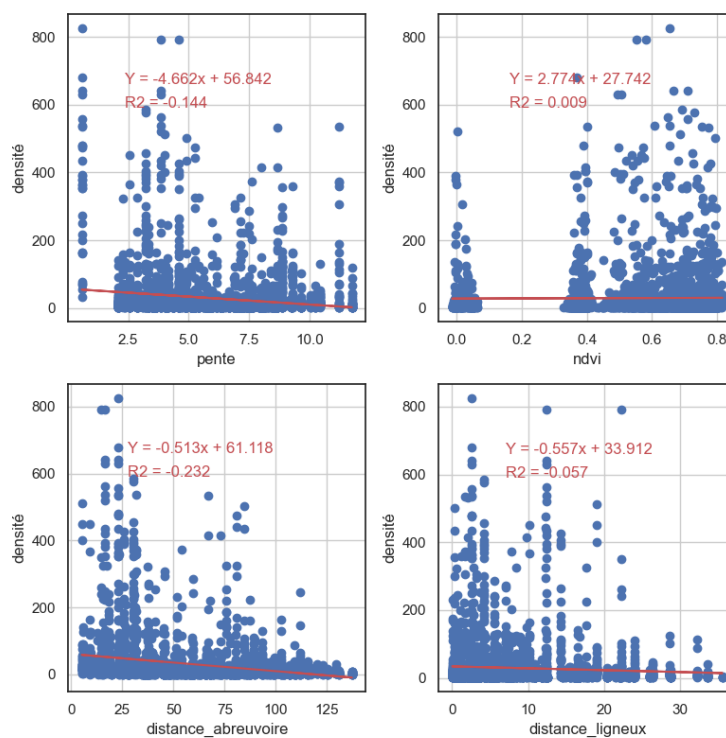


Figure A.12 – Résultats de la régression linéaire entre l'indicateur de densité et les variables spatiaux du milieu pour les agnelles ART pâturant l'été.

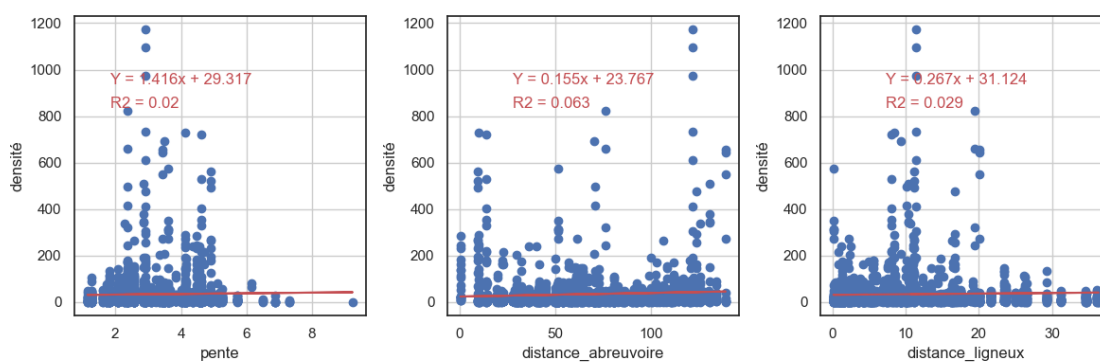


Figure A.13 – Résultats de la régression linéaire entre l'indicateur de densité et les variables spatiaux du milieu pour les agnelles MAT pâturant l'automne.

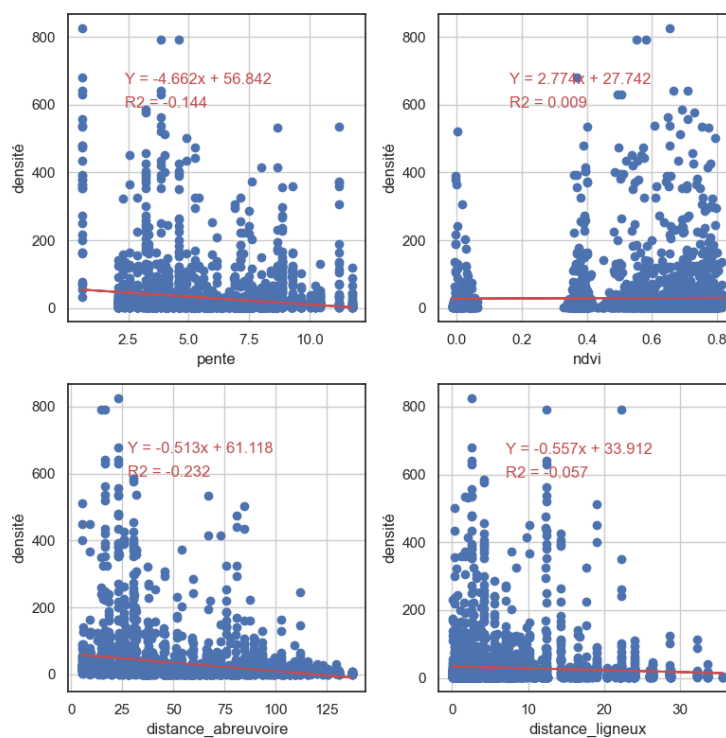


Figure A.14 – Résultats de la régression linéaire entre l'indicateur de densité et les variables spatiales du milieu pour les agnelles ART pâturant l'automne.

Résultats de régression linéaire entre indicateurs spatiaux et paramètres météorologiques

Indicateurs d'étendue explorée

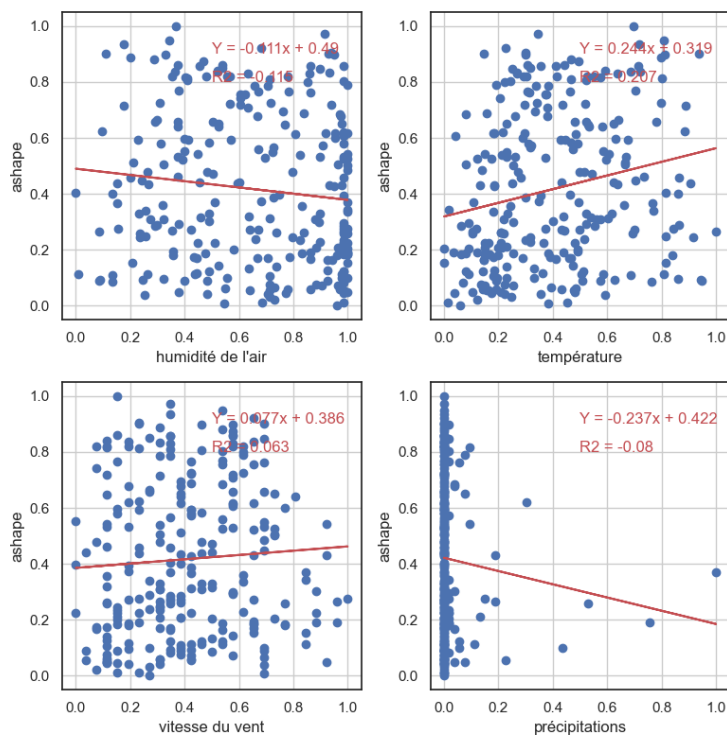


Figure A.15 – Résultats de la régression linéaire entre l'indicateur d'étendue explorée (Alpha shape) et les paramètres météorologiques pour les agnelles MAT pâturant l'été.

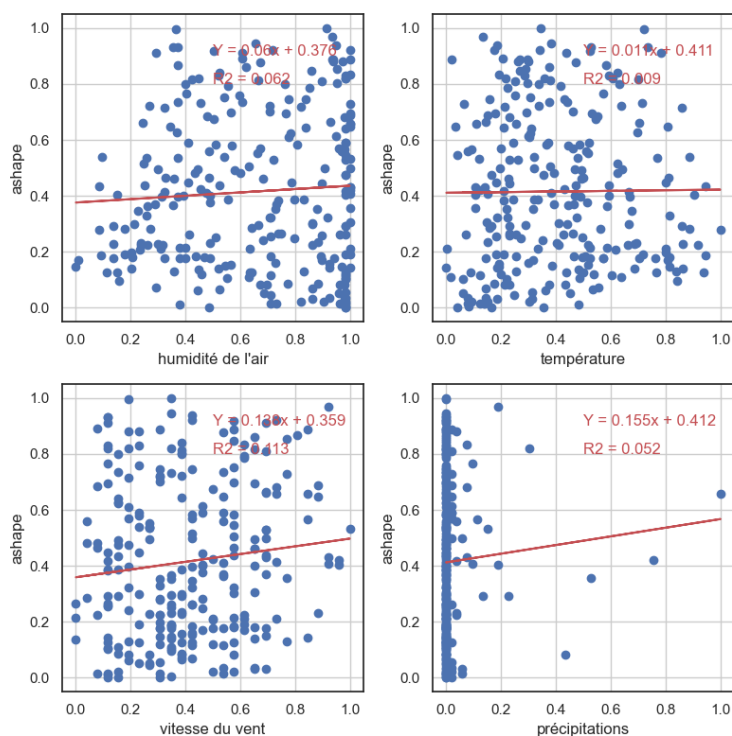


Figure A.16 – Résultats de la régression linéaire entre l'indicateur d'étendue explorée(Alpha shape) et les paramètres météorologiques pour les agnelles ART pâturant l'été.

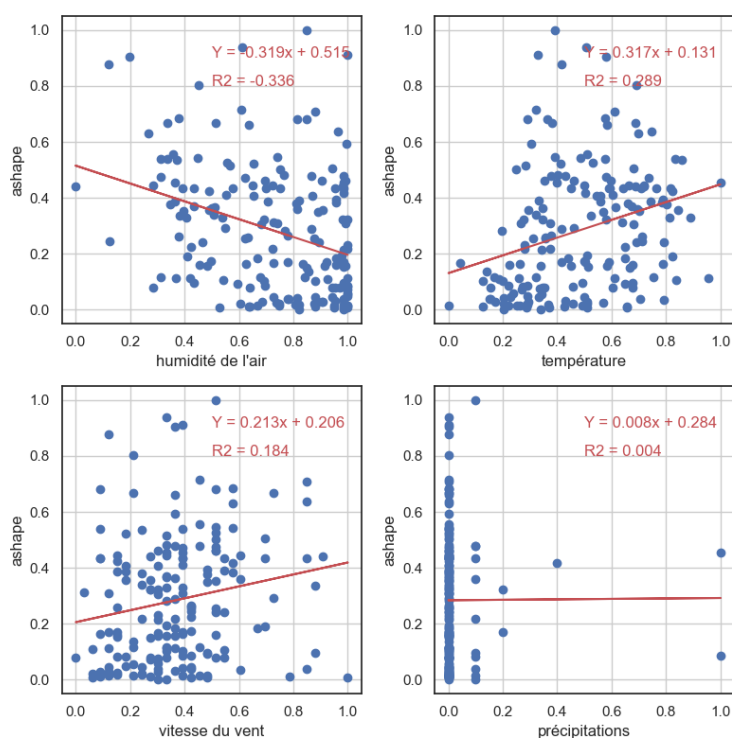


Figure A.17 – Résultats de la régression linéaire entre l'indicateur d'étendue explorée(Alpha shape) et les paramètres météorologiques pour les agnelles MAT pâturant l'automne.

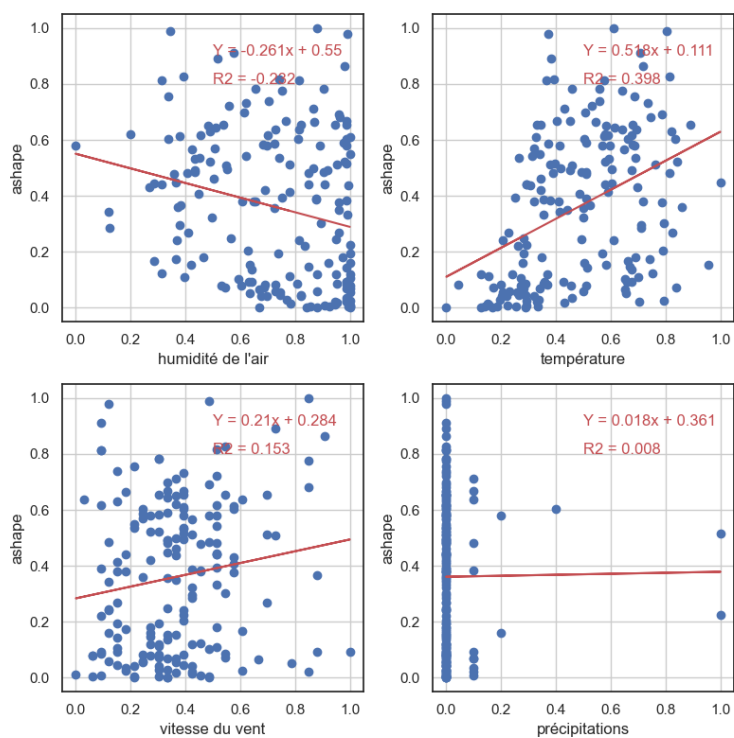


Figure A.18 – Résultats de la régression linéaire entre l'indicateur d'étendue explorée (Alpha shape) et les paramètres météorologiques pour les agnelles ART pâturant l'automne.

Indicateurs de dispersion

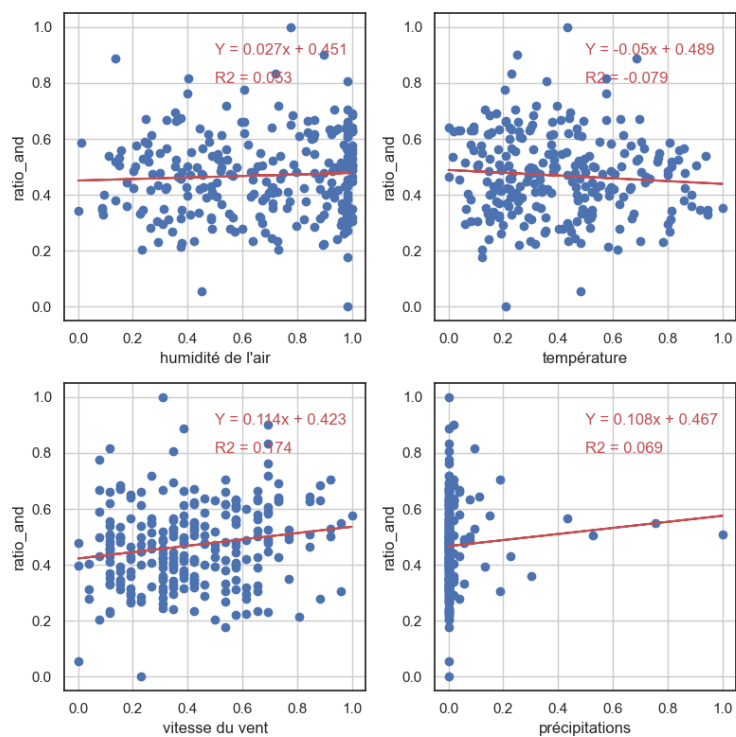


Figure A.19 – Résultats de la régression linéaire entre l'indicateur de dispersion (AND) et les paramètres météorologiques pour les agnelles MAT pâturant l'été.

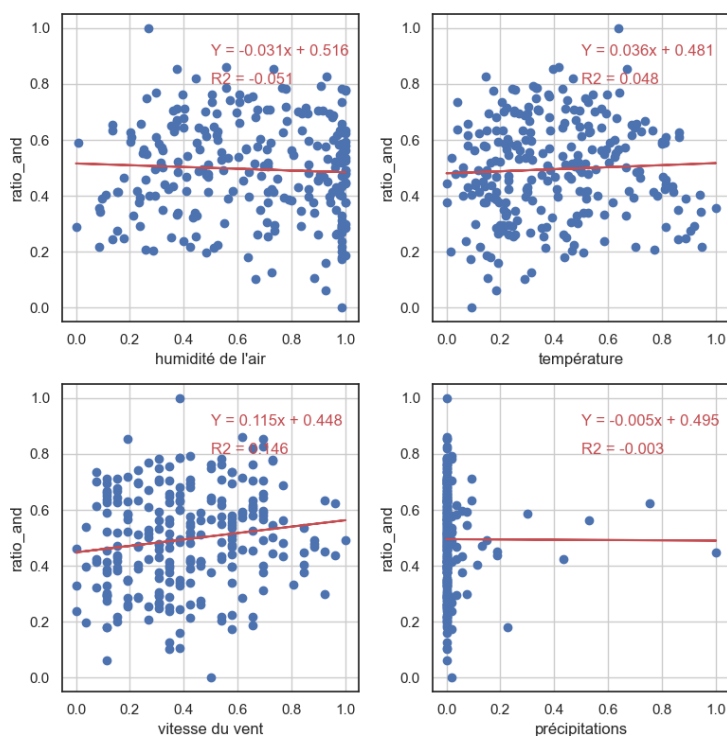


Figure A.20 – Résultats de la régression linéaire entre l'indicateur de dispersion (AND) et les paramètres météorologiques pour les agnelles ART pâturant l'été.

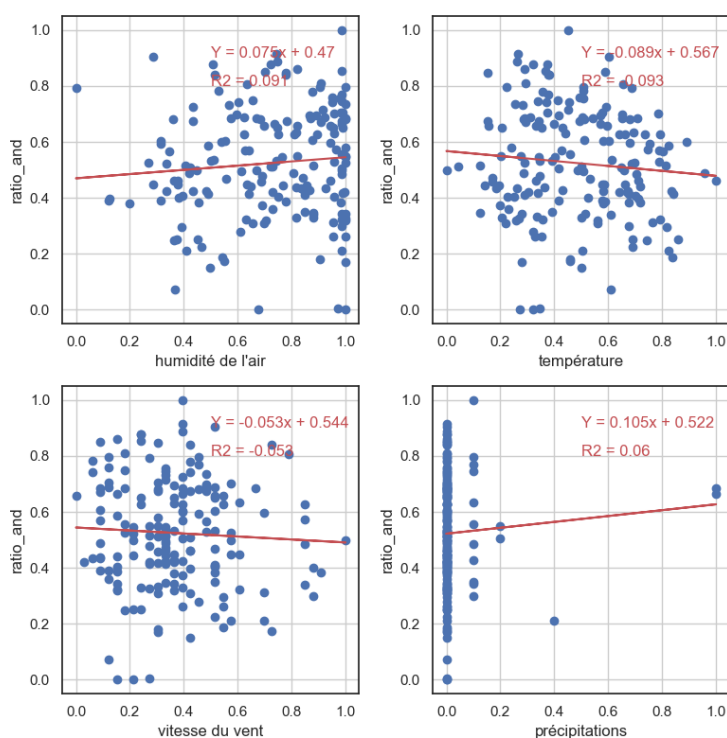


Figure A.21 – Résultats de la régression linéaire entre l'indicateur de dispersion (AND) et les paramètres météorologiques pour les agnelles MAT pâturant l'automne.

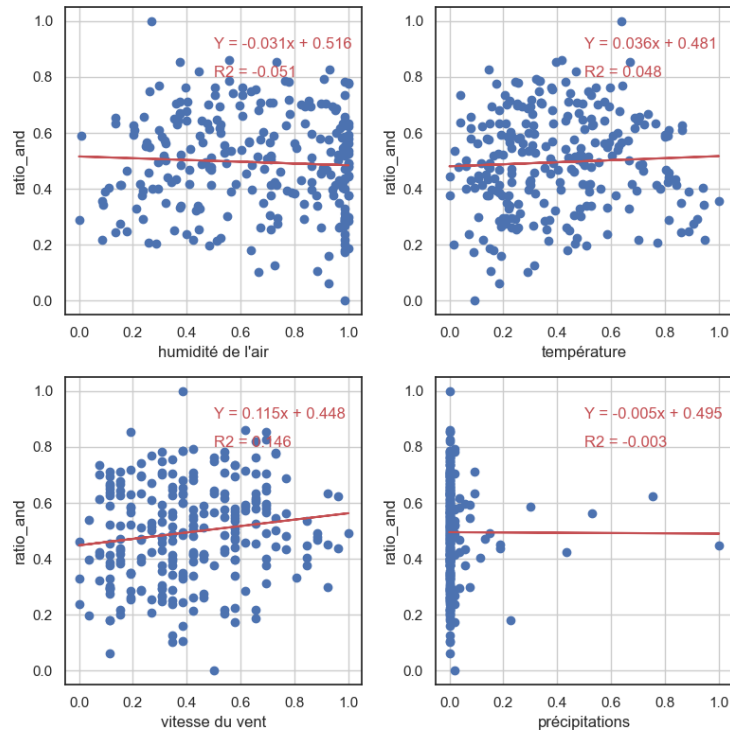


Figure A.22 – Résultats de la régression linéaire entre l'indicateur de dispersion (AND) et les paramètres météorologiques pour les agnelles ART pâturant l'automne.

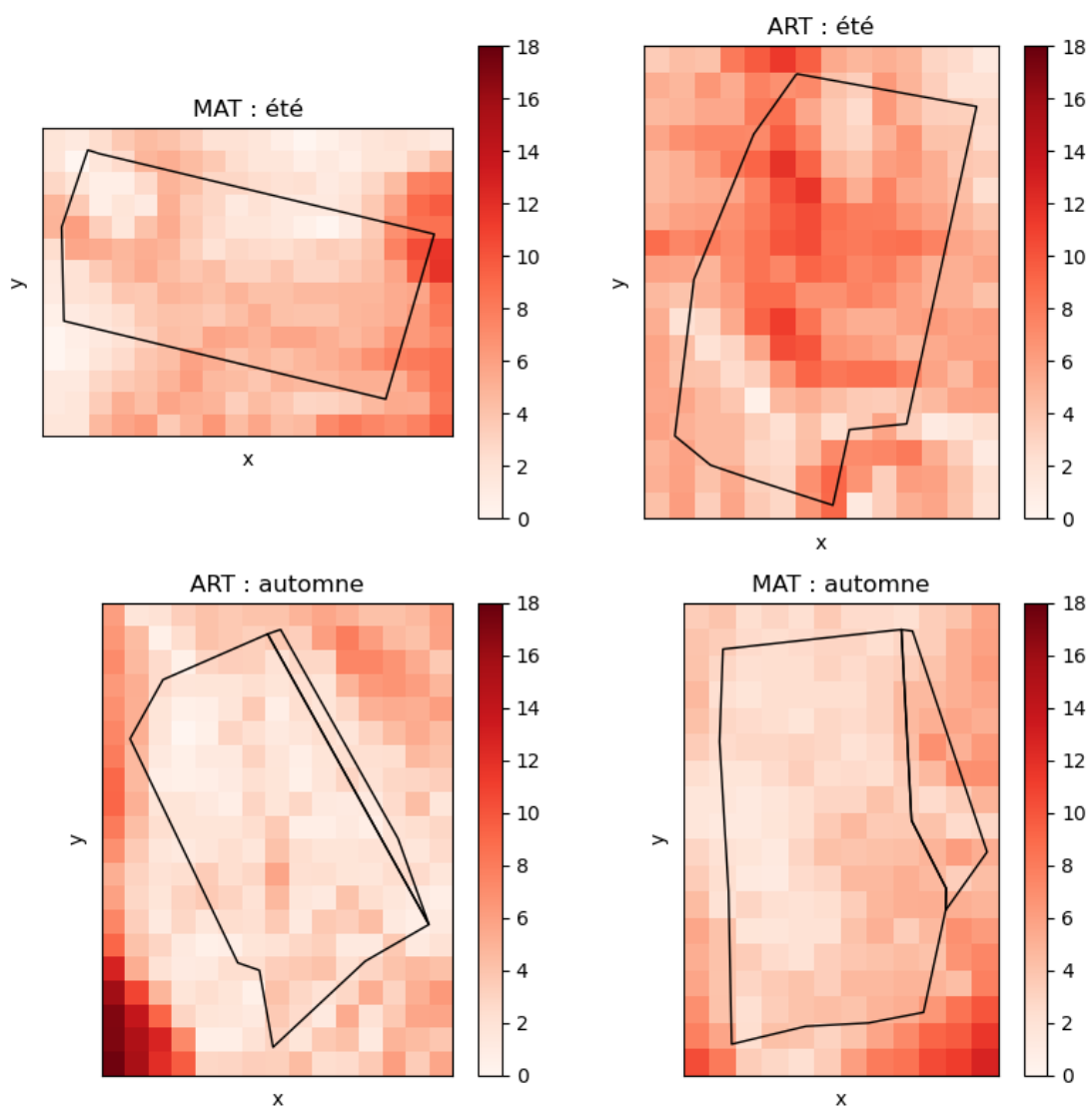
Cartes des variables spatiales utilisées

Figure A.23 – Cartes de pente des parcs utilisés par les agnelles. En haut à gauche : parc utilisé en été par les agnelles allaitées naturellement. En haut à droite : parc utilisé en été par les agnelles allaitées artificiellement. En bas à gauche : parc utilisé en été par les agnelles allaitées artificiellement. En bas à droite : parc utilisé en été par les agnelles allaitées naturellement

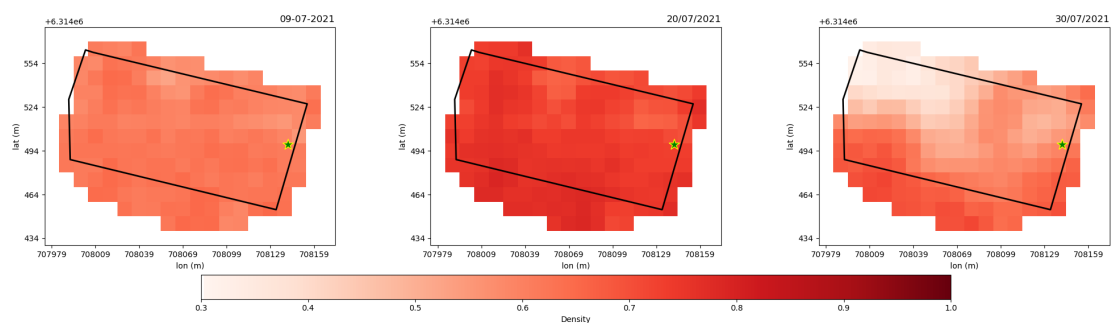


Figure A.24 – Cartes de NDVI de trois dates correspondant au début des phases de pâturage d'été au sein du parc utilisé par les agnelles naturellement allaitées. L'étoile verte bordée de jaune représente la localisation de l'auge.

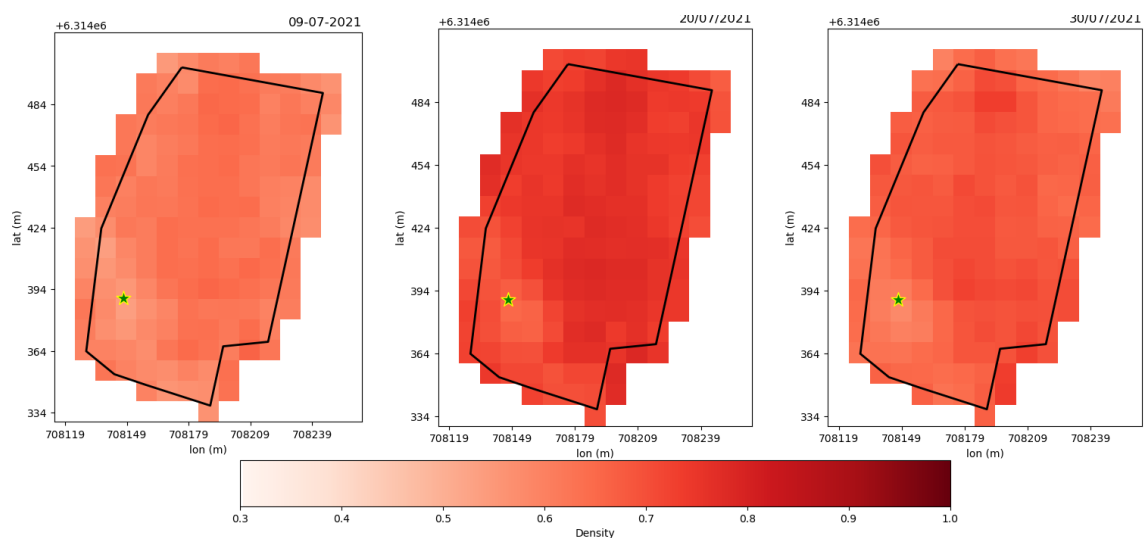


Figure A.25 – Cartes de NDVI de trois dates correspondant au début des phases de pâturage d'été au sein du parc utilisé par les agnelles artificiellement allaitées. L'étoile verte bordée de jaune représente la localisation de l'auge.

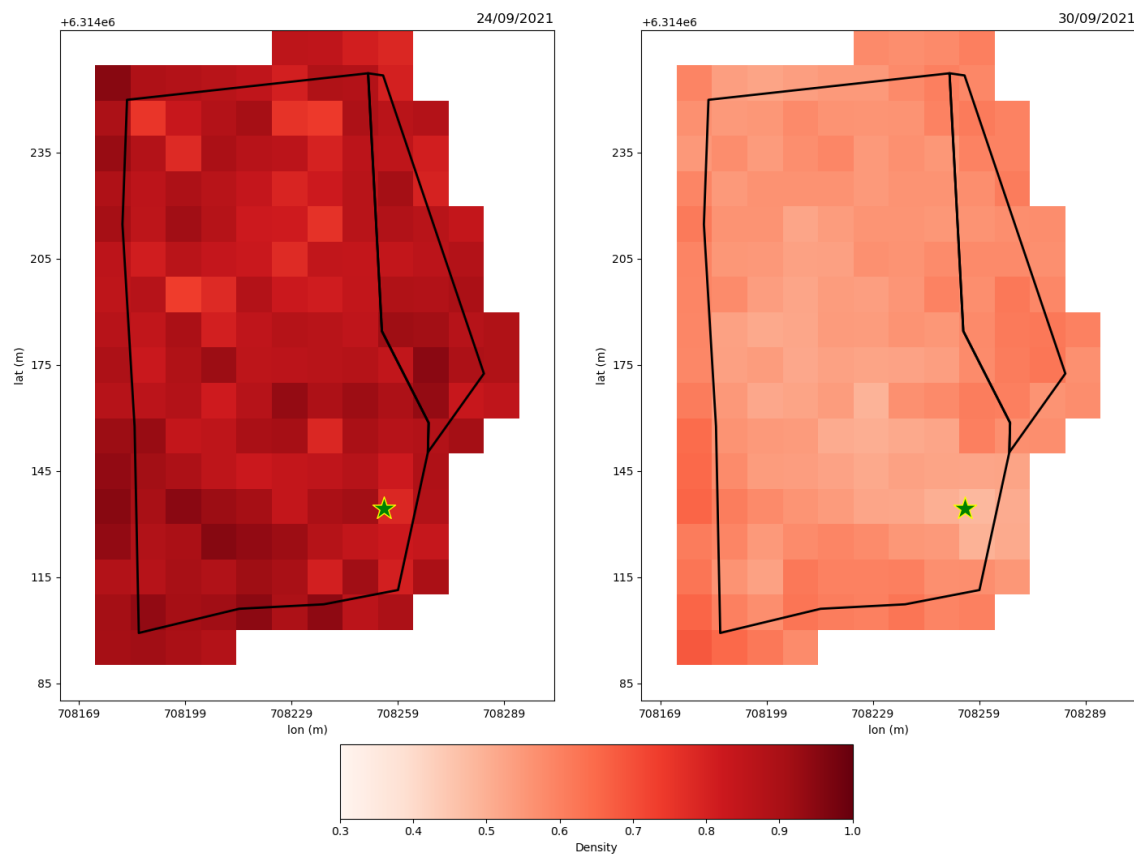


Figure A.26 – Cartes de NDVI de deux dates correspondant au début des phases de pâturage d'automne au sein du parc utilisé par les agnelles naturellement allaitées. L'étoile verte bordée de jaune représente la localisation de l'auge.

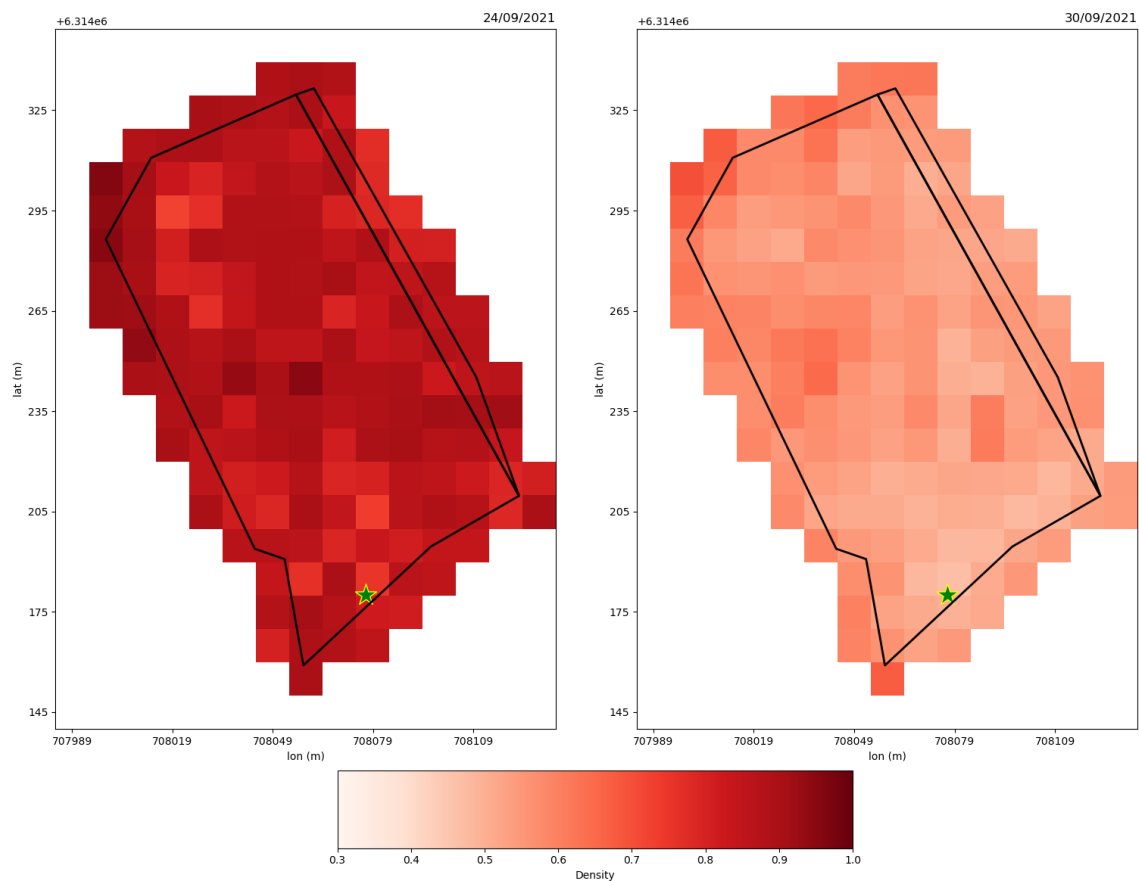


Figure A.27 – Cartes de NDVI de deux dates correspondant au début des phases de pâturage d'automne au sein du parc utilisé par les agnelles artificiellement allaitées. L'étoile verte bordée de jaune représente la localisation de l'auge.

Indicateurs journaliers produits à partir des données du domaine expérimental du Merle et d'une exploitation privée de chèvres

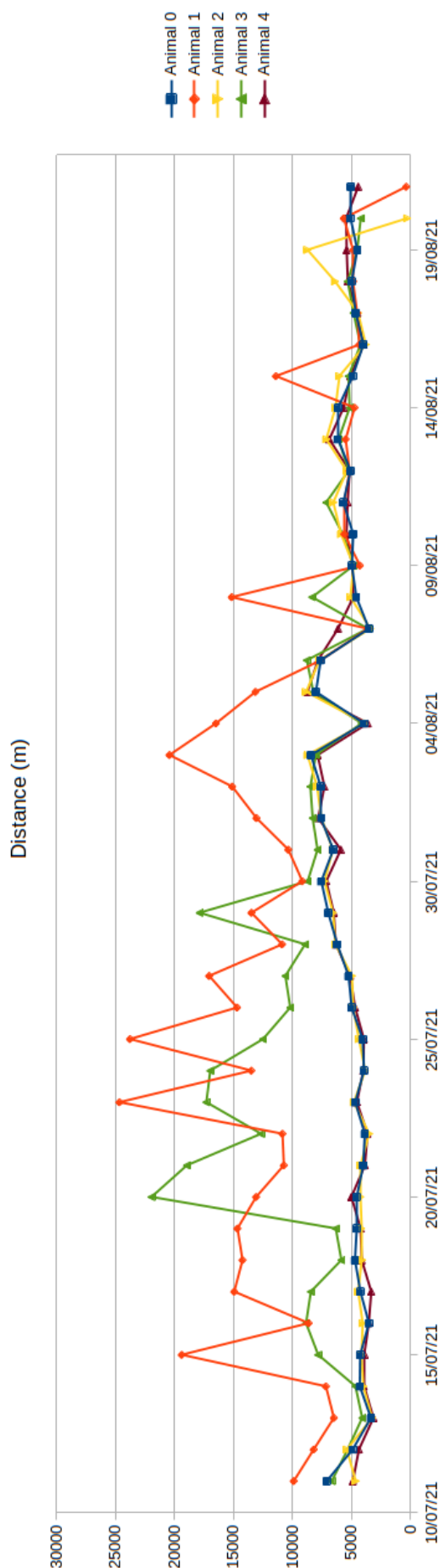


Figure A.28 – Distances journalières parcourues par cinq brebis équipées au domaine du Merle et d'une exploitation privée de chèvres

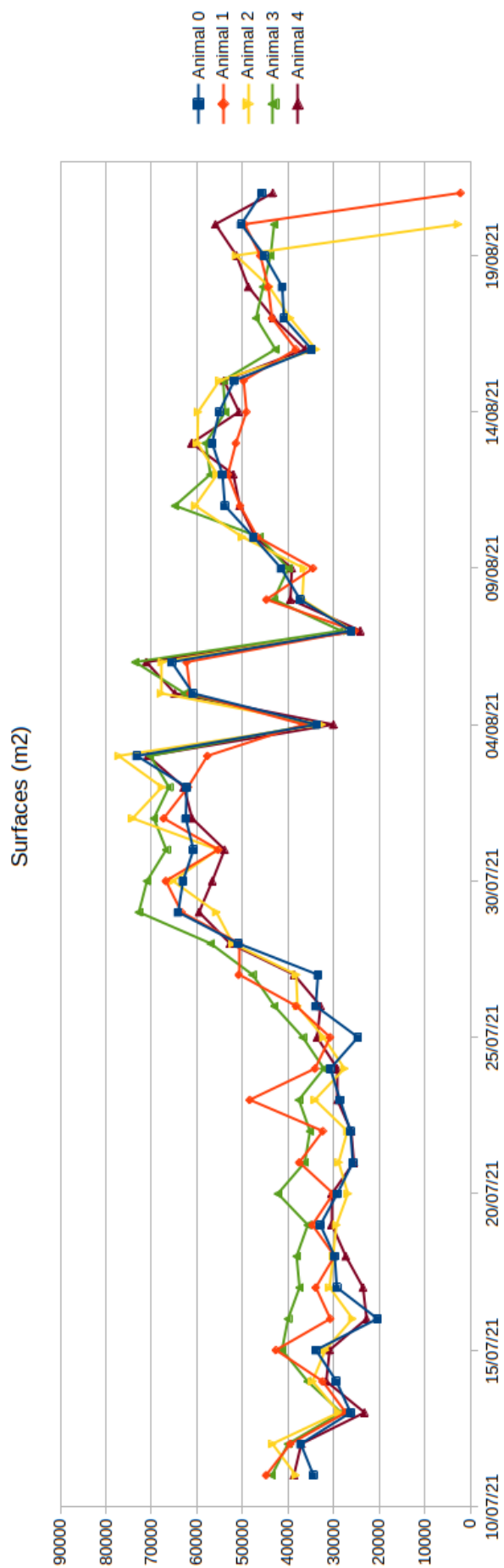


Figure A.29 – Surfaces journalières parcourues par cinq brebis équipées au domaine du Merle. Cet indicateur est beaucoup moins sensible aux erreurs de mesure du GPS.

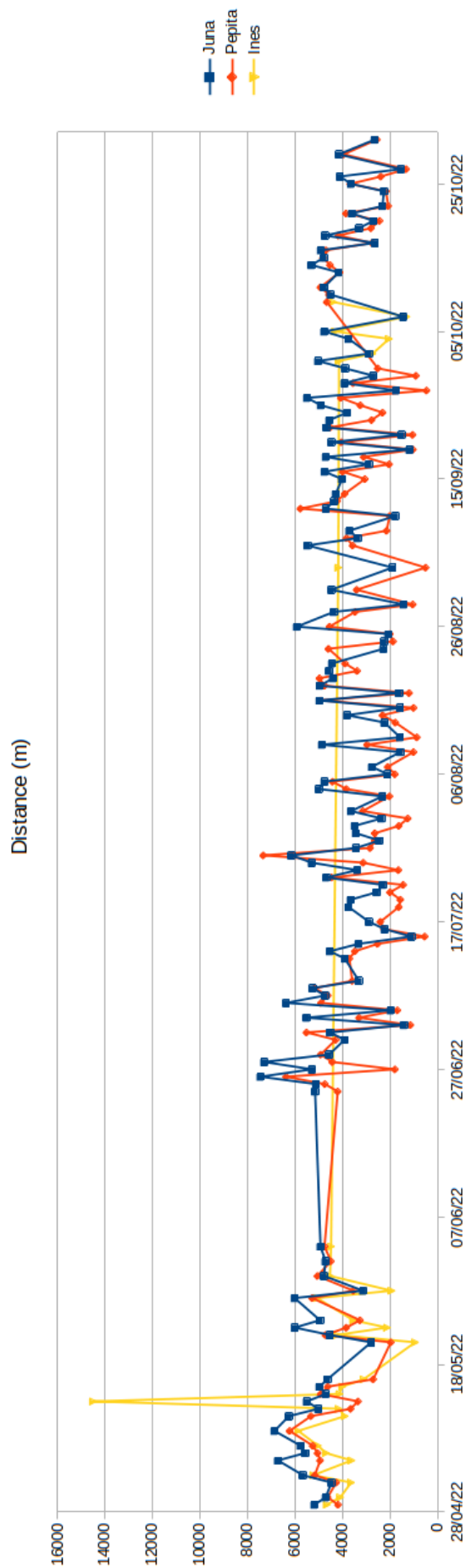


Figure A.30 – Distances journalières parcourues par 3 chèvres équipées dans une exploitation privée d'avril à octobre 2022.

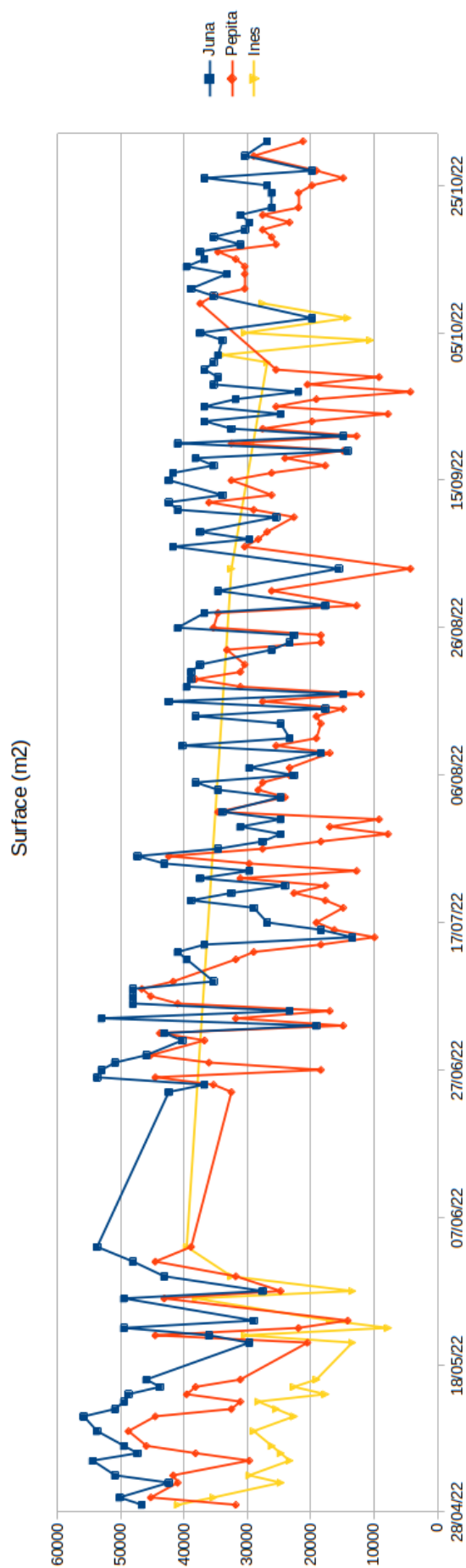


Figure A.31 – Surfaces journalières parcourues par 3 chèvres équipées dans une exploitation privée d'avril à octobre 2022.

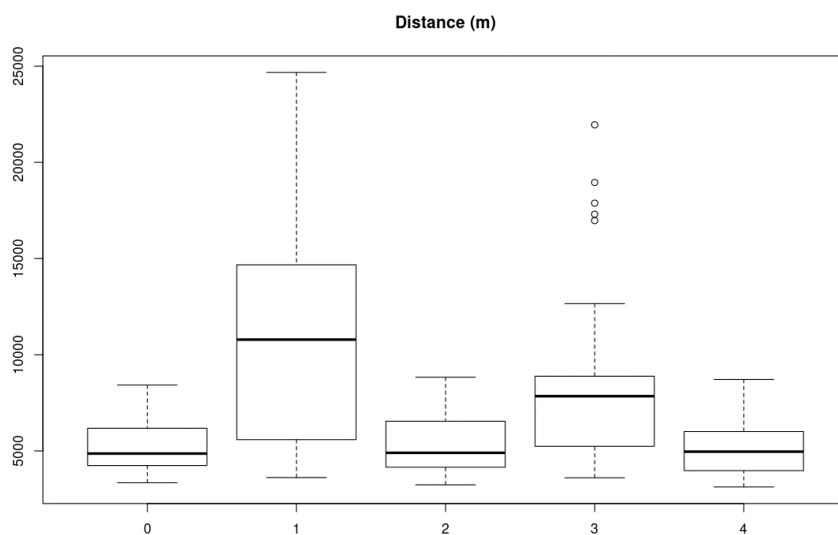


Figure A.32 – Distribution des distances journalières parcourues par cinq brebis équipées au domaine du Merle.

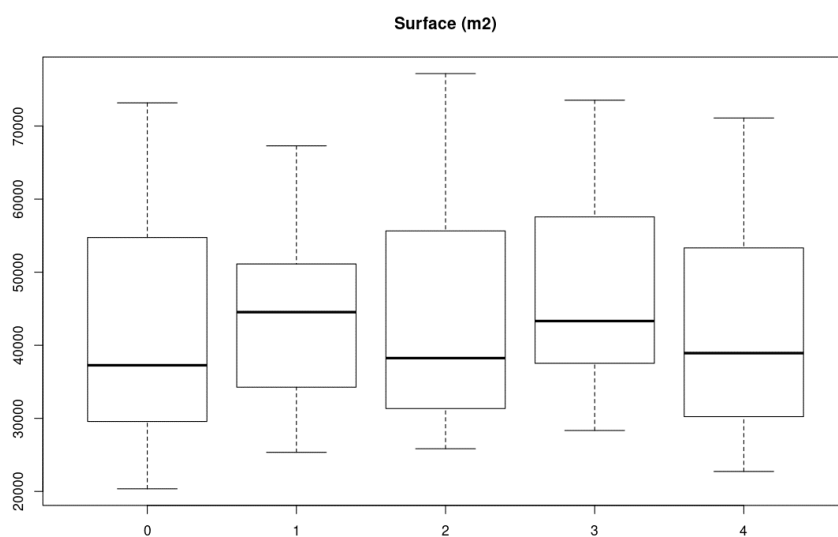


Figure A.33 – Distribution des surfaces journalières parcourues par cinq brebis équipées au domaine du Merle.

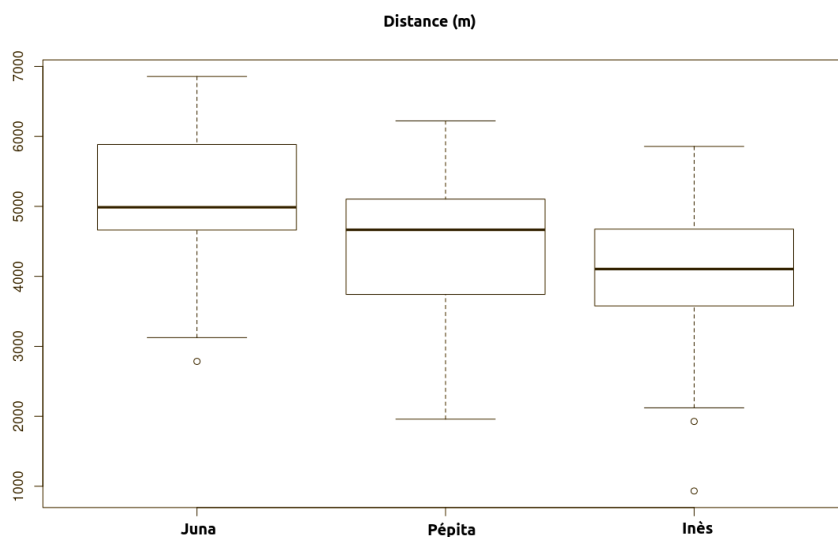


Figure A.34 – Distribution des distances journalières parcourues par trois chèvres équipées dans une exploitation privée.

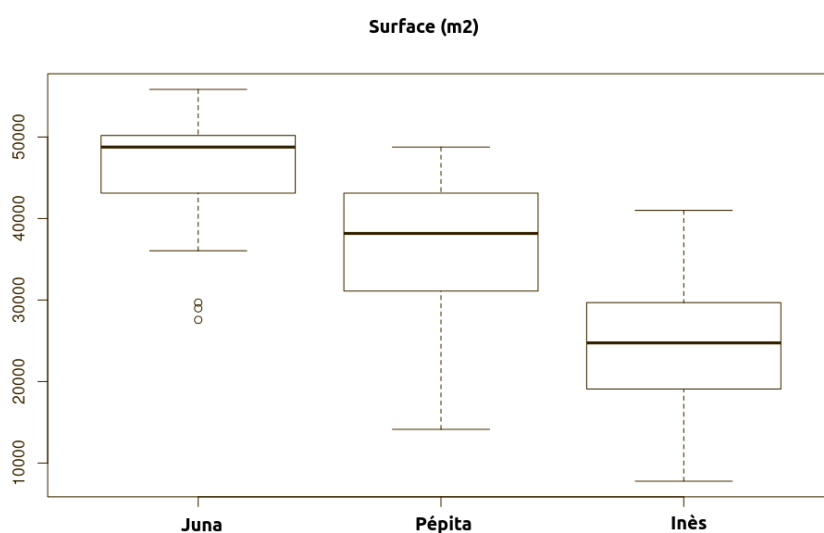


Figure A.35 – Distribution des surfaces journalières parcourues par trois chèvres équipées dans une exploitation privée.

Annexe B

Tableaux

Table B.1 – Résultats des P-valeurs des ANOVAs sur les paires de créneaux horaires entre les phases de pâturage

AND								
Eté						Automne		
Phases	Agnelles MAT			Agnelles ART			Agnelles MAT	Agnelles ART
	P 1 - P 2	P 1 - P 3	P 2 - P 3	P 1 - P 2	P 1 - P 3	P 2 - P 3	P 1 - P 2	P 1 - P 2
00-03h	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900
03-06h	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900
06-09h	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900
09-12h	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900
12-15h	0.414	0.900	0.900	0.900	0.900	0.900	0.900	0.900
15-18h	0.580	0.900	0.900	0.900	0.900	0.900	0.900	0.900
18-21h	0.900	0.900	0.900	0.900	0.149	0.447	0.891	0.900
21-00h	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900
Alpha shape								
Eté						Automne		
Phases	Agnelles MAT			Agnelles ART			Agnelles MAT	Agnelles ART
	P 1 - P 2	P 1 - P 3	P 2 - P 3	P 1 - P 2	P 1 - P 3	P 2 - P 3	P 1 - P 2	P 1 - P 2
00-03h	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900
03-06h	0.900	0.900	0.860	0.900	0.900	0.900	0.840	0.900
06-09h	0.498	0.753	0.900	0.900	0.900	0.900	0.827	0.900
09-12h	0.900	0.900	0.900	0.900	0.191	0.112	0.900	0.900
12-15h	0.161	0.318	0.900	0.900	0.742	0.001	0.900	0.296
15-18h	0.308	0.900	0.494	0.900	0.900	0.900	0.900	0.900
18-21h	0.900	0.900	0.900	0.900	0.113	0.900	0.576	0.900
21-00h	0.900	0.548	0.900	0.900	0.900	0.900	0.900	0.900

Références

- Aggarwal, V., Gupta, V., Singh, P., Sharma, K., Sharma, N., 2019. Detection of spatial outlier by using improved z-score test, in : 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), pp. 788–790. doi :[10.1109/ICOEI.2019.8862582](https://doi.org/10.1109/ICOEI.2019.8862582).
- Agreil, C., Meuret, M., Vincent, M., 2004. GRENOUILLE : une méthode pour gérer les ressources alimentaires pour des ovins sur milieux embroussaillés , 17.
- Agreste, 2015. L'utilisation du territoire en 2014. Teruti-Lucas. Agreste Chiffres et Données Agriculture , 105.
- Akasbi, Z., Oldeland, J., Dengler, J., Finckh, M., 2012. Analysis of GPS trajectories to assess goat grazing pattern and intensity in Southern Morocco. *The Rangeland Journal* 34, 415.
- Akkiraju, N., Edelsbrunner, H., Facello, M., Mocke, E., Fu, P., Varela, C., 1995. Alpha shapes : definition and software, in : *Software Workshop*, Minneapolis.
- Allen, V., Batello, C., Berretta, E., Hodgson, J., Kothmann, M., Li, X., McIvor, J., Milne, J., Morris, C., Peeters, A., Sanderson, M., The Forage and Grazing Terminology Committee, 2011. An international terminology for grazing lands and grazing animals. *Grass and Forage Science* 66, 2–28. URL : <https://onlinelibrary.wiley.com/doi/10.1111/j.1365-2494.2010.00780.x>, doi :[10.1111/j.1365-2494.2010.00780.x](https://doi.org/10.1111/j.1365-2494.2010.00780.x).
- Alvarenga, F., Borges, I., Oddy, V., Dobos, R., 2020. Discrimination of biting and chewing behaviour in sheep using a tri-axial accelerometer. *Computers and Electronics in Agriculture* 168, 105051. URL : <https://linkinghub.elsevier.com/retrieve/pii/S016816991930780X>, doi :[10.1016/j.compag.2019.105051](https://doi.org/10.1016/j.compag.2019.105051).
- Anderson, D.M., 2007. Virtual fencing - past, present and future. *The Rangeland Journal* 29, 65.
- Antle, J.M., Jones, J.W., Rosenzweig, C.E., 2017. Next generation agricultural system data, models and knowledge products : Introduction. *Agricultural Systems* 155, 186–190. URL : <https://linkinghub.elsevier.com/retrieve/pii/S0308521X16304887>, doi :[10.1016/j.agsy.2016.09.003](https://doi.org/10.1016/j.agsy.2016.09.003).

- Archer, K.J., Kimes, R.V., 2008. Empirical characterization of random forest variable importance measures. *Computational Statistics & Data Analysis* 52, 2249–2260.
- Aubron, C., Garambois, N., Nozieres-Petit, M.O., 2019. L'économie agropastorale revisitée. Formes et conditions de développement de systèmes agropastoraux conciliant création de richesse et d'emploi et entretiens des écosystèmes. volume Hors-série. Pastum ed., AFP et Cardère éditeur.
- Aubron, C., Peglion, M., Nozières, M.O., Boutonnet, J.P., 2014. Démarches qualité et pastoralisme en France : Synergies et paradoxes. *Revue de géographie alpine* 102, 17.
- Bahlo, C., Dahlhaus, P., Thompson, H., Trotter, M., 2019. The role of interoperable data standards in precision livestock farming in extensive livestock systems : A review 156, 459–466.
- Bailey, D.W., Mosley, J.C., Estell, R.E., Cibils, A.F., Horney, M., Hendrickson, J.R., Walker, J.W., Launchbaugh, K.L., Burritt, E.A., 2019. Synthesis Paper : Targeted Livestock Grazing : Prescription for Healthy Rangelands. *Rangeland Ecology & Management* 72, 865–877.
- Bailey, D.W., Provenza, F.D., 2008. Mechanisms Determining Large-Herbivore Distribution, in : Prins, H.H.T., Van Langevelde, F. (Eds.), *Resource Ecology*. Springer Netherlands, Dordrecht, pp. 7–28.
- Bailey, D.W., Stephenson, M.B., Pittarello, M., 2015. Effect of terrain heterogeneity on feeding site selection and livestock movement patterns. *Animal Production Science* 55, 298.
- Banhazi, T.M., Lehr, H., Black, J.L., Crabtree, H., Schofield, P., Tschärke, M., Berckmans, D., 2012. Precision Livestock Farming : An international review of scientific and commercial aspects. *Biol Eng* 5, 9.
- Barrière, O., 2015. Patrimonialisation de la pâture : Entre marginalisation et valeur universelle. *Techniques & culture* 63, 182–201.
- Barrière, O., Bes, C., 2017. Droit foncier et pastoralisme, entre propriété et territoire. *VertigO* 17.
- Barroso, F., Alados, C., Boza, J., 2000. Social hierarchy in the domestic goat : effect on food habits and production. *Applied Animal Behaviour Science* 69, 35–53.
- Bataille, J.F., Launay, F., Aussibal, G., Moulin, C.H., Nozieres, M.O., Lasseur, J., Tchakerian, E., 2016. Des scénarios de futurs possibles pour des systèmes d'élevage pastoraux méditerranéens. *Innovations Agronomiques* 49, 309–325.
- Bekker, M., Long, J., 2000. User Involvement in the Design of Human—Computer Interactions : Some Similarities and Differences between Design Approaches, in : *People and*

- Computers XIV - Usability or Else! (Proceedings of HCI'2000). Springer-Verlag, London, pp. 135–147.
- Belgiu, M., Drăguț, L., 2016. Random forest in remote sensing : A review of applications and future directions 114, 24–31.
- Ben-Said, M., 2021. Spatial point-pattern analysis as a powerful tool in identifying pattern-process relationships in plant ecology : an updated review. *Ecological Processes* 10, 56. URL : <https://ecologicalprocesses.springeropen.com/articles/10.1186/s13717-021-00314-4>, doi :10.1186/s13717-021-00314-4.
- Berckmans, D., 2017. General introduction to precision livestock farming. *Animal Frontiers* 7, 6–11. URL : <https://academic.oup.com/af/article/7/1/6/4638786>, doi :10.2527/af.2017.0102.
- Bernués, A., Ruiz, R., Olaizola, A., Villalba, D., Casasús, I., 2011. Sustainability of pasture-based livestock farming systems in the European Mediterranean context : Synergies and trade-offs. *Livestock Science* 139, 44–57.
- Blanc, J., 2009. Savoirs relationnels et « engagement » avec le vivant : les dimensions oubliées du métier d'éleveur? *Natures Sciences Sociétés* 17, 29–39. URL : <https://www.cairn.info/revue-natures-sciences-societes-2009-1-page-29.htm>. place : Les Ulis Publisher : EDP Sciences.
- Bocquier, F., 2006. Device and method for spatially and temporarily counting animals. WO/2006/111680, 2006, INRA, SupAgro, CIRAD. Fin du soutien de l'INRA en Mars 2013. .
- Bocquier, F., Debus, N., Lurette, A., Maton, C., Viudes, G., Moulin, C.H., Jouven, M., 2014. Elevage de précision en systèmes d'élevage peu intensifiés. *INRA Productions Animales* 27, 101–112.
- Bocquier, F., Jouven, M., 2017. Quelle place pour l'élevage de précision dans le contexte du bassin méditerranéen? *Watch Letter - Lettre de veille du CIHEAM* 38, 8.
- Bouju, S., 2000. Evolution des systèmes d'élevage de part et d'autre de la Méditerranée : une difficile conciliation avec des objectifs de développement durable. Quelques réflexions à partir de deux études de cas en France (Préalpes de Digne) et en Tunisie (Khroumirie), in : Bourbouze, A., Qarro, M. (Eds.), *Rupture : nouveaux enjeux, nouvelles fonctions, nouvelle image de l'élevage sur parcours*, Montpellier. pp. 145–157.
- Bourbouze, A., 2018. Les grandes transformations du pastoralisme méditerranéen et l'émergence de nouveaux modes de production. *Watch Letter - Lettre de veille du CIHEAM* 39, 7–12.

- Bourdoncle, M., 2022. La pandémie et l'agriculture : un virus accélérateur de mutations? : Thierry Pouch et Marine Raffray. *Paysans & société* N° 393, 51–53. URL : <https://www.cairn.info/revue-paysan-et-societe-2022-3-page-51.htm?ref=doi>, doi :10.3917/pes.393.0051.
- Boussahoua, M., Boussaid, O., Bentayeb, F., 2017. Optimisation des performances dans les entrepôts de données NoSQL orientés colonnes, in : 17ème Journées Francophones Extraction et Gestion des Connaissances, EGC 2017, Hermann Éditions, Grenoble, France. pp. 69–80. URL : <https://hal.archives-ouvertes.fr/hal-02053993>.
- Breiman, L., 2001. Random forests. *Machine Learning* 45, 5–32.
- Buerkert, A., Schlecht, E., 2009. Performance of three GPS collars to monitor goats' grazing itineraries on mountain pastures. *Computers and Electronics in Agriculture* 65, 85–92.
- Cabaret, J., 2004. Parasitisme helminthique en élevage biologique ovin : réalités et moyens de contrôle. *INRAE Productions Animales* 17, 145–154. URL : <https://productions-animales.org/article/view/3562>, doi :10.20870/productions-animales.2004.17.2.3562.
- Caja, G., Castro-Costa, A., Salama, A.A., Oliver, J., Baratta, M., Ferrer, C., Knight, C.H., 2020. Sensing solutions for improving the performance, health and wellbeing of small ruminants. *Journal of Dairy Research* 87, 34–46. doi :10.1017/S0022029920000667.
- Candela, I., 2022. Calcul d'indicateur pour l'aide à la conduite de troupeaux sur parcours grâce aux SIG. Technical Report. Institut Agro Montpellier. Montpellier, France.
- Capalbo, S.M., Antle, J.M., Seavert, C., 2017. Next generation data systems and knowledge products to support agricultural producers and science-based policy decision making 155, 191–199.
- Capinha, C., Pateiro-López, B., 2014. Predicting species distributions in new areas or time periods with alpha-shapes. *Ecological Informatics* 24, 231–237. URL : <https://linkinghub.elsevier.com/retrieve/pii/S1574954114000594>, doi :10.1016/j.ecoinf.2014.06.001.
- Castro, M., Ameray, A., Castro, J.P., 2020. A new approach to quantify grazing pressure under mediterranean pastoral systems using GIS and remote sensing 41, 5371–5387.
- Cedea, Csao/Ocde, 2008. Elevage et marché régional au Sahel et en Afrique de l'Ouest. Potentialités et défis. Editions club du sahel et de l'afrique de l'ouest/ocde ed.
- Celik, M., Dadaser-Celik, F., Dokuz, A.s., 2011. Anomaly detection in temperature data using dbscan algorithm, in : 2011 International Symposium on Innovations in Intelligent Systems and Applications, pp. 91–95. doi :10.1109/INISTA.2011.5946052.

- CEN-L-R, 2017. Terres pastorales : diversité et valeurs des milieux ouverts méditerranéens. Cen l-r / rouergue ed., CEN L-R / ROUERQUE, France.
- Charbonier, Q. (Ed.), 1972. La loi pastorale. La Cardère, France.
- Chebli, Y., Otmani, S.E., Chentouf, M., Hornick, J.L., Bindelle, J., Cabaraux, J.F., 2020. Foraging behavior of goats browsing in southern mediterranean forest rangeland. *Animals* 10. URL : <https://www.mdpi.com/2076-2615/10/2/196>.
- Cheng, H., Jin, B., Luo, K., Pei, J., Zhang, X., Zhang, Y., Tang, J., Yang, Q., Sun, G., 2021. Vegetation Response to Goats Grazing Intensity in Semiarid Hilly Grassland of the Loess Plateau, Lanzhou, China. *Sustainability* 13, 3569. URL : <https://www.mdpi.com/2071-1050/13/6/3569>, doi :10.3390/su13063569.
- Chevalier, M., El Malki, M., Teste, O., Kopliku, A., Tournier, R., 2017. Un benchmark enrichi pour l'évaluation des entrepôts de données noSQL volumineuses et variables, in : Journées Francophones sur les Entrepôts de Données et l'Analyse en ligne (EDA 2017), Lyon, France. pp. 11–26. URL : <https://hal.archives-ouvertes.fr/hal-02885343>.
- Clark, P.J., Evans, F.C., 1954. Distance to Nearest Neighbor as a Measure of Spatial Relationships in Populations. *Ecology* 35, 445–453. URL : <http://www.jstor.org/stable/1931034>, doi :10.2307/1931034. publisher : Ecological Society of America.
- Colhoun, O., 2018. Data Warehouse, in : Gressner, A.M., Arndt, T. (Eds.), *Lexikon der Medizinischen Laboratoriumsdiagnostik*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–1. URL : http://link.springer.com/10.1007/978-3-662-49054-9_826-1, doi :10.1007/978-3-662-49054-9_826-1.
- Corbellini, A., Mateos, C., Zunino, A., Godoy, D., Schiaffino, S., 2017. Persisting big-data : The NoSQL landscape. *Information Systems* 63, 1–23. URL : <https://linkinghub.elsevier.com/retrieve/pii/S0306437916303210>, doi :10.1016/j.is.2016.07.009.
- Cowling, R.M., Rundel, P.W., Lamont, B.B., Arroyo, M.K., Arianoutsou, M., 1996. Plant diversity in mediterranean-climate regions. *Trends in Ecology & Evolution* 11, 362–366. URL : <https://www.sciencedirect.com/science/article/pii/0169534796100446>, doi :[https://doi.org/10.1016/0169-5347\(96\)10044-6](https://doi.org/10.1016/0169-5347(96)10044-6).
- Dardel, C., Kergoat, L., Hiernaux, P., Mougin, E., Manuela, G., Tucker, C., 2014. Re-greening Sahel : 30 years of remote sensing data and field observations (Mali, Niger). *Remote Sensing of Environment* 140, 350–364. doi :10.1016/j.rse.2013.09.011.
- Davies, J., Herrera, P., Ruiz-Mirazo, J., Mohamed-Katerere, J., Hannam, I., Nuesri, E., 2017. Améliorer la gouvernance des terres pastorales : Mettre en œuvre les Directives volontaires pour une gouvernance responsable des régimes fonciers applicables aux terres, aux pêches et aux forêts dans le contexte de la sécurité alimentaire nationale. Number 6 in GUIDE TECHNIQUE POUR LA GOUVERNANCE DES RÉGIMES FONCIERS.

- Davoudian, A., Chen, L., Liu, M., 2019. A Survey on NoSQL Stores. *ACM Computing Surveys* 51, 1–43. URL : <https://dl.acm.org/doi/10.1145/3158661>, doi :10.1145/3158661.
- De Cao, G., Ickowicz, A., Toure, I., Gerber, P., 2008. An information and early warning system designed for sahelian pastoral systems : the example of SIPSA implementation in Senegal». *Journal of Agriculture and Environment for International Development* 102, 141–159.
- De Haan, C., Steinfeld, H., Blackburn, H., 1999. *Elevage et Environnement. A la recherche d'un équilibre.* Fao ed.
- Delegido, J., Verrelst, J., Meza, C., Rivera, J., Alonso, L., Moreno, J., 2013. A red-edge spectral index for remote sensing estimation of green LAI over agroecosystems. *European Journal of Agronomy* 46, 42–52.
- Dewi, C., Chen, R.C., 2019. RANDOM FOREST AND SUPPORT VECTOR MACHINE ON FEATURES SELECTION FOR REGRESSION ANALYSIS. *International Journal of Innovative Computing, Information and Control* 15, 2027–2037.
- Dixon, J., 2010. Pentaho, Hadoop, and Data Lakes URL : <https://jamesdixon.wordpress.com/2010/10/14/pentaho-hadoop-and-data-lakes/>.
- Dixon, P.M., 2001. Nearest Neighbor Methods 51, 28.
- Djibo Saibou, B., 2020. Elaboration d'un modèle d'utilisation spatialisée de parcours méditerranéens pâturés par des herbivores domestiques. Ph.D. thesis. Institut Agro Montpellier. Montpellier, France.
- Dono, G., Cortignani, R., Dell'Unto, D., Deligios, P., Doro, L., Lacetera, N., Mula, L., Pasqui, M., Quaresima, S., Vitali, A., Roggero, P.P., 2016. Winners and losers from climate change in agriculture : Insights from a case study in the Mediterranean basin 147, 65–75.
- Dumont, B., Andueza, D., Niderkorn, V., Lüscher, A., Porqueddu, C., Picon-Cochard, C., 2015. A meta-analysis of climate change effects on forage quality in grasslands : specificities of mountain and Mediterranean areas. *Grass and Forage Science* 70, 239–254.
- Edelsbrunner, H., Kirkpatrick, D., Seidel, R., 1983. On the shape of a set of points in the plane. *IEEE Transactions on Information Theory*. v. 29 (4).-p. 551-559 .
- Erturk, E., Shi, H., 2016. Natural Language Processing using Hadoop and KOSHIK. *ArXiv abs/1608.04434*. ArXiv : 1608.04434.
- Escadafal, R., 1989. Remote sensing of arid soil surface color with Landsat thematic mapper. *Advances in Space Research* 9, 159–163.
- Etienne, L., Jouven, M., Launay, F., 2021. Des outils et méthodes pour enseigner et accompagner .

- Evangelou, C., Yiakoulaki, M., Papanastasis, V., 2014. Spatio-Temporal Analysis of Sheep and Goats Grazing in Different Forage Resources of Northern Greece. *Hacquetia* 13, 205–213.
- Eychenne, C., 2014. Ecopastoralisme et écopâturage : éléments de définition et de discussion, in : *Rencontres nationales de l'écopâturage* organisées par l'Association Entretien, Nature & Territoire, Association Entretien, Nature & Territoire, Saint Herblain, France. URL : <https://halshs.archives-ouvertes.fr/halshs-01862262>.
- Fang, H., 2015. Managing Data Lakes in Big Data Era , 5.
- Farrié, B., Jouven, M., Launay, F., Moreau, J.C., Moulin, C.H., Piquet, M., Taverne, M., Tchakérian, E., Thénard, V., Martin, G., 2015. Rangeland Rummy – A board game to support adaptive management of rangeland-based livestock systems. *Journal of Environmental Management* 147, 236–245. URL : <https://linkinghub.elsevier.com/retrieve/pii/S0301479714004320>, doi :10.1016/j.jenvman.2014.08.018.
- Farrugia, A., Claxton, R., Thompson, S., 2016. Towards social network analytics for understanding and managing enterprise data lakes, in : *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, IEEE, San Francisco, CA, USA. pp. 1213–1220. URL : <http://ieeexplore.ieee.org/document/7752393/>, doi :10.1109/ASONAM.2016.7752393.
- Favre, C., Bentayeb, F., Boussaid, O., Darmont, J., Gavin, G., Harbi, N., Kabachi, N., Loudcher, S., 2013. Les entrepôts de données pour les nuls... ou pas!, in : *2ème Atelier alde à la Décision à tous les Etages (EGC/AIDE 13)*, Toulouse, France. pp. –. URL : <https://hal.archives-ouvertes.fr/hal-00783638>.
- Fawagreh, K., Gaber, M.M., Elyan, E., 2014. Random forests : from early developments to recent advancements. *Systems Science & Control Engineering* 2, 602–609.
- Feldt, T., Schlecht, E., 2016. Analysis of GPS trajectories to assess spatio-temporal differences in grazing patterns and land use preferences of domestic livestock in southwestern Madagascar. *Pastoralism* 6, 5.
- Fern, R.R., Foxley, E.A., Bruno, A., Morrison, M.L., 2018. Suitability of NDVI and OSAVI as estimators of green biomass and coverage in a semi-arid rangeland. *Ecological Indicators* 94, 16–21. URL : <https://linkinghub.elsevier.com/retrieve/pii/S1470160X18304692>, doi :10.1016/j.ecolind.2018.06.029.
- Fogarty, E.S., Swain, D.L., Cronin, G.M., Moraes, L.E., Trotter, M., 2020. Behaviour classification of extensively grazed sheep using machine learning. *Computers and Electronics in Agriculture* 169, 105175. URL : <https://linkinghub.elsevier.com/retrieve/pii/S0168169919318319>, doi :10.1016/j.compag.2019.105175.

- Fournier-Viger, P., Lin, J.C.W., 2017. A Survey of Sequential Pattern Mining. *Ubiquitous International* 1, 24.
- Fournier-Viger, P., Nkambou, R., Mephu Nguifo, E., 2008. A Knowledge Discovery Framework for Learning Task Models from User Interactions in Intelligent Tutoring Systems, in : *MICAL 2008 : Advances in Artificial Intelligence*, Atizapán de Zaragoza, Mexico. pp. 765–778. URL : <https://hal.archives-ouvertes.fr/hal-02134394>, doi : [10.1007/978-3-540-88636-5_72](https://doi.org/10.1007/978-3-540-88636-5_72).
- Frenay, B., Verleysen, M., 2014. Classification in the Presence of Label Noise : A Survey. *IEEE Transactions on Neural Networks and Learning Systems* 25, 845–869. URL : <http://ieeexplore.ieee.org/document/6685834/>, doi : [10.1109/TNNLS.2013.2292894](https://doi.org/10.1109/TNNLS.2013.2292894).
- García, I., Martínez-Prieto, M.A., Bregón, A., Álvarez, P.C., Díaz, F., 2017. Towards a Scalable Architecture for Flight Data Management :, in : *Proceedings of the 6th International Conference on Data Science, Technology and Applications*, SCITEPRESS - Science and Technology Publications. pp. 263–268. URL : <http://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0006473402630268>, doi : [10.5220/0006473402630268](https://doi.org/10.5220/0006473402630268).
- Gatrell, A.C., Bailey, T.C., Diggle, P.J., Rowlingson, B.S., 1996. Spatial Point Pattern Analysis and Its Application in Geographical Epidemiology. *Transactions of the Institute of British Geographers* 21, 256. URL : <https://www.jstor.org/stable/622936?origin=crossref>, doi : [10.2307/622936](https://doi.org/10.2307/622936).
- Gautier, D., Garde, L., Aussibal, G., 2011. Des systèmes d'élevage de petits ruminants basés sur le pâturage des parcours embroussaillés., in : Bernués, A., Boutonnet, J., Casaasus, I., Chentouf, M., Gabifia, D., Joy, M., Lopez-François, A., Morand-Fehr, P., Pachero, F. (Eds.), *Economic, social and environmental sustainability in sheep and goat production systems. options méditerranéennes : série a. séminaires méditerranéens ; n . 1 00 ed.*. Options Méditerranéennes : Série A. Séminaires Méditerranéens ; n . 1 00, Zaragoza : CIHEAM / FAO / CITA-DGA. volume 100 of *Options Méditerranéennes : Série A. Séminaires Méditerranéens ; n . 1 00*, pp. 243–248.
- Ginane, C., Dumont, B., Baumont, R., Praches, S., Farrugia, A., 2008. Comprendre le comportement alimentaire des herbivores au pâturage : intérêts pour l'élevage et l'environnement, in : *Rencontres autour des Recherches sur les Ruminants*, Paris. pp. 315–322.
- Gorelik, A., 2019. *The Enterprise Big Data Lake*. O'reilly media, inc. ed., O'Reilly Media, Inc.
- Grisot, P.G., Llaria, A., Philibert, A., Guinamard, C., Menassol, J.B., Flégar, M., Jouven, M., 2022. Clochète - caractérisation du comportement et localisation des ovins et des caprins grâce aux technologies embarquées : aide à la conduite des troupeaux et à la valorisation des surfaces pastorales , 6.

- Guérin, G., Agreil, C., Aussibal, G., Garde, L., Gautier, D., Meuret, M., 2007. Qualifier les surfaces pastorales pour combiner le renouvellement des ressources alimentaires et la maîtrise des couverts végétaux. *Acquis, enjeux et questions actuelles*, 9.
- Guérin, G., Bellon, S., Gautier, D., 2001. Valorisation et maîtrise des surfaces pastorales par le pâturage. *Fourrages, Association Française pour la Production Fourragère*, 239–256.
- Hai, R., Geisler, S., Quix, C., 2016. Constance : An Intelligent Data Lake System, in : *Proceedings of the 2016 International Conference on Management of Data - SIGMOD '16*, ACM Press, San Francisco, California, USA. pp. 2097–2100. URL : <http://dl.acm.org/citation.cfm?doid=2882903.2899389>, doi :10.1145/2882903.2899389.
- Handcock, R., Swain, D., Bishop-Hurley, G., Patison, K., Wark, T., Valencia, P., Corke, P., O'Neill, C., 2009. Monitoring Animal Behaviour and Environmental Interactions Using Wireless Sensor Networks, GPS Collars and Satellite Remote Sensing. *Sensors* 9, 3586–3603.
- Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics SMC-3*, 610–621. URL : <http://ieeexplore.ieee.org/document/4309314/>, doi :10.1109/TSMC.1973.4309314.
- Hiernaux, P., Diawara, M., Gangneron, F., 2014. Quelle accessibilité aux ressources pastorales du Sahel? L'élevage face aux variations climatiques et aux évolutions des sociétés sahéliennes. *Afrique contemporaine* 249, 21–35. URL : <https://www.cairn.info/revue-afrique-contemporaine1-2014-1-page-21.htm>, doi :10.3917/afco.249.0021. place : Louvain-la-Neuve Publisher : De Boeck Supérieur.
- Hill, M.J., Donald, G.E., Hyder, M.W., Smith, R.C., 2004. Estimation of pasture growth rate in the south west of Western Australia from AVHRR NDVI and climate data. *Remote Sensing of Environment* 93, 528–545. URL : <https://linkinghub.elsevier.com/retrieve/pii/S0034425704002457>, doi :10.1016/j.rse.2004.08.006.
- Honnutagi, P.S., 2014. The Hadoop distributed file system. *International Journal of Computer Science and Information Technologies* 5, 6238–6243.
- Hubert, B., Deverre, C., Meuret, M., 2010. Deux siècles de changements radicaux pour le parcours du sud de la France, in : Meuret, M. (Ed.), *Un savoir-faire de bergers*. Educagri & Quæ, France.
- Ibra, T., Ickowicz, A., A., W., Garba, I., Gerber, P.J., 2012. Atlas of Trends in pastoral systems in the Sahel 1970-2012. SIPSA. FAO-CIRAD, 32 pages (<http://umr-selmet.cirad.fr/publications-et-ressources/documents-techniques>).
- Inglada, J., Vincent, A., Arias, M., Tardy, B., Morin, D., Rodes, I., 2017. Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image

- Time Series. *Remote Sensing* 9, 95. URL : <http://www.mdpi.com/2072-4292/9/1/95>, doi : [10.3390/rs9010095](https://doi.org/10.3390/rs9010095).
- Jago, J., Eastwood, C., Kerrisk, K., Yule, I., 2013. Precision dairy farming in Australasia : adoption, risks and opportunities. *Animal Production Science* 53, 907. URL : <http://www.publish.csiro.au/?paper=AN12330>, doi : [10.1071/AN12330](https://doi.org/10.1071/AN12330).
- Javelle, A., Jouven, M.J.P., 2015. Comment l'agroécologie ré-interroge les pratiques agricoles ? L'enjeu autour de la construction et de la transmission des savoirs, in : *L'agroécologie : du nouveau pour le pastoralisme ?*. association française de pastoralisme (afp) ed.. Association Française de Pastoralisme (AFP), Montpellier, France, p. 2.
- Jouven, M., 2018. Utilisation des végétations spontanées par les élevages d'herbivores. Une approche fonctionnelle pour accompagner la transition agroécologique. HDR. Montpellier SupAgro. Montpellier, France.
- Jouven, M., Casasus, I., Franca, A., 2019. Improving the use of grazed forage in Mediterranean ruminant systems issues, options and perspectives. *Options Méditerranéennes Serie A* 125, 237–256.
- Jouven, M., Lapeyronie, P., Moulin, C.H., Bocquier, F., 2010. Rangeland utilization in Mediterranean farming systems. *animal* 4, 1746–1757.
- Jáuregui, B.M., García, U., Osoro, K., Celaya, R., 2009. Sheep and Goat Grazing Effects on Three Atlantic Heathland Types. *Rangeland Ecology & Management* 62, 119–126.
- Júnior, R.F.d.V., Siqueira, H.E., Valera, C.A., Oliveira, C.F., Fernandes, L.F.S., Moura, J.P., Pacheco, F.A.L., 2019. Diagnosis of degraded pastures using an improved NDVI-based remote sensing approach : An application to the Environmental Protection Area of Uberaba River Basin (Minas Gerais, Brazil). *Remote Sensing Applications : Society and Environment* 14, 20–33. URL : <https://www.sciencedirect.com/science/article/pii/S2352938518303409>, doi : <https://doi.org/10.1016/j.rsase.2019.02.001>.
- Kafando, R., Decoupes, R., Sautot, L., Teisseire, M., 2020. Spatial Data Lake for Smart Cities : From Design to Implementation. *AGILE : GIScience Series* 1, 1–15.
- Kawamura, K., Akiyama, T., Yokota, H.o., Tsutsumi, M., Yasuda, T., Watanabe, O., Wang, S., 2005. Quantifying grazing intensities using geographic information systems and satellite remote sensing in the Xilingol steppe region, Inner Mongolia, China. *Agriculture, Ecosystems & Environment* 107, 83–93.
- Khalilia, M., Chakraborty, S., Popescu, M., 2011. Predicting disease risks from highly imbalanced data using random forest. *BMC Medical Informatics and Decision Making* 11, 51.

- Khine, P.P., Wang, Z.S., 2018. Data lake : a new ideology in big data era. ITM Web of Conferences 17, 03025. URL : <https://www.itm-conferences.org/10.1051/itmconf/20181703025>, doi :10.1051/itmconf/20181703025.
- Khoshgoftaar, T.M., Golawala, M., Hulse, J.V., 2007. An Empirical Study of Learning from Imbalanced Data Using Random Forest, IEEE, Patras, Greece. pp. 310–317.
- Kowal, V., Ahlborn, J., Jamsranjav, C., Avirmed, O., Chaplin-Kramer, R., 2021. Modeling Integrated Impacts of Climate Change and Grazing on Mongolia's Rangelands. Land 10, 397. URL : <https://www.mdpi.com/2073-445X/10/4/397>, doi :10.3390/land10040397.
- Laca, E.A., 2009. Precision livestock production : tools and concepts. Revista Brasileira de Zootecnia 38, 123–132.
- Lasanta, T., Khorchani, M., Pérez-Cabello, F., Errea, P., Sáenz-Blanco, R., Nadal-Romero, E., 2018. Clearing shrubland and extensive livestock farming : Active prevention to control wildfires in the Mediterranean mountains 227, 256–266.
- Lebaudy, G., 2004. Gravures et graffiti des bergers de la plaine de la Crau : un patrimoine fragile et méconnu. Ecologia mediterranea 30, 35–45. URL : https://www.persee.fr/doc/ecmed_0153-8756_2004_num_30_1_1494, doi :10.3406/ecmed.2004.1494.
- Li, L., Ban, H., Wechsler, S.P., Xu, B., 2018. Spatial Data Uncertainty, in : Huang, B. (Ed.), Comprehensive Geographic Information Systems. Elsevier, Oxford, pp. 313–340. URL : <https://www.sciencedirect.com/science/article/pii/B978012409548909610X>, doi :<https://doi.org/10.1016/B978-0-12-409548-9.09610-X>.
- Lin, J., Keogh, E., Wei, L., Lonardi, S., 2007. Experiencing sax : a novel symbolic representation of time series. Data Mining and Knowledge Discovery 15, 107–144.
- Lin, W., Wu, Z., Lin, L., Wen, A., Li, J., 2017. An Ensemble Random Forest Algorithm for Insurance Big Data Analysis 5, 16568–16575.
- Llave, M.R., 2018. Data lakes in business intelligence : reporting from the trenches. Procedia Computer Science 138, 516–524. URL : <http://www.sciencedirect.com/science/article/pii/S1877050918317046>, doi :10.1016/j.procs.2018.10.071.
- Lozano, F.D., Schwartz, M.W., 2005. Patterns of rarity and taxonomic group size in plants. Biological Conservation 126, 146–154. URL : <https://www.sciencedirect.com/science/article/pii/S0006320705002090>, doi :<https://doi.org/10.1016/j.biocon.2005.04.024>.
- Lécrivain, E., Leroy, A., Savini, I., Deffontaines, J.P., 1993. Les formes de troupeau au pâturage : Genèse et diversité. Études et Recherches sur les Systèmes Agraires et le Développement , 237–263.

- López De La Cruz, J., Gutiérrez, M., 2008. Spatial statistics of pitting corrosion patterning : Quadrat counts and the non-homogeneous Poisson process. *Corrosion Science* 50, 1441–1448. URL : <https://linkinghub.elsevier.com/retrieve/pii/S0010938X08000346>, doi :10.1016/j.corsci.2008.01.003.
- Ma, Q., Chai, L., Hou, F., Chang, S., Ma, Y., Tsunekawa, A., Cheng, Y., 2019. Quantifying Grazing Intensity Using Remote Sensing in Alpine Meadows on Qinghai-Tibetan Plateau. *Sustainability* 11, 417. URL : <http://www.mdpi.com/2071-1050/11/2/417>, doi :10.3390/su11020417.
- MacEachren, A., Robinson, A., Hopper, S., Gardner, S., Murray, R., Gahegan, M., Hetzler, E., 2005. Visualizing geospatial information uncertainty : What we know and what we need to know. *Cartography and Geographic Information Science* 32, 139–160. doi :10.1559/1523040054738936.
- Madera, C., 2019. L'évolution des systèmes et architectures d'information sous l'influence des données massives : les lacs de données. Ph.D. thesis. Université Montpellier.
- Malzac, G., Bonnet, O., Garde, L., Genevet, E., Micola, S., Rocher, C., Etienne, L., Launay, F., 2018. StratPasto, un outil de caractérisation des systèmes d'alimentation des élevages à composante pastorale. », Plaquette de description de l'outil. URL : <https://cerpam.com/guides-plaquettes-etudes/stratpasto/>.
- Mathis, C., 2017. Data Lakes. *Datenbank-Spektrum* 17, 289–293. URL : <http://link.springer.com/10.1007/s13222-017-0272-7>, doi :10.1007/s13222-017-0272-7.
- Maton, C., Bouquet, P.M., Laville, Y., Bocquier, F., 2006. Automate de tri pour les ovins basé sur l'identification électronique. *Rencontres Recherches Ruminants* 13, 167–167.
- Maton, C., Montagnac, D., Viudes, G., Bouquet, P., Bocquier, F., 2008. Les applications de l'identification électronique des petits ruminants au service de l'élevage biologique. *Inovations Agronomiques* 4.
- McAfee, A., Brynjolfsson, E., 2012. Big Data : The Management Revolution. *Harvard Business Review* 90, 60–68.
- McGavin, S.L., Bishop-Hurley, G.J., Charmley, E., Greenwood, P.L., Callaghan, M.J., 2018. Effect of GPS sample interval and paddock size on estimates of distance travelled by grazing cattle in rangeland, Australia. *The Rangeland Journal* 40, 55–64.
- Meuret, M., 2007. Les pratiques pastorales entre temps court de l'alimentation des troupeaux et temps long des ressources et des milieux. *Pastum, Association Française de Pastoralisme* , 5–23.
- Meuret, M., Tichit, M., Hostiou, N., 2013. Elevage et pâturage « de précision » : l'animal sous surveillance électronique. *Courrier de l'environnement de l'INRA* 63, 13–24.

- Micola, S., Genevet, E., 2022. PASTOSCOPIE. Un panorama du pastoralisme en région Occitanie.
- Monod, M.O., Faure, P., Moiroux, L., Rameau, P., 2008. A virtual fence for animals management in rangelands, in : MELECON 2008 - The 14th IEEE Mediterranean Electrotechnical Conference, pp. 337–342. doi :[10.1109/MELCON.2008.4618457](https://doi.org/10.1109/MELCON.2008.4618457).
- Nozieres, M., Etienne, L., Launay, F., Moulin, C., 2021. Les grands traits de l'élevage pastoral aujourd'hui en France. Revue de l'association francophone pour les prairies et les fourrages , 3–11.
- O'Grady, M.J., O'Hare, G.M., 2017. Modelling the smart farm. Information Processing in Agriculture 4, 179–187. URL : <https://linkinghub.elsevier.com/retrieve/pii/S2214317316301287>, doi :[10.1016/j.inpa.2017.05.001](https://doi.org/10.1016/j.inpa.2017.05.001).
- O'Sullivan D., Unwin, D., 2010. Point Pattern Analysis. John Wiley and Sons, Ltd. chapter 5. pp. 121–156. arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470549094.ch5>.
- Pelletier, C., Valero, S., Inglada, J., Champion, N., Dedieu, G., 2016. Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas. Remote Sensing of Environment 187, 156–168. URL : <https://linkinghub.elsevier.com/retrieve/pii/S0034425716303820>, doi :[10.1016/j.rse.2016.10.010](https://doi.org/10.1016/j.rse.2016.10.010).
- Pereira, M.G., Trigo, R.M., da Camara, C.C., Pereira, J.M., Leite, S.M., 2005. Synoptic patterns associated with large summer forest fires in Portugal. Agricultural and Forest Meteorology 129, 11–25.
- Petit, M., Agabriel, J., D'Hour, P., Garel, J.P., 1994. Quelques caractéristiques des races bovines allaitantes de type rustique. INRAE Productions Animales 7. doi :[10.20870/productions-animales.1994.7.4.4171](https://doi.org/10.20870/productions-animales.1994.7.4.4171).
- Pierce, F.J., Nowak, P., 1999. Aspects of Precision Agriculture, in : Advances in Agronomy. Elsevier. volume 67, pp. 1–85. URL : <https://linkinghub.elsevier.com/retrieve/pii/S0065211308605131>, doi :[10.1016/S0065-2113\(08\)60513-1](https://doi.org/10.1016/S0065-2113(08)60513-1).
- Pierpaoli, E., Carli, G., Pignatti, E., Canavari, M., 2013. Drivers of Precision Agriculture Technologies Adoption : A Literature Review 8, 61–69.
- Poilecot, P., Saidi, S., Daget, P., 2013. Stratégie de pâturage d'une population d'éléphants en saison sèche : le cas du parc de Zakouma au Tchad. BOIS & FORETS DES TROPIQUES 317, 59.
- Pokou, Y.J.M., Fournier-Viger, P., Moghrabi, C., 2016. Authorship attribution using small sets of frequent part-of-speech skip-grams., in : International Florida Artificial Intelligence Research Society Conference, AAAI. pp. 86–91.

- Porqueddu, C., Ates, S., Louhaichi, M., Kyriazopoulos, A.P., Moreno, G., del Pozo, A., Ovalle, C., Ewing, M.A., Nichols, P.G.H., 2016. Grasslands in 'Old World' and 'New World' Mediterranean-climate zones : past trends, current status and future research priorities. *Grass and Forage Science* 71, 1–35.
- Pouget, M., Madeira, E., Floch, L., Kamal, S., 1990. Caractéristiques spectrales des surfaces sableuses de la région côtière Nord-Ouest de l’Egypte : application aux données satellitaires SPOT, in : 2ème Journée de Télédétection : caractérisation et suivi des milieux terrestres en régions arides et tropicales, ORSTOM, Paris, France. pp. 4–6.
- Putfarken, D., Dengler, J., Lehmann, S., Härdtle, W., 2008. Site use of grazing cattle and sheep in a large-scale pasture landscape : A GPS/GIS assessment. *Applied Animal Behaviour Science* 111, 54–67.
- Pégliion, M., Launay, F., Rocher, C., Genevet, E., Girardin, S., Dessailly, G., Pedron, T., 2017. LIFE + MIL’OUV : Communication et partage de savoirs pour la préservation des milieux ouverts pastoraux méditerranéens. *Pastum* 108, 17–21.
- Quix, C., Hai, R., 2019. Data Lake, in : Sakr, S., Zomaya, A. (Eds.), *Encyclopedia of Big Data Technologies*. Springer International Publishing, Cham, pp. 1–8.
- de Rancourt, M., Fois, N., Lavín, M., Tchakérian, E., Vallerand, F., 2006. Mediterranean sheep and goats production : An uncertain future. *Small Ruminant Research* 62, 167–179.
- Redman, T.C., 1998. The impact of poor data quality on the typical enterprise. *Communications of the ACM* 41, 79–82. URL : <https://dl.acm.org/doi/10.1145/269012.269025>, doi :10.1145/269012.269025.
- Rey, S., Kang, W., Shao, H., Wolf, L.J., Seth, M., Gaboardi, J., Arribas-Bel, D., 2019. pysal/pointpats : pointpats 2.1.0. URL : <https://doi.org/10.5281/zenodo.3265637>, doi :10.5281/zenodo.3265637.
- Rey, S.J., Anselin, L., 2007. PySAL : A Python Library of Spatial Analytical Methods. *The Review of Regional Studies* 37, 5–27.
- Riedel, J., Bernués, A., Casasús, I., 2013. Livestock Grazing Impacts on Herbage and Shrub Dynamics in a Mediterranean Natural Park. *Rangeland Ecology & Management* 66, 224–233.
- Ripley, B.D., 1988. *Statistical Inference for Spatial Processes*. 1 ed., Cambridge University Press. URL : <https://www.cambridge.org/core/product/identifier/9780511624131/type/book>, doi :10.1017/CB09780511624131.
- Ritchie, H., Rosado, P., Roser, M., 2017. Meat and dairy production. *Our World in Data* <https://ourworldindata.org/meat-production>.

- Rochon, J., Goby, J., 2009. Gestion pastorale du complexe lagunaire de Can et (si te N atu ra 2000 FR 9101465) par des ovins : influence du mode de conduite sur l'utilisation de l'espace, in : Pachero, F., Morand-Fehr, P. (Eds.), Changes in sheep and goat farming systems at the beginning of the 21st century : research, tools, methods and initiatives in favour of a sustainable development. options méditerranéennes : série a. séminaires méditerranéens ; n. 91 ed.. Options Méditerranéennes : Série A. Séminaires Méditerranéens ; n. 91, Zaragoza : CIHEAM / DRAP-Norte / FAO. Options Méditerranéennes : Série A. Séminaires Méditerranéens ; n. 91, pp. 145–149.
- Rondeaux, G., Steven, M., Baret, F., 1996. Optimization of soil-adjusted vegetation indices. Remote Sensing of Environment 55, 95–107. URL : <https://www.sciencedirect.com/science/article/pii/S0034425795001867>, doi :[https://doi.org/10.1016/0034-4257\(95\)00186-7](https://doi.org/10.1016/0034-4257(95)00186-7).
- Roudaut, F., Aussibal, G., Beylier, B., Gardel, L., Gautier, D., 2007. La broussaille, un atout pour le pâturage 14, 4.
- Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring Vegetation Systems in the Great Plains with ERTS. NASA Special publication 351, 309.
- Rutter, S., Beresford, N., Roberts, G., 1997. Use of GPS to identify the grazing areas of hill sheep. Computers and Electronics in Agriculture 17, 177–188.
- Saeys, Y., Inza, I., Larranaga, P., 2007. A review of feature selection techniques in bioinformatics. Bioinformatics 23, 2507–2517.
- Saha, B., Srivastava, D., 2014. Data quality : The other face of Big Data, in : 2014 IEEE 30th International Conference on Data Engineering, IEEE, Chicago, IL, USA. pp. 1294–1297. URL : <http://ieeexplore.ieee.org/document/6816764/>, doi :[10.1109/ICDE.2014.6816764](https://doi.org/10.1109/ICDE.2014.6816764).
- Sawadogo, P.N., Scholly, E., Favre, C., Ferey, E., Loudcher, S., Darmont, J., 2019. Metadata Systems for Data Lakes : Models and Features, in : 1st International Workshop on BI and Big Data Applications, Springer, Bled, Slovenia. pp. 440–451.
- Schellberg, J., Hill, M.J., Gerhards, R., Rothmund, M., Braun, M., 2008. Precision agriculture on grassland : Applications, perspectives and constraints. European Journal of Agronomy 29, 59–71. URL : <https://www.sciencedirect.com/science/article/pii/S1161030108000646>, doi :<https://doi.org/10.1016/j.eja.2008.05.005>.
- Senseman, G.M., Tweddale, S.A., Anderson, A.B., Bagley, C.F., 1996. Correlation of land condition trend analysis (LCTA) rangeland cover measures to satellite-imagery-derived vegetation indices. US Army Corps of Engineering Research Laboratories (USACERL) 97, 29.

- Serrano, J., Shahidian, S., Paixão, L., Marques da Silva, J., Morais, T., Teixeira, R., Domingos, T., 2021. Spatiotemporal Patterns of Pasture Quality Based on NDVI Time-Series in Mediterranean Montado Ecosystem. *Remote Sensing* 13, 3820. URL : <https://www.mdpi.com/2072-4292/13/19/3820>, doi :10.3390/rs13193820.
- Serrano, J., Shahidian, S., Marque da Silva, J., 2013. Precision agriculture '13. Wageningen academic publishers ed., Wageningen Academic Publishers, Wageningen, Netherlands.
- Shakesby, R., 2011. Post-wildfire soil erosion in the Mediterranean : Review and future research directions. *Earth-Science Reviews* 105, 71–100.
- Shaqura, I., Lasseur, J., 2019. Characterizing extensively grazed areas between forests and pastureland to better assess the resilience of Mediterranean livestock activities. Lessons learned from a study on the scale of the Provence Alpes Côte d’Azur region (France)., in : FAO-CIHEAM Network on Sheep and Goats and Mediterranean Pastures "Efficiency and resilience of forage resources and small ruminant production to cope with global challenges in Mediterranean areas", Meknes, Morocco. p. 2. URL : <https://hal.inrae.fr/hal-02611577>.
- Shipley, L.A., 1999. Grazers and Browsers : How Digestive Morphology Affects Diet Selection, in : Launchbaugh, K., Sanders, K., Mosley, J. (Eds.), *Grazing Behavior of Livestock and Wildlife*. University of Idaho, Moscow, p. 8.
- Spielman, S.E., 2017. Point Pattern Analysis, in : *International Encyclopedia of Geography*. John Wiley & Sons, Ltd, pp. 1–9. URL : <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118786352.wbieg0849>, doi :<https://doi.org/10.1002/9781118786352.wbieg0849>. _eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118786352.wbieg0849>.
- Srikant, R., Agrawal, R., 1996. Mining sequential patterns : Generalizations and performance improvements, in : Apers, P., Bouzeghoub, M., Gardarin, G. (Eds.), *Advances in Database Technology — EDBT '96*, Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 1–17.
- Tchakerian, E., 2008. Pastoralisme méditerranéen : état des lieux et perspectives pour huit régions méditerranéennes .
- Terrasson, G., Llaria, A., Marra, A., Voaden, S., 2016. Accelerometer based solution for precision livestock farming : geolocation enhancement and animal activity identification 138, 012004.
- Terrasson, G., Villeneuve, E., Pilnière, V., Llaria, A., 2017. Precision Livestock Farming : A Multidisciplinary Paradigm, Venice, Italy. pp. 55–59.
- Thang, T.M., Kim, J., 2011. The anomaly detection by using dbscan clustering with multiple parameters, in : 2011 International Conference on Information Science and Applications, pp. 1–5. doi :10.1109/ICISA.2011.5772437.

- Vaintrub, M.O., Levit, H., Chincarini, M., Fusaro, I., Giammarco, M., Vignola, G., 2021. Review : Precision livestock farming, automats and new technologies : possible applications in extensive dairy sheep farming. *Animal* 15, 100143. URL : <https://www.sciencedirect.com/science/article/pii/S1751731120301452>, doi :<https://doi.org/10.1016/j.animal.2020.100143>.
- Van Rijsbergen, C., 1979. *Information Retrieval*. 2ème édition ed., Butterworth-Heinemann, Londres.
- Varga, A., Molnár, Z., 2014. The role of traditional ecological knowledge in managing wood-pastures, in : *European wood-pastures in transition*. Routledge, pp. 203–220.
- Velázquez, E., Martínez, I., Getzin, S., Moloney, K.A., Wiegand, T., 2016. An evaluation of the state of spatial point pattern analysis in ecology. *Ecography* 39, 1042–1055. URL : <https://onlinelibrary.wiley.com/doi/10.1111/ecog.01579>, doi :[10.1111/ecog.01579](https://doi.org/10.1111/ecog.01579).
- Manteca i Vilanova, X., J. Smith, A., 2014. *Comportement, conduite et bien-être animal*. éditions Quae.
- di Virgilio, A., Morales, J.M., Lambertucci, S.A., Shepard, E.L., Wilson, R.P., 2018. Multi-dimensional Precision Livestock Farming : a potential toolbox for sustainable rangeland management 6, e4867.
- Wane, A., Kerven, C., Dutilly-Diane, C., Biber, J.P., Merega, J.L., Odhiambo, M.O., Behnke, R., Gura, S., 2006. *Global Review of the Economics of Pastoralism*. International Union for Conservation of Nature , 47.
- Wang, G., 2019. Machine learning for inferring animal behavior from location and movement data. *Ecological Informatics* 49, 69–76. URL : <https://www.sciencedirect.com/science/article/pii/S1574954118302036>, doi :<https://doi.org/10.1016/j.ecoinf.2018.12.002>.
- Wang, J., Han, J., Li, C., 2007. Frequent closed sequence mining without candidate maintenance. *IEEE Transactions on Knowledge and Data Engineering* 19, 1042–1056. doi :[10.1109/TKDE.2007.1043](https://doi.org/10.1109/TKDE.2007.1043).
- Wang, L., Wang, D., Bai, Y., Jiang, G., Liu, J., Huang, Y., Li, Y., 2009a. Spatial distributions of multiple plant species affect herbivore foraging selectivity. *Oikos* 119, 401–408.
- Wang, L., Wang, D., Bai, Y., Jiang, G., Liu, J., Huang, Y., Li, Y., 2009b. Spatial distributions of multiple plant species affect herbivore foraging selectivity. *Oikos* 119, 401–408.
- Wathes, C., Kristensen, H., Aerts, J.M., Berckmans, D., 2008. Is precision livestock farming an engineer's daydream or nightmare, an animal's friend or foe, and

- a farmer's panacea or pitfall? *Computers and Electronics in Agriculture* 64, 2–10. URL : <https://linkinghub.elsevier.com/retrieve/pii/S0168169908001476>, doi :10.1016/j.compag.2008.05.005.
- Wiegand, T., Moloney, A., 2014. *Handbook of Spatial Point-Pattern Analysis in Ecology*. Chapman and hall/crc press ed., USA.
- Wilson, S., Bekker, M., Johnson, P., Johnson, H., 1997. Helping and hindering user involvement — a tale of everyday design, in : *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*, ACM, Atlanta Georgia USA. pp. 178–185. URL : <https://dl.acm.org/doi/10.1145/258549.258699>, doi :10.1145/258549.258699.
- Zhang, N., Wang, M., Wang, N., 2002. Precision agriculture—a worldwide overview. *Computers and Electronics in Agriculture* 36, 113–132. URL : <https://linkinghub.elsevier.com/retrieve/pii/S0168169902000960>, doi :10.1016/S0168-1699(02)00096-0.